Oracle® Financial Services

Compliance Studio User Guide

Release 8.1.2.9.0
G13663-05
June 2025

Oracle Financial Services Compliance Studio User Guide, Release 8.1.2.9.0

G13663-05

Contents

5 Administration

6 Modeling

7 Orchestration

8 More

9 Using Compliance Studio Workspace

A Appendix

Document Control

The following table describes document control of this guide.

Table Document Control

Version Number	Revision Date	Change Log
8.1.2.9.0	April 2025	Added ER Type as **CSA_8129** in the Creating Match/Merge Ruleset and Creating Data Survival Rules sections.
8.1.2.8.0	August 2024	Added User Defined method for data survival in the Creating Data Survival Rules section. Added ER Type as **CSA_8128** in the Creating Match/Merge Ruleset and Creating Data Survival Rules sections.

Preface

Compliance Studio User Guide provides information about configuring and publishing workspace for the application.

About this Guide

This user guide is provided for Data Scientists working in financial crime, allowing them to create sandboxes for model development and management of the end-to-end model lifecycle, including deployment and scheduling. It will also be used by analysts and investigators who provide business input into the entity resolution and manual decision-making processes.

Audience

This document is intended for the system administrators and users configuring the Workspace and models in OFS Compliance Studio.

Related Resources

This section identifies additional resources to the OFS Compliance Studio. You can access additional documents from the Oracle Help Center.

Abbreviations

The following table lists the abbreviations used in this document.

Table 1 Abbreviations

Abbreviation	Meaning
BA	Business Analysts
Infodom	Information Domain
Navigation Tree Menu or Navigation Menu	Left-hand side menu
OFS AAI	Oracle Financial Services Analytical Application Infrastructure
OFSAA	Oracle Financial Services Analytical Applications
Production Infodom	Production Information Domain
Sandbox Infodom	Sandbox Information Domain
SA	System Administrator
URL	Uniform Resource Locator
UI	User Interface
BD	Behavior Detection
ECM	Enterprise Case Management

Table 1 (Cont.) Abbreviations

Abbreviation	Meaning
PGX	Parallel Graph Analytics / Graph Analytics Engine
FCGM	Financial Crime Graph Model

Documentation Accessibility

For information about Oracle's commitment to accessibility, visit the Oracle Accessibility Program website at http://www.oracle.com/pls/topic/lookup?ctx=acc&id=docacc.

Access to Oracle Support

Oracle customer access to and use of Oracle support services will be pursuant to the terms and conditions specified in their Oracle order for the applicable services.

Diversity and Inclusion

Oracle is fully committed to diversity and inclusion. Oracle respects and values having a diverse workforce that increases thought leadership and innovation. As part of our initiative to build a more inclusive culture that positively impacts our employees, customers, and partners, we are working to remove insensitive terms from our products and documentation. We are also mindful of the necessity to maintain compatibility with our customers' existing technologies and the need to ensure continuity of service as Oracle's offerings and industry standards evolve. Because of these technical constraints, our effort to remove insensitive terms is ongoing and will take time and external cooperation.

Conventions

The following text conventions are used in this document:

Convention	Meaning
boldface	Boldface type indicates graphical user interface elements associated with an action, or terms defined in text or the glossary.
italic	Italic type indicates book titles, emphasis, or placeholder variables for which you supply particular values.
`monospace`	Monospace type indicates commands within a paragraph, URLs, code in examples, text that appears on the screen, or text that you enter.

Comments and Suggestions

Please give us feedback about Oracle Applications Help and guides! You can send an e-mail to: https://support.oracle.com/portal/.

1

About OFS Compliance Studio

OFS Compliance Studio is an advanced analytics application that supercharges anti-financial crime programs for better customer due diligence, transaction monitoring, and investigations by leveraging the latest innovations in artificial intelligence, open-source technologies, and data management. It combines Oracle's Parallel Graph Analytics (PGX), Machine Learning for AML, Entity Resolution, and notebook-based code development and enables Contextual Investigations in one platform with complete and robust model management and governance functionality.

Capabilities Offered by Compliance Studio

Compliance Studio has inbuilt advanced analytics for fighting financial crime on a robust platform that also allows for management and governance of user defined models and close integration with the Oracle Financial Services Crime and Compliance Management suite of applications.

This section lists the Compliance Studio capabilities:

- **Specific Use Cases for Financial Crime**
  - Behavioral ML models and Custom rules-based scenario frameworks for identifying suspicious patterns of behaviour and generating alerts for review
  - Sanctions and AML Event Scoring for false positive prediction and disposition
  - Automated Scenario Calibration and Scenario Conversion Utility for Oracle AML Scenarios
  - Customer Risk Scoring
  - Customer Segmentation and Anomaly Detection
  - Typology Detection
- **The following specific use cases are supported by the ML Foundation for Financial Crimes**
  - Integrated with Oracle Financial Crime Application Data and readily usable across the enterprise financial crime data lake
  - Pre-engineered features and transformations to address each use case
  - Simplified APIs for each stage of the modeling lifecycle
  - Leverage the power of Graph, Supervised ML, and Unsupervised ML to build typology detection models, detect anomalies, and risk score customers or events
  - Ongoing Monitoring of Model Performance and Concept Drift
- **Entity Resolution for Detection and Investigations**
  - Entity Resolution to enhance monitoring effectiveness and provide a single customer view
  - Linking and Resolution across internal and external data to improve single entity detection and enhance investigations
  - Allows for Scenario/Model detection across internal data

  - Multi-attribute enabled with ML boosts for Name/Address models
  - Prebuilt Integrations and easily configurable for Data Sources like ICIJ, Safari, etc
- **Graphs**
  - Graph Pipeline feature allows you to view the data relationships in a graphical format.
  - Graph Analytics will give Financial Institutions the ability to monitor the data financial institutions effectively. The data is organized as nodes, relationships, and properties (property data is stored on the nodes or relationships). The results of analytics algorithms are stored as transient properties of nodes and edges in the Graph.
- **Model Management and Governance**
  - End-to-end management from model creation to model deployment
    * Data Ingestion (Oracle DB, Graph, Hive)
    * Model Development
    * Supports virtually all open source packages, interpreters, etc.
    * Process in Database or Big Data
    * Model Training
    * Model Performance Evaluation
    * Model Explainability
    * Model Tracking and Audit
    * Approval Mechanisms
    * Model Deployment
    * Scheduling
    * Ongoing Monitoring
    * Analytics of Choice
    * Choose from our proprietary models or bring your own
    * Fully embedded Graph Analytics Engine and Financial Crime Model
    * Embedded with a highly scalable in-memory Graph Analytics Engine (PGX)
    * Industry's most intuitive Graph Query Language to gain rapid insights
- **Analytics of Choice**
  - Choose from our proprietary models or bring your own
  - Fully embedded Graph Analytics Engine and Financial Crime Model
  - Embedded with a highly scalable in-memory Graph Analytics Engine (PGX)
  - Industry's most intuitive Graph Query Language to gain rapid insights
- **Integrated with Oracle Financial Crime and Compliance Applications**
  - Fully defined and sourced Financial Crime Graph Model supporting detection and investigation
  - Integration with ECM and Investigation Toolkit to provide meaningful guidance to investigators for rules-based and ML-generated alerts
  - Enterprise-ready and compatible with the underlying OFSAA framework

- Works with earlier 8.0.x releases of Oracle Financial Crime and Compliance Management Anti Money Laundering (AML), Enterprise Case Management, and Fraud applications

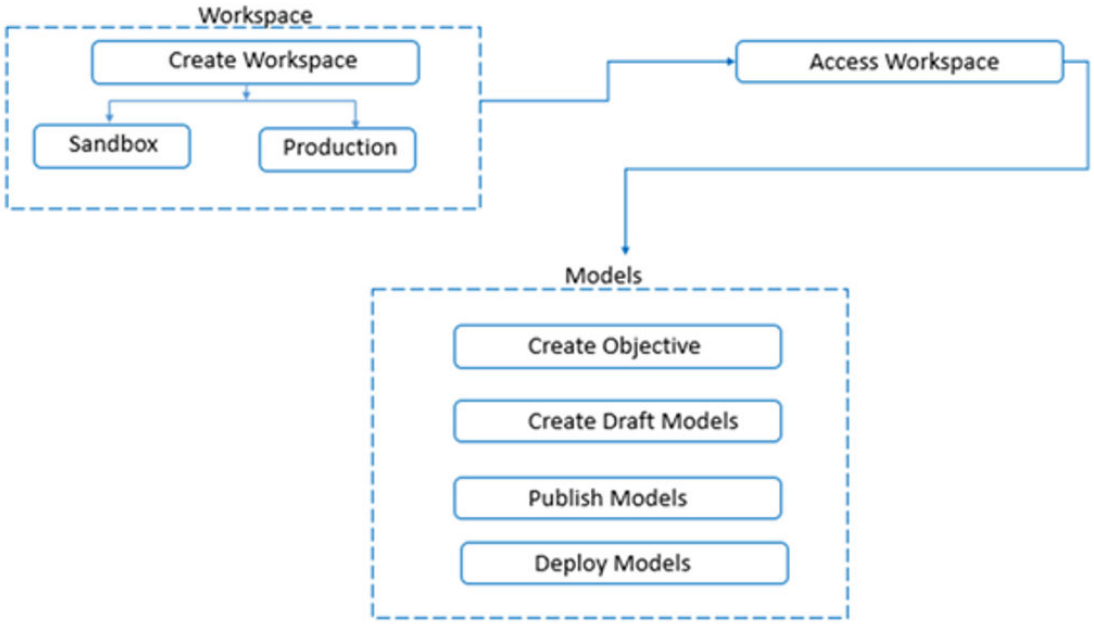
Architecture Overview

For more information on Compliance Studio architecture, see the OFS Compliance Studio Architecture Guide.

Process Flow

The process flow for building models in Compliance Studio involves the creation of Sandboxes and the creation of Models mapped to the Sandboxes. Models can be deterministic rules or trained machine learning models. You can use these models to perform model visualizations and test for the outcomes. You can then publish a model into production and make it available to users after you have determined that the models and the parameters used to construct the models meet the requirements of your business logic.

Figure 1-1 Process Flow of Compliance Studio



2
Getting Started

This section provide details on how to set up user groups and how users can access the application.

It also provides an overview of the menu items in the user interface and their purpose.

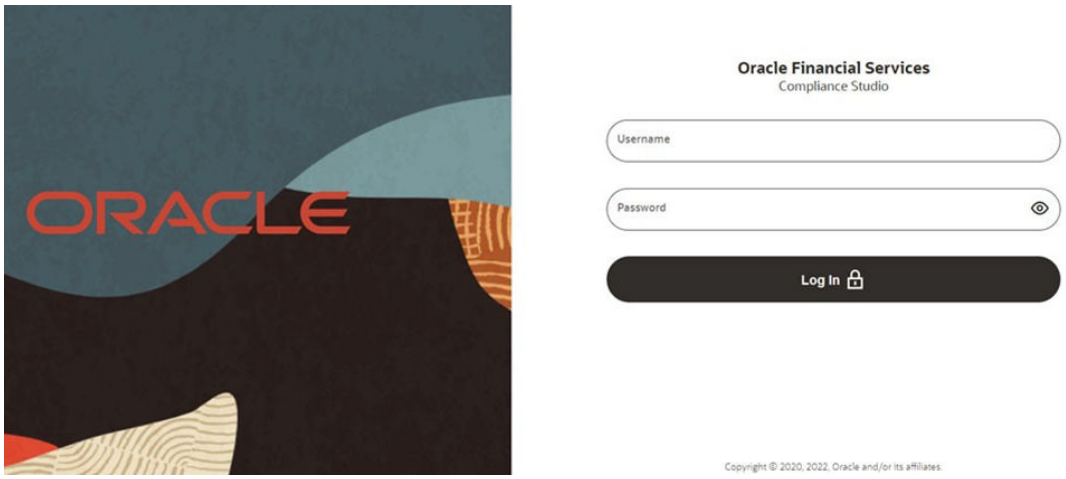
2.1 Accessing the OFS Compliance Studio Application

Users can access the OFS Compliance Studio application, once the application is installed and configured.

To access the OFS Compliance Studio, follow these steps:

1. Enter the application URL in your browser. The Login page is displayed.

Figure 2-1 Login Page for Compliance Studio



> **Note:**
>
> - The login page appears only in case of FCCMRealm.
> - In case of SAMLRealm, the user will see the login page of their Identity Provider.

2. Enter your **Username** and **Password**.

> **Note:**
>
> The Username is case-sensitive. If you are providing improper case then it will lead to access restrictions.

3. Click **Login**. The Workspace Summary page is displayed.

2.2 Mapping User Groups

This section provides information about user groups that are required to access the application.

For information on User Groups and Mappings, see the **User Access and Permissioning Management** section in the OFS Compliance Studio Administration and Configuration Guide.

> **Note:**
>
> - At the first time login, User Group mappings are initialized from AAI/ Identity Management for the newly provisioned users. These will be reflected in Compliance Studio Admin Console in the next Compliance Studio login.
> - If User Group mappings are deleted in AAI/ Identity Management, they will not delete in Compliance Studio Admin Console. Admin needs to delete this in Compliance Studio Identity screens too.
> - The user must be given **DSADMNGRP** permission to create a notebook using DSUSER login. To verify and request permission, perform the following:
>   - Login to Compliance Studio as an Administrator user.
>   - Navigate to the top-right corner and select **Admin**. The Identity Management window is displayed.
>   - For more information on permissions, see OFS Admin Console User Guide.

2.3 Using the OFS Compliance Studio Application

This section describes the following key capabilities and concepts in detail.

Workspaces

Workspace Management is where the workspace Administrators define datasets and make them available to modelers in OFS Compliance Studio. It is an environment where the data is prepared for modelers to use in their modeling creation activity.

Workspaces are provisioned with the data required for modeling by administrators, who configure workspaces with subsets of production data. The data in the Workspace is made available to modelers using datasets, who then build models with the dataset without any exposure being provided to the physical data tables and columns in the database. In effect, the data is ready for the modelers, and they do not have to undergo the arduous task of accessing and querying the database. See the Workspaces section for information on how to use Workspace Management in the application.

Dataset

Datasets allow for the creation of a dataframe, a data structure that captures the logic that organizes data into a 2-dimensional table of rows and columns. Datasets allow for the reuse of these dataframes across models, as well as the features derived from the data source(s).

The Dataset allows you to manage the entire operation related to data set. You can perform the following two things using the Dataset window:

- Define a metadata on how you want to create a dataset.
- In addition, a mechanism where you can actually take a snapshot of real time data and store it. So, it can be used later in the pipeline.
  For more information, See the Dataset section for more information.

Pipelines

Modeling refers to the process of designing a prototype based on a structured data model for statistical analysis and for simulating actual events and functions. A user with access to the Workspace can create or modify models in a workspace. Model versions are preserved in the Workspace, along with execution and output histories. Once a model has been validated in the Workspace and considered fit for use, modelers can request to push the Model into the production environment.

See the Pipelines to create models from pre-defined models to predict business trends and validate the existing models with the help of Compliance Studio.

Model Actions

The Model Actions window allows you to view the list of models associated with the Workspace.

For more information, see the Model Actions section.

Graphs (Graph Pipeline)

Unlike traditional relational database management systems, the Graph Pipeline feature allows you to view the data relationships in a graphical format.

This feature uses the latest technology to harness the power of Graph Analytics to give Financial Institutions the ability to monitor the data financial institutions effectively. The data is organized as nodes, relationships, and properties (property data is stored on the nodes or relationships). The results of analytics algorithms are stored as transient properties of nodes and edges in the Graph.

For more information, see the Graphs section.

Scheduler Service

The Scheduler Service is a service in the Infrastructure system that automates behind-the-scenes work that is necessary to sustain various enterprise applications and functionalities. This automation helps the applications to control unattended background jobs program execution.

The Scheduler Service contains a graphical user interface and a single point of control for the definition and monitoring of background executions.

For more information, see the Scheduler Service section.

Audit Trail

The Audit Trail window provides the complete details of model. This shows the information such as, when Model was created, who created the Model, workflow of Model, for example when this Model became champion or deployed, and so on.

For more information, see the Audit Trial section.

Data Studio Options

Compliance Studio contains an underlying Notebook Server which has the following configurable options:

- Interpreters
- Tasks
- Permissions
- Credentials
- Templates

For more information, see the Data Studio Options section.

Object Migration

Object Migration is the process of migrating or moving objects between environments.

You may want to migrate objects for several reasons such as managing global deployments on multiple environments or creating multiple environments so that you can separate the development, testing, and production processes.

For more information, see the Object Migration section.

Model Libraries

The Model Library information is used to bind a particular technique and its details to one unique Model Library. The Starting point for the Model builder is to register a model library with application after it is properly set up in the python environment to be used.

For more information, see the Model Libraries section.

Model Techniques

Model Technique is the algorithm/technique used to create python model using the library/ package which was created in the Model Library Screen. It is the actual information captured in the application that helps in training the model (Upload and Build).

For more information, see the Model Techniques section.

Model Catalog

The Model Catalog page allows you to add and manage the Model Technique, Model Library, and Model Objectives. You can either import the models from external sources or create, train and label it as champion model.

For more information, see the Model Catalog section.

Ruleset Details

The Ruleset facilitates identifying the similarity between two entities (customer, account, etc) and derives a match. A Ruleset is a set of rules applied to the defined source and target entities and compares the entities' attributes to derive a match.

For more information, see the Creating Rulesets section.

Manual Decisioning

The Manual Decisioning in Compliance Studio allows you to make manual decisions on similarity edges from the threshold set on the Match Ruleset page. This flexibility provides an Ad-hoc experience to improve the graph analytic capability and identify the match to analyze risk and take required actions.

For more information, see the Using Manual Decisioning section.

Merge and Split Global Entities

Merge and Split Global Entities in Compliance Studio allows you to manually break, merge, create, and re-arrange parties into global parties. This flexibility provides an Ad-hoc experience to enable users to make decisions on borderline cases or where data issues or rules logic has grouped records incorrectly.

For more information, see the Using Merge and Split Global Entities section.

ML4AML Model Templates

The prepackaged Model template allows you to perform the following:

- Customer Risk Scoring
  - Model segmentation (model grouping)
  - Load and Preview data
  - User-defined transformation (deriving additional features)
  - EDA
  - Feature selections
  - Model training and evaluation
  - Model Agnostics (Explainability)
  - On-going validations
- Customer Segmentation
  - Anomaly detection (within Cluster)
  - Deviation detection (across Cluster)
- AML Event Scoring
  - Train and Score ECM events
  - Classify events into Low, Medium and High
- Typology Scenarios
  - Shell Account detection
- Automated Scenario Calibration
  - Prod scenario analysis
    - Scenario Execution
    - Above the Line (ATL) Analysis
    - Below the Line (BTL) Analysis
    - Outcome/Impact Analysis
    - Transaction Analysis

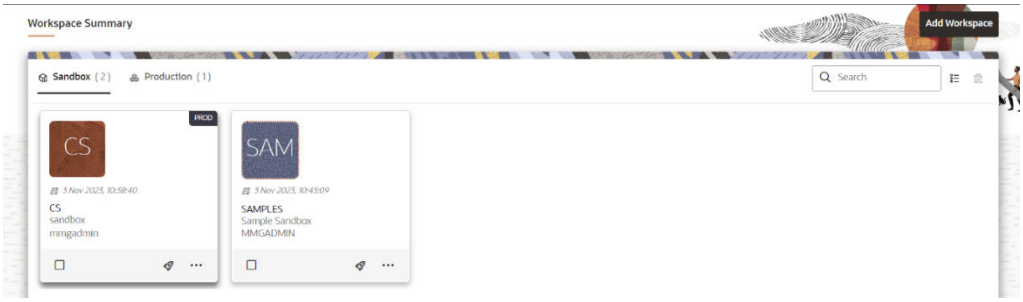
- Pre-prod Scenario Analysis
  - Scenario Execution
  - Pre prod Analysis
  - Transaction Analysis

For more information, see the OFS Compliance Studio ML4AML Use Case Guide.

3
User Interface

Access all features from the home page. Familiarize yourself with the menus, icons, and their location to quickly complete your tasks.

Figure 3-1 Workspace Summary page



The following table provides descriptions for the fields and icons on the Workspace Summary page.

Table 3-1 Fields and Icons of the Workspace Summary page

Field or Icon	Description
Search	The field to search for Workspace. Enter specific terms in the field for which you want to search, and press Enter on the keyboard to display the results. You can search a Workspace using Workspace Code, Owner, Creation Date, or Workspace Type fields. The search list shows the Workspace details with Workspace Code, Owner, Creation Date and time, and Type.
Card and Grid View	Click the Card View or Grid View on the top-right corner of this pane to view the workspace as a block or a table, respectively.
Delete	Click Delete to delete the Workspaces.
Workspace Code	The code of the Workspace. NOTE: Use up to 20 characters consisting of alphanumeric and space characters only. Do not use "ALL" as a workspace code.
Description	The description for the Workspace.
Owner	The owner of the Workspace.
Creation Date	The date and time on which the Workspace was created.
Data Source Type	Shows the type of the Data Source.

Table 3-1 (Cont.) Fields and Icons of the Workspace Summary page

Field or Icon	Description
Add Workspace	Click Add Workspace to create a new Workspace. See the Create a Workspace section for more information.
Launch Workspace	Click Launch next to corresponding workspace to display the Workspace Dashboard with application configuration and model creation menu. See the Launch a Workspace section for more information.
View Workspace	Click Action next to corresponding Workspace and select View to view the workspace with dataset data. See the View a Workspace section for more information.
Populate Workspace	Click Action next to corresponding Workspace and select Populate to populate the workspace with dataset data. See the Populate a Workspace section for more information.
Edit Workspace	Click Action next to corresponding Workspace and select Edit to edit the Workspace.
Delete Workspace	Click Action next to corresponding Workspace and select Delete to delete the Workspace. See the Delete a Workspace section for more information.
Download Workspace	Click Action next to corresponding Workspace and select Download to download the Workspace. See the Download a Workspace section for more information. The file is downloaded in .cfg format.

3.1 Home Page Components

The Home page contains the following sections:

- Navigation Menu
- Home
- Data Source Summary
- API Configurations
- Conda Environments
- Kafka Topics
- Health
- Notification
- Profile

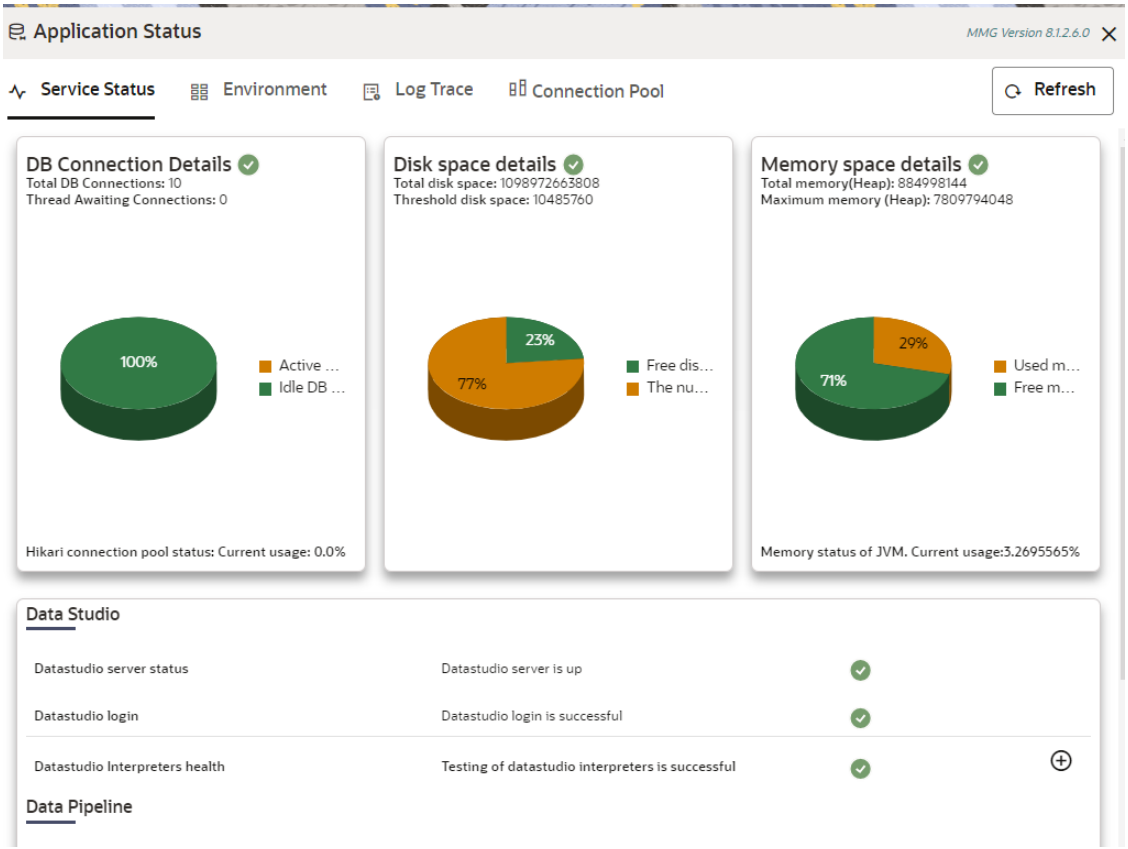
3.1.1 Home

Click **Home** to navigate to **Workspace Summary** page from any other window in the application. You can add and manage Workspaces from this page.

3.1.2 Health

Click **Health** icon to view the application status and version details. The Health page consists of four tabs:

Figure 3-2 Application Status



- **Service Status**:The Database connection, Disk space, Memory allocation details are displayed in graphical chart format. Hover on the chart to view more details. In addition, the status information of the services that are running such as Data Studio, Data Pipeline, and Graph Services are displayed. Click on the + icon to view individual services status.

  > **Note:**
  >
  > The Data Pipeline Service Health status is displayed as "Not healthy" when the response is other than "UP" status. You must be mapped to the Graph role 'GRPSUM' as the Data Pipeline Health check is present in Graph services.

- **Environment**: The System and Application properties details are displayed.
- **Log Trace**: The individual logs are displayed that will help you during analysis and debugging of the issue. You can set the last lines of the log based on your requirements. For example, if you wish to see only 10 last lines of the log, enter 10 in the Show last lines text box. By default, the files are displayed in alphabetical order and **mmg-service.log** is displayed. Click on the required log file to view specific details.

> **Note:**
>
> Access logs are available in the following location for UI and Services.
> $LOG_HOME/services/access/mmg-ui-access.log
>
> $LOG_HOME/services/access/mmg-service-access.log

- **Connection Pool**: The information of connections used for the data source are displayed. Select the data source from the drop-down list to view the connections and the maximum capacity. This will help you in monitoring the connections of the specific data source. Clicking on the **Refresh** icon will refresh the application status details.

> **Note:**
>
> Logs can be downloaded from Health UI screen, under log trace tab. User can download the zip files of individual logs, and Bulk download is also supported. The user needs to have LOGADMN role to access the download option.

3.1.3 Notification

Click **Notification** icon to display the notifications from the application. In the current release, only data population related notifications are displayed. You can delete the notifications by clicking on the **Delete** when no longer needed.

To add or modify the notification details, perform the following:

1. In the header of the Home page, click on the **Notifications** icon.

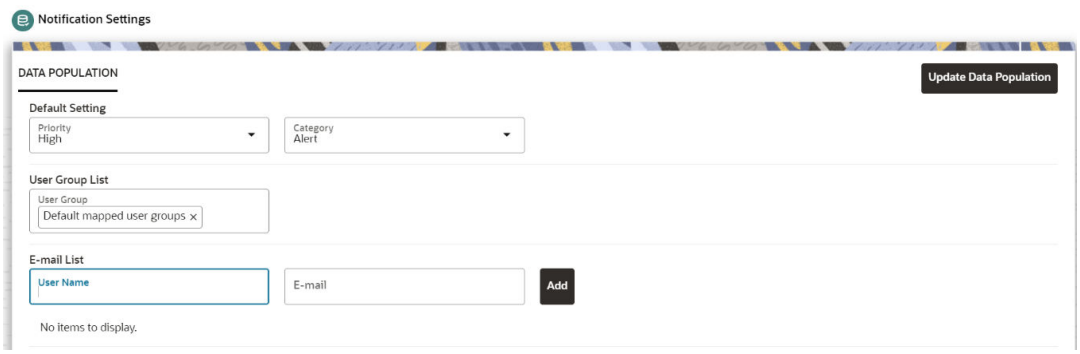
   The Notifications are displayed.

2. In the Notifications screen, click **Notification Settings** icon.

   The Notifications Settings page is displayed.

3. Select the required options from the **Notification Settings** page and click **Update Data Population**. If you select the User group, they will be notified via User Interface with the alerts under Notification icon and if you want to get the notifications in email, add the user name and email ID so that whenever a data population is triggered, an email notification is shared to the user.

Figure 3-3 Notification Settings



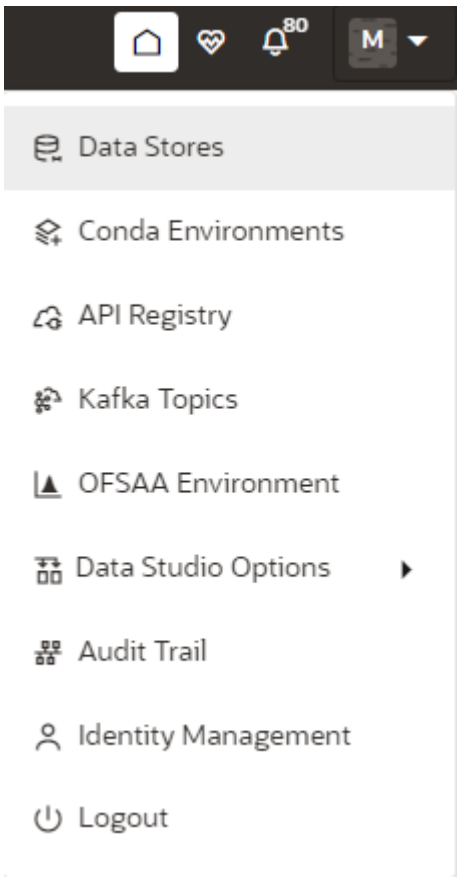
The notification alerts details are updated.

> **Note:**
>
> An email notification is sent to the default email id added in AAICL_SMS_USER_PROFILE present in the Config Schema.

3.1.4 Profile

The Header displays icons, buttons, and text for generic information and access to the OFSAA application's features. The following user-interface elements are displayed.

Figure 3-4 User Interface Elements



- **User Name**: Displays the first letter of the logged-in user name. Hover over to view the complete name of the user. Click this to select from the following options in the drop-down list:
  - **Data Stores**: Click to navigate to Data Store Summary page. You can add and manage data stores from this page.
  - **Conda Environments**: Click to navigate to Environment Summary page. You can configure and manage conda environments from this page.
  - **API Registry:** Click to navigate to API Summary page. You can configure and manage APIs from this page.

- **Kafka Topics**: Click to navigate to Kafka Topics page. You can configure and manage kafka topics and clusters from this page.
- **OFSAA Environment**: Click to navigate to OFSAA Environment page. You can register and OFSAA environments (Simulation and Production) from this page.
- **Data Studio Options**: Contains an underlying Notebook Server which has the following configurable options:
  * Interpreters
  * Tasks
  * Permissions
  * Credentials
  * Templates

  Clicking on the option will navigate you to the specific page. For more details, see Data Studio Options section.
- **Admin**: Navigates to IDCS page.
- **Audit Trail**: The Audit Trail window provides the information such as start and stop of UI and Service, add or delete of datasource, API Registry, Kafka topics, Conda Environments and so on are displayed. Click on **Audit Timeline View** to display the sequence of actions performed is listed. In addition, the Audit Table view (graphical representation) can also be viewed. You can search the model by entering the name in the search box.
- **Identity Management**: Click to navigate to IDCS page. You can manage users, groups, roles, and functions from this page.
- **Logout**: Select to log out of the application.

3.2 Mega Menu Components

To view the Mega Menu components, launch a Workspace. The Dashobaord page is displayed with menu options such as Modeling, Orchestration, More, and other sections of the Application.

3.2.1 Modeling

The Modeling component contains the following features:

- **Datasets**: Explore data, engineer features and monitor data drifts.
- **Model Libraries**: Manage python libraries for model creation and usage.
- **Model Techniques**: Manage techniques and algorithms from registered model libraries.
- **Model Catalog**: Centralized hub for all your machine learning models.
- **Pipelines**: Design, deploy and govern end-to-end machine learning pipelines for seamless model development.
- **Graphs**: Analyze complex relationships within your data through interactive graph representations.

For more details, see Modeling section.

3.2.2 Orchestration

The Orchestration component contains the following features:

- **Scheduler Dashboard**: Get an overview of scheduled tasks and processes.
- **Define Batch**: Manage and configure batch definitions.
- **Define Tasks**: Create tasks, configure parameters and set execution dependencies within a batch process.
- **Schedule Batch**: Set execution schedules for your batch processes.
- **Monitor Batch**: Track and monitor batch process executions.

For more details, see Orchestration section.

3.2.3 More

The More component contains the following features:

- **Export Objects**: Utilize Object Migration feature to efficiently export and save objects.
- **Import Objects**: Import previously exported objects into the system.
- **Model Actions**: Review and approve model deployments in bulk.
- **Audit Trail**: Monitor all system activities and changes.
- **Data Pipelines**: Create and manage data pipelines for efficient data flow and processing.

For more details, see More section.

4
Workspaces

Use the Dashboard to create and manage workspaces. By default, a Sample workspace is created which you can provision and use it to understand the workspace workflow.

Workspaces can be in the production environment (deployed), or they can exist in a separate instance on their own (on local for testing purposes) with a copy of data that comes from the desired data source (production or external data source).

You can view the following details on Workspace Summary page:

- Number of Sandbox Workspaces
- Number of Production Workspaces
- File Manager
  File manager is a ftpshare home path. It has option to Upload a new file., create a new folder, preview the file, delete and download options.

> **Note:**
>
> For File Manager to be visible on Workspace Summary, File Read, File Write, and File Delete roles must be mapped.

The Workspace Dashboard allows you to view the models of the launched workspace. To access the Dashboard window, follow these steps:

1. Navigate to Workspace Summary page.
   The page displays workspace records in a tile format.
2. Click next to corresponding Workspace to Launch Workspace.
   The **Dashboard** window is displayed with application configuration and model creation menu.

   The Workspace Dashboard shows the following details of a launched workspace:

   - Mega Menu
   - Recently Used
   - Most Used Tags
   - Models Status
   - Job Status
   - Models Timeline

4.1 Create a Workspace

The Workspace creation requires entry of the source of dataset, validation, and deployment. Besides, the application may require users of different function groups to create and approve a Workspace. In other words, a user associated with the modeler function group creates a Workspace and the approval and deployment are done by a user associated with the Modeling Administrator function group.

UGDOMMAP function should be mapped to the user performing sandbox creation operation. Otherwise, the create operation will fail.

Two new objects are added during data sourcing in workspace creation:

- Materialized View
- Triggers

Create workspace process can be time-consuming, typically may take between 2 to 3 hours. However, if certain scenarios are not carefully considered, the creation may encounter failures, leading to unwanted delays. To prevent such setbacks and ensure a seamless creation, it is crucial to address these potential issues beforehand.

The possible issue and solution are as follows:

- **Adding Views, Materialized Views, or Triggers without selecting their respective parent tables**
  - **Issue**: If a user adds views, materialized views, or triggers without confirming the selection of their corresponding parent tables, it can lead to the workspace creation failure.
  - **Solution**: It is essential to verify that all dependent objects, such as Views, Materialized views, and Triggers, have their corresponding parent tables included in the selection before initiating the creation process.

To create a Workspace, follow these steps:

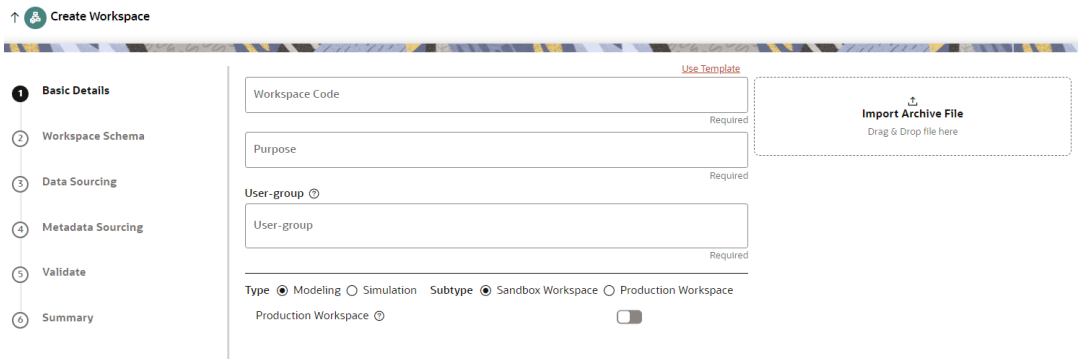
1. Navigate to **Workspace Summary**page.

   The page displays workspace records in a table.

2. Click **Add Workspace**.

   The **Create Workspace** page is displayed.

Figure 4-1 Create Workspace



The window displays a progress indicator at the left that indicates the active window where you are entering details. Click **Previous** to go back a step and click **Next** to go to the next step.

The following steps show the various phases from workspace creation to deployment:

a. Basic Details

b. Workspace Schema

c. Data Sourcing

d. Metadata Sourcing

e. Validate

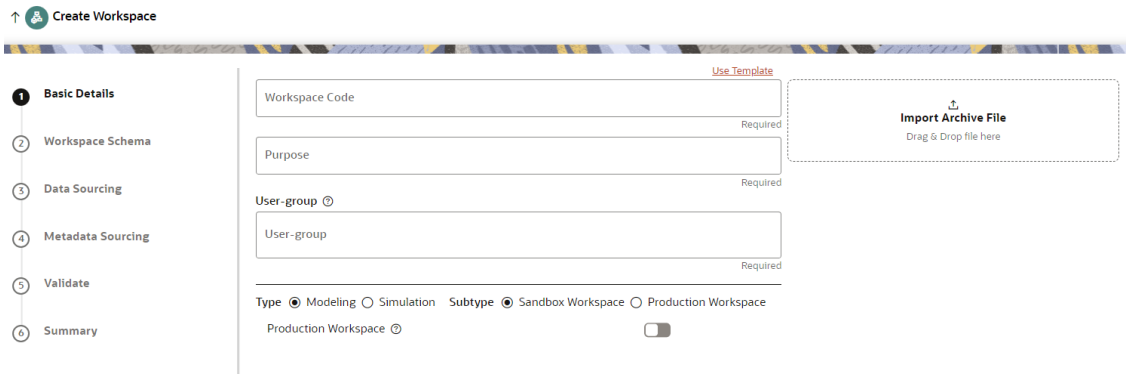
f. Summary

> **Note:**
>
> During the Workspace creation, simulation segment code parameter is added. Simulation segment code can be provided as anything. It should be 10Char Name and it can be same as the workspace name.

4.1.1 Configure Basic Details

Enter basic configuration details in this window.

Figure 4-2 Basic Details



To configure the basic details for the workspace, follow these steps:

1. Enter the required details in the **Basic Details** pane as shown in the following table:

Table 4-1 Details for Basic Details pane

Field	Description
Use Template	Click on this field to create a workspace from a template. The zip files stored on the path: <Installation path> /scratch/ofsaadb/ftpshare/mmg/seeded/workspace-templates will be displayed in the **Library** drop-down. On selecting the template, any pre-filled values will be overridden with the template provided values.

Table 4-1 (Cont.) Details for Basic Details pane

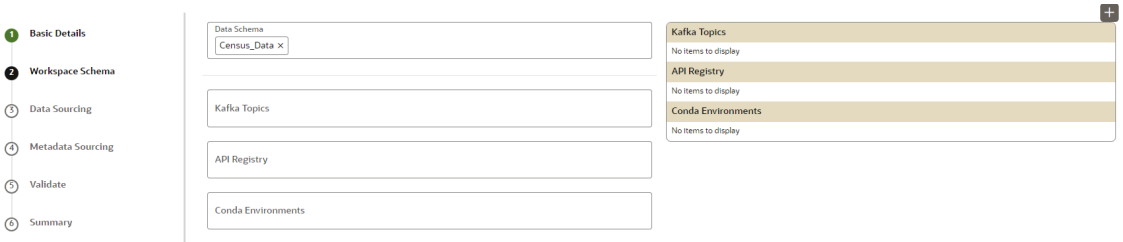
Field	Description
Workspace Code	Enter the code of the workspace. **Note:** Use up to 20 characters consisting of alphanumeric and space characters only. Do not use "ALL" as a workspace code.
Purpose	Enter the purpose of the creation of the Workspace.
User-group	Click on this field to display a list of User-group values. Select the required value. For Example: Modeling Approver, Modeling Reviewer, and Modeling User.
Type	Select the type of Workspace as Modeling or Simulation. NOTE: If Simulation is selected, fields to add the simulation data are displayed. The simulation option is displayed only when the user is mapped to **SIMULATIONACCESS** role. Select the subtype of Workspace as Sandbox or Production.
Production Workspace	NOTE: This option is not displayed when the **Type** field is selected as Production. Move the toggle switch to the right to enable this option. Enabling it attaches the production schema, which selects the Workspace being created for automatic model promotion. Based on the selection of the Source Workspace, the model is promoted to the environment.
Import archive file	Enter the archived file to import for basic details. If you use this feature, the other fields described in the preceding rows are auto populated. Click on the box to open the file selector dialog and select the required configuration file or drag the file from its directory and drop it in the box.

2. Click **Next** to go to the next step.

4.1.2 Configure Workspace Schema

Enter schema operation and enter connection details in this window.

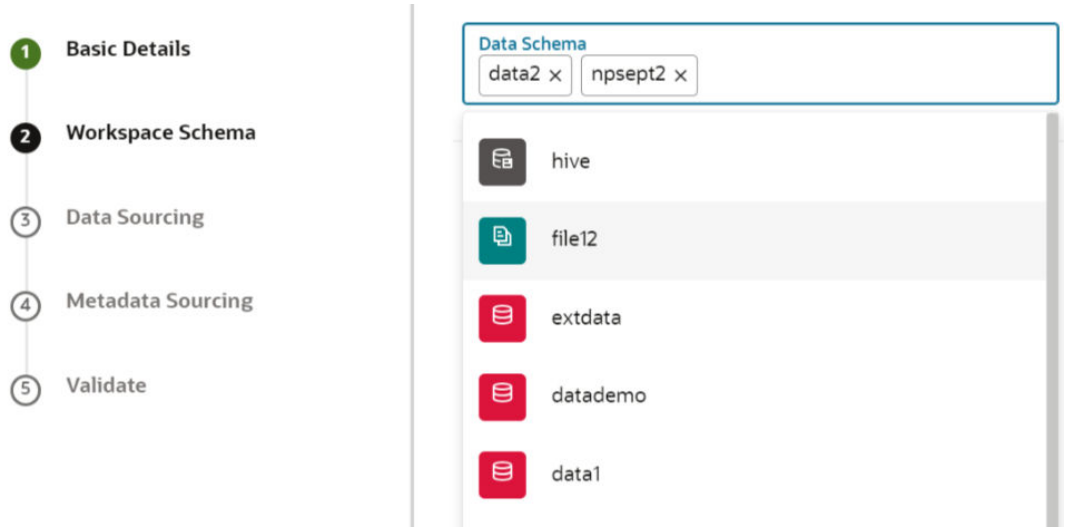
Figure 4-3 Workspace Schema



To configure the Workspace Schema:

1. Select the required schemas from the **Data Schema** drop-down list. This is the actual data used for model building. You can use multiple Data Schemas for one Meta Source. You can select an existing Schema or add a new Schema by clicking on + icon. The data source options such as Oracle, Hive, and File are displayed.

Figure 4-4 Data Schema



> **Note:**
>
> For details on adding a new data store, see Data Store section.

2. Select the required options from the **Kafka Topics** list. The selected Kafka topics are listed in the right side menu with the order of selection. For more details, see Kafka Topics.
3. Select the required options from the **API Registry** list. The selected API Registry are listed in the right side menu with the order of selection. For more details, see API Registry.
4. Select the required options from the **Conda Environments** list. For more details, see Conda Environments.
5. Click **Next** to go to the next step.

   OR

   Click **Skip** to skip the step.

4.1.3 Configure Data Sourcing

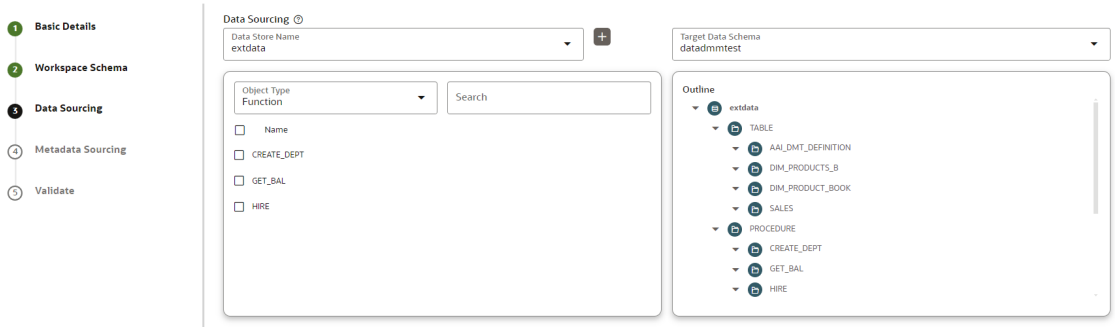
The schema type selected in the previous step requires the definition of database objects to be used for model creation. Data Sourcing step of Workspace provisioning allows to select tables from Oracle, Hive, or File based data sources from which data has to pulled into the Oracle based Workspace Data Schemas. However, unlike the data sourcing from RDBMS data sources, the tables will not get physicalized in the target schema and hence it is expected that the tables with compatible structures are already present in target RDBMS Schema. You can also select DMM operations such as Procedures, Functions, Sequences, and Package while configuring Data Sourcing. Once a workspace has been provisioned using DMM, it is stored in ftpshare path- ftpshare/dmm/DATE.

In case any of the selected tables are not present in the target schema, those tables are included in the failed objects count in workspace provisioning summary.

This window shows the different icons for Oracle, File, and Hive data sources.

Enter the details in this window.

Figure 4-5 Data Sourcing - External Data Source



To configure Data Sourcing, follow these steps:

1. You can select Data Source from Data Source Name drop-down list or create a new Data Source. To create a new Data Source, see the Configure Workspace Schema section.
2. Select the Target Data Schema. You can select multiple Data Sources for a Target Data Schema.
3. For example, if there are D1, D2 and D3 Data Sources, then you can select the tables from all these Data Sources, tables from two Data Sources, tables from one Data Source, or as required. Here, multiple combination of tables are possible with Data Source and Target Data Source.
4. If two Data Sources are having same tables (from different Data Sources), then the columns from the first selected table will be used. For example:
5. If table A has columns C1, C2, C3 and Table B has columns C1, C2, and C4,then the data from the first table will be used.
6. During the data population, only columns C1 and C2 will be used and those will be marked in Green color.
7. Select the type of objects to be displayed in the pane that follows the drop-down list. The Object Type drop-down list will be enabled after selecting the Data Source from Data Source Name drop-down list. The following are the options in the drop-down list:
   - Table
   - View
   - Synonym
   - Function
   - Procedure
   - Package
   - Sequence
8. Click **Next** to go to the next step. or Click **Skip** to skip the step.

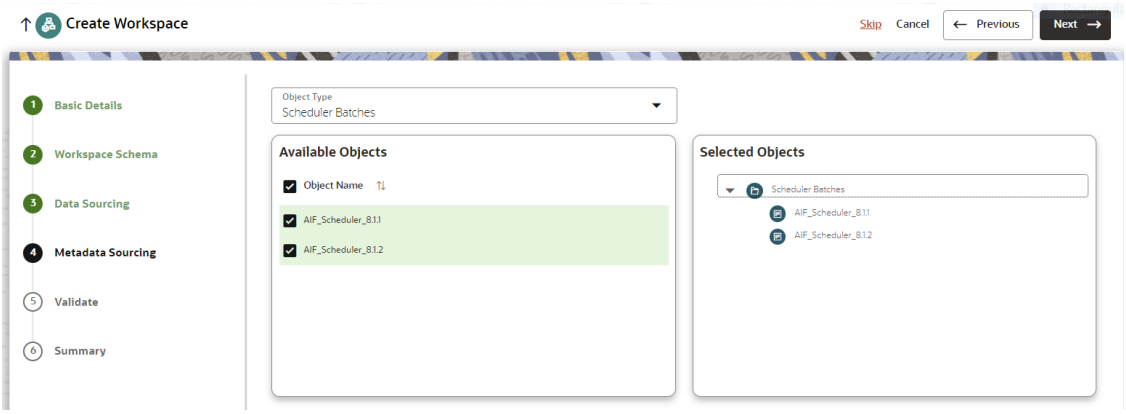
4.1.4 Configure Metadata Sourcing

The database objects selected in the previous step can be added with metadata for selected objects. Metadata Sourcing is a stage during Workspace provisioning to allow seeding of metadata like scheduler batches at the time of workspace provisioning.

Also, by default there will be seeded metadata. However, if you wish to seed the metadata, navigate to <installed path>/ftpshare/mmg/seeded/batches and drag and drop the metadata in SQL format.

> **Note:**
>
> This step is optional.

Figure 4-6 Metadata Sourcing



To configure Metadata Sourcing, follow these steps:

1. Select the required schema from the **Object Type (Optional)** drop-down list.

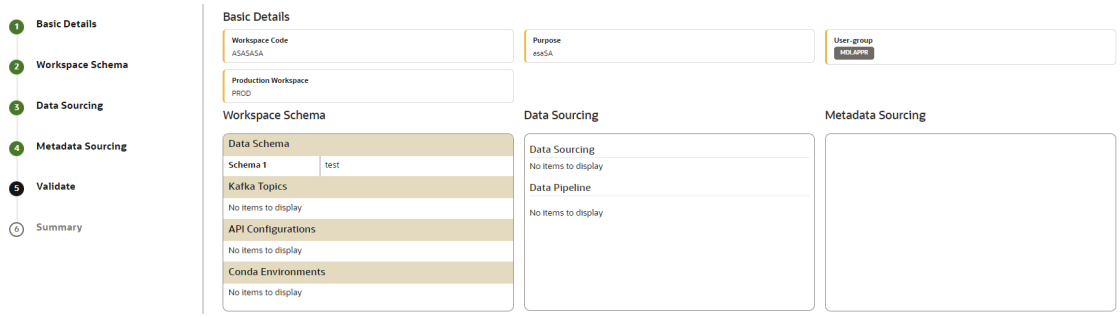
   The **Available Objects** are displayed in **Selected Objects** pane.

2. Click **Next** to go to the next step. or Click **Skip** to skip the step.

4.1.5 Validate Workspace

The **Validate** pane displays a preview of the configuration values entered in the previous panes.

Figure 4-7 Validate Workspace



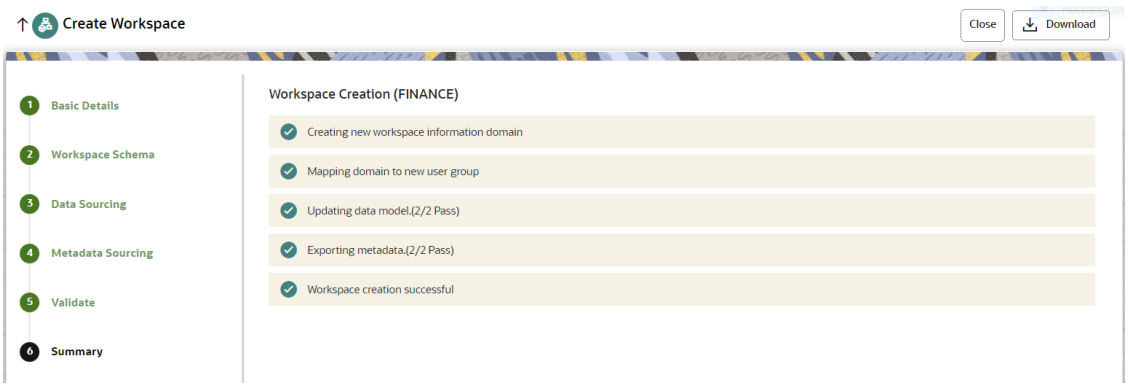
To validate the Workspace and deploy, follow these steps:

1. Review the details in the **Validate** pane. You can edit the Workspace by clicking on Edit.
2. Click **Finish** to creation of the Workspace using **Physicalize Workspace** option or **Download Configurable Archive**.
   - When you click **Download Configurable Archive**, it exports the metadata information of the workspace in .zip format which can be used later using the Import option.
   - When you click **Physicalize Workspace**, it creates actual workspace.

4.1.6 Display Summary

The **Summary** pane displays the status of the workspace creation.

Figure 4-8 Workspace Creation Summary



- Click **Download** to download the Deployment Report.

  During workspace creation, if any of the DMM operations such as Procedures, Functions, Packages, and so on fails, workspace creation will be rolled back. You can verify the details in MMG_SANDBOX_DETAILS, MMG_SANDBOX_MASTER, MMG_SANDBOX_SCHEMA tables.

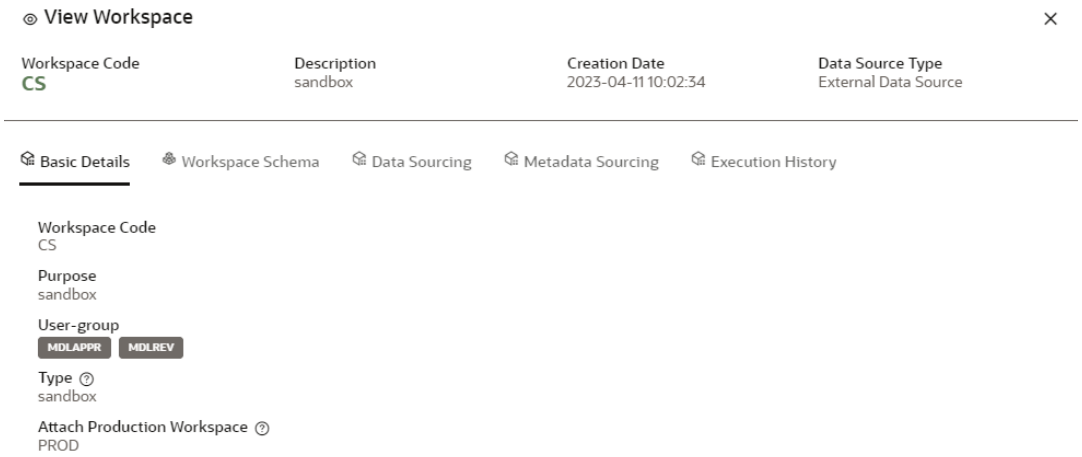
4.2 View a Workspace

To view a workspace, follow these steps:

1. Navigate to **Workspace Management** to display the **Workspace Summary** page. The page displays workspace records in a table.
2. Click **Action** next to corresponding workspace and select **View** to view the workspace.

   The Basic Details of the Workspace is displayed.

Figure 4-9 View Workspace window



3. Navigate to the Workspace Schema, Data Sourcing, Metadata Sourcing, and Execution History tab to view the respective details.

> **Note:**
>
> You can view the data population logs under the Execution History page.

4.3 Populate a Workspace

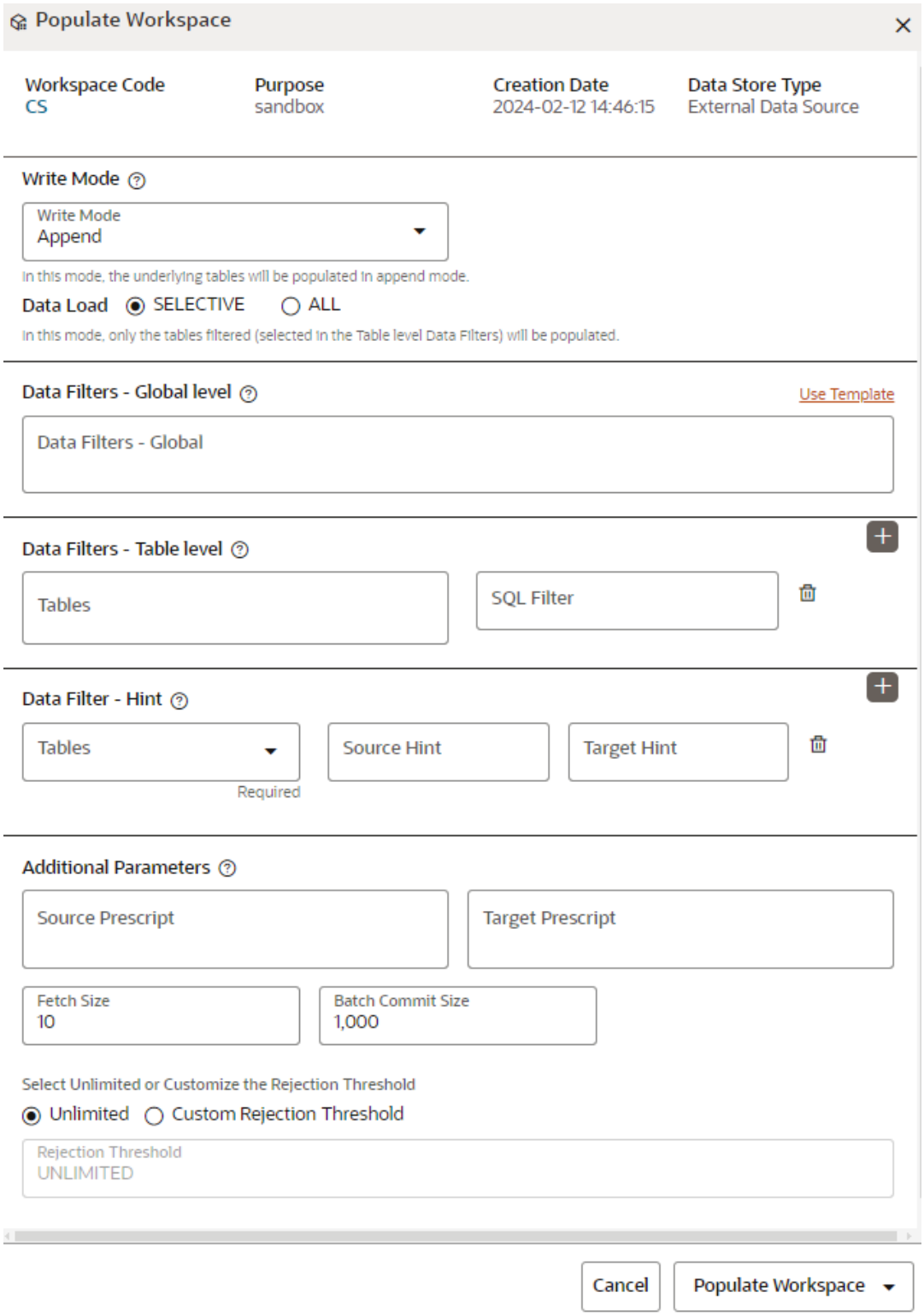
The workspace is populated with data from the datasets in External sources.

To populate the Workspace, follow these steps:

1. Navigate to the **Workspace Summary** page.

   The page displays Workspace records in a table.
2. Click Action next to corresponding Workspace and select **Populate** to populate the Workspace with data from a dataset in the **Populate Workspace** window.

Figure 4-10 Populate Workspace window



The following table provides descriptions for the fields in the Populate Workspace window.

Table 4-2 Details for the Populate Sandbox Window

Field	Description
Workspace Code	The code of the Workspace.
Purpose	The description for the Workspace.
Creation Date	The date and time on which the Workspace was created.
Data Source Type	The source of data. The value can be the OFSAA Data Schema or an external data source.
Write Mode	You can either overwrite the existing data (truncate and insert) or append to the existing data. • **Overwrite**: In this mode, the underlying tables are truncated (overwritten on existing data) followed by an insert operation. • **Append**: In this mode, the underlying tables will be populated (added to the existing data) in the append mode.
Data Load	You can select few tables or add all the tables from the source. • **Selective**: In this mode, only the tables filtered (selected in the Table level Data Filters) are populated. • **ALL**: In this mode, all the underlying tables mapped to the workspace are populated along with the filters mentioned below for specific tables.
Data Filter- Global Level	Enter the data filter that needs to be applied on all the tables selected for data sourcing. For example: If MISDATE is equal to Today, then it is applied to all tables (wherever it is available) for selected Data Sources during population. If this field is not found (MISDATE) in the tables, it is not updated. OR Click **Use Template** to select a library. On selecting the template, any pre-filled values will be overridden with the template provided values.
Data Filter- Table level	You can provide the data filters individually on the tables here. Select the table and enter the SQL filter. **Note**: You can provide multiple table names for the same SQL filter. For example, there are two tables called Student and Employee in the target data source, and below filters are applied MISDATE as Today for Student and Employee tables ID as 1 for Student table Then, Student table will be populated with MISDATE and ID filters and Employee table will be populated with only MISDATE filter. **Note**: Global filters are not applicable for those tables on which filters have been applied individually. If the same table name is provided in more than one rows here, then filter condition is generated as a conjunction of all the provided filters.

Table 4-2 (Cont.) Details for the Populate Sandbox Window

Field	Description

Data Filter - Hint	You can provide database Hints at table-level and SQL Prescripts at schema level for data load performance improvement during Workspace Population.
	Select the table and enter source and target hint filters. You can add multiple filters by clicking on Add icon.
	Note:
	You can also pass runtime values in workspace population for the user-defined parameters in **Data Filter – Global/Table**.
	Example: Table Filter /Global Filter
	`[{"id":1,"filter":"CUSTOMER_ID =$CUSTID and INCOME =$income and CUSTOMER_NAME =$customerName","tables": ["CUSTOMERS"]}]`
	Additional Parameters:
	`{"fetch_size" :10, "batch_commit_size" :1000, "rejection_threshold" :"UNLIMITED", "write_mode" :"APPEND", "dynamic_parameters": [\ {"key":"$CUSTID","value":"125"} , {"key":"$income$","value":"30000"}, {"key":"$customerName","value":"Cust125"}]}`
	The Runtime parameters can be passed as part of additional parameters json .Key = "dynamic_parameters" .It will replace all the $ values in Table filter /Global Filters.

Source Prescript	Enter the source prescript of JDBC properties for data upload.
Target Prescript	Enter the target prescript of JDBC properties for data upload.
Fetch Size	Enter the Fetch size of JDBC properties for data upload
Batch Commit Size	Enter the Batch Commit size of JDBC properties for data upload
Write Mode	You can either overwrite the existing data (truncate and insert) or to append to the existing data. You can choose to either overwrite the data or append to the existing data.
Rejection Threshold	Following two options are available: • Unlimited - Here, all the errors will be ignored during the data population. • Custom Rejection Threshold - Enter the maximum of number of inserts that may fail for any of the selected tables. You can provide the maximum number of inserts that can fail while loading data to a given table from all the sources. In case of threshold breach, all the inserts into the particular target schema will be rolled back. However, it will continue with populating the next target schema.

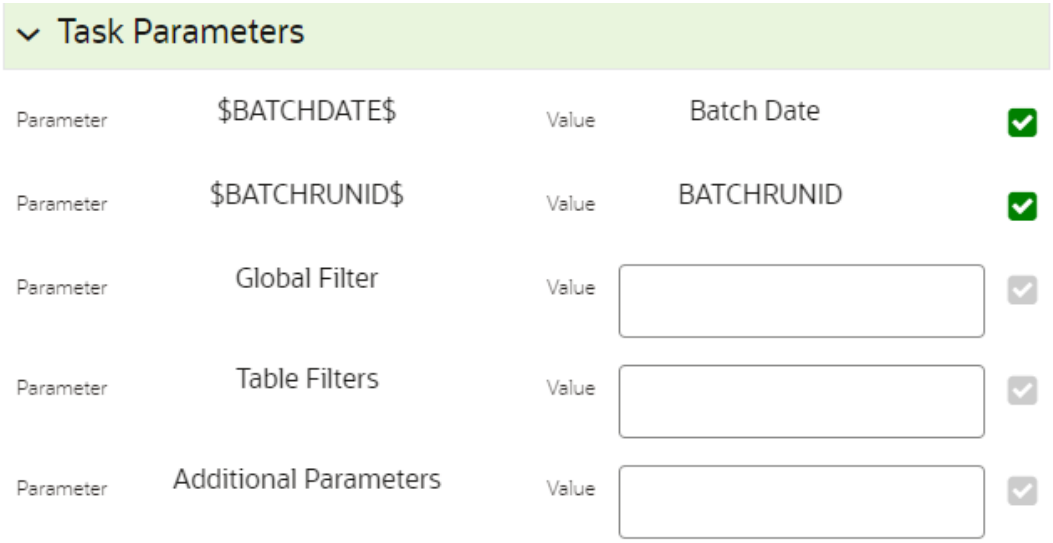
3. Click **Populate Workspace** to start the process.

   Here, you can create the batch using Create Batch, or create and execute using Create and Execute Batch option. On selecting either of these options, a workspace population task gets added to the batch.

   - When you select Create and Execute Batch option, it allows you to create batch and triggers the batch as well.
   - When you select 'Create Batch' option, it allows you to prepare the batch and then execute or schedule the batch at a later time through Scheduler Service window. The Workspace population task execution can be tracked in the 'Monitor Batch' window in the Scheduler Service.

- For a workspace population task, you have to mandatory check the logs whether the target table is getting populated. Workspace population logs can be found under: <MMG_HOME>/OFS_MMG/logs/execution/<MISDATE>/<WORKSPACENAME>/workspace-population/

Figure 4-11 Task Parameters



4. Enter the following parameters for workspace population.

   This step is optional.

   - **Additional Parameters** Enter the Additional Parameters in following format: {"fetch_size" :10, "batch_commit_size" :1000, "rejection_threshold" :"UNLIMITED", "write_mode" :"OVERWRITE"}
   - **Global Filter**
     Provided input will be applied as a data filter on all the tables selected for data sourcing.
   - **Table Filter**
     You can provide data filters individually on the tables here. You must provide multiple table names for the same SQL filter. Global Filters will not be applicable for those tables on which filters have been applied individually. In case the same table name is provided in more than one rows here, the filter condition will be generated as a conjunction of all the provided filters.

     Enter the Table filters in following format:

     [{"id":1,"filter":"<filter condition>","tables":["TABLE1",

     "TABLE2"]}, {"id":2,"filter":"<filter condition>","tables":["TABLE2"]}]

> **Note:**
>
> You can run workspace population for a given workspace any number of times. New tables may be added to the definition. Any new table added to the definition, that is not present in the target schema will be physicalized on update of the workspace. Also, user can add new sources if required.
> Any table that is deselected from the data sourcing definition will **NOT**be dropped.

4.4 Edit a Workspace

The Edit Workspace process can be time-consuming, typically may take between 2 to 3 hours. However, if certain scenarios are not carefully considered, the update may encounter failures, leading to unwanted delays. To prevent such setbacks and ensure a seamless update, it is crucial to address these potential issues beforehand.

The possible issue and solution are as follows:

- **Adding Views, Materialized Views, or Triggers without selecting their respective parent tables**
  - **Issue**: If a user adds views, materialized views, or triggers without confirming the selection of their corresponding parent tables, it can lead to the failure of the edit workspace update.
  - **Solution**: It is essential to verify that all dependent objects, such as Views, Materialized views, and Triggers, have their corresponding parent tables included in the selection before initiating the update process.
- **Manually adding additional tables in the target Workspace Schema instead of using Workspace functionality**
  - **Issue**: If a user manually creates a table (for example, ABC) in both the Source and Target schemas, independent of the workspace functionality, and then tries to map it to an existing workspace, the edit process may fail.
  - **Solution**: To resolve this issue, execute the following API with the respective headers and payload.
    - **HTTP Method**: POST
    - **URL**: http(s)://<MMG_BE_HOSTNAME>:<MMG_BE_PORT>/<MMG_CONTEXT_NAME>/dmm-service/v1/modelupload/dbcatalog
    - **Header Parameter**:
      - **ofs_workspace_id**: MMG Workspace Code.
      - **ofs_remote_user**: Login User of the application.
      - **locale**: Locale in languageCode-countryCode format. The values should be en-US.
      - **ofs_service_id**: Target Workspace Data Store Name.
      - **ofs_connection_type**: Connection type and the value shoule be OFSAA-DATA.
    - **Payload**:

* Case 1: This will retrieve all table definitions, and since the list is empty, it will fetch all tables, making the process time-consuming.

```
{
            "tableNames":[]
    }
```

* Case 2: This will retrieve only the tables specified in the request body. If you want to include specific tables, ensure they are listed.

```
{
            "tableNames":["A","B"]
    }
```

– **Response**: The response will serve as an acknowledgment that the request has been received, and it will be asynchronous due to the time-consuming nature of the process.

* **Success**:

```
{
            "status": "received",
            "errorMessage": null,
            "additionalInfo": null
    }
```

* **Error**:

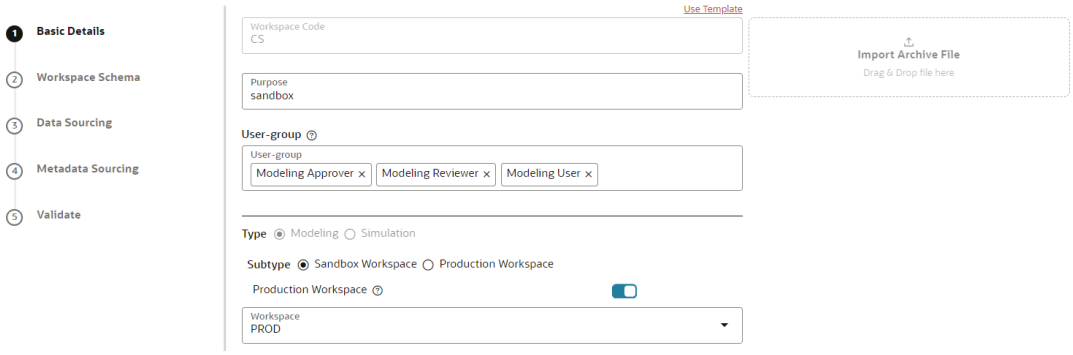
```
{
            "status": "error",
            "errorMessage": <Error Message>,
            "additionalInfo": null
    }
```

Once the execution is successful, utilize the workspace edit functionality to update the workspace by selecting the externally created table. This update is expected to proceed without issues.

To edit a workspace, follow these steps:

1. Navigate to **Workspace Management** to display the **Workspace Summary**page. The page displays Workspace records in a table.
2. Click Action next to corresponding workspace and select **Edit** to edit the workspace.

Figure 4-12 Edit Workspace window



> **Note:**
>
> You can modify the Workspace Type from Sandbox to Production or vice versa.

For more details on the Basic Details fields, see Configure Basic Details section.

> **Note:**
>
> While updating a workspace, if any of the DMM operations such as Procedures, Functions, Packages, and so on fails, DMM operations alone will be rolled back and workspace will not be rolled back. You can verify the details in MMG_SANDBOX_DETAILS, MMG_SANDBOX_MASTER, MMG_SANDBOX_SCHEMA tables.

3. Click **Save**.

> **Note:**
>
> The **Details** option has been added, replacing Edit Workspace in Action list of workspace. In the Details screen, the Edit Workspace is available.
> The background concept has been added while adding/editing workspace. Once a user clicks on update option, the confirmation message will pop up and the user will be redirected to Sandbox Summary screen where the workspace will display as loading, indicating add/edit process is going on. Meanwhile, the user can proceed with other tasks and operations, instead of waiting for the workspace to be created or edited.

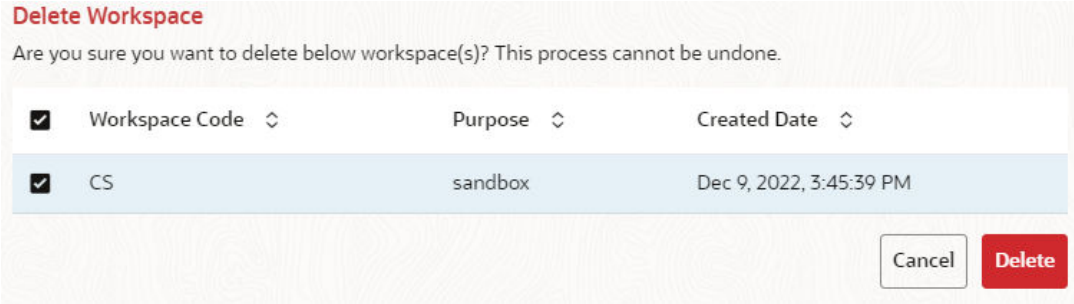
4.5 Delete a Workspace

> **Note:**
>
> When you delete a Workspace, all the underlying objects such as Dataset, Scheduler service metadata and so on from the associated tables are deleted.

To delete a Workspace, follow these steps:

1. Navigate to **Workspace Management** to display the **Workspace Summary** page. The page displays workspace records in a table.
2. Click Action next to corresponding Workspace and select **Delete** to delete the workspace.

Figure 4-13 Delete Workspace window



The following table provides descriptions for the fields in the Delete Workspace window.

Table 4-3 Details on the Delete Sandbox window

Field	Description
Workspace Code	The code of the workspace.
Purpose	The description for the workspace.
Created Date	The date on which the workspace was created.

3. Click **Delete**.
4. Click **OK** on the confirmation dialog box to confirm or click **Cancel** to cancel.

> **Note:**
>
> Ensure that you de-link the Sandbox workspaces from Production Workspace before deleting the Production Workspace.

4.6 Download a Workspace

Downloading a workspace allows you to export the workspace metadata definition (without underlying objects such as Models, Pipeline, Graphs, and so on) in a zip format. Further, the same can be used to re-create the workspace.

To download a workspace, follow these steps:

1. Navigate to **Workspace Management**to display the **Workspace Summary**page. The page displays Workspace records in a table.
2. Click Action next to corresponding workspace and select **Download**to download the workspace.

   OR

   Navigate to the Summary screen during the creation of a Workspace and click Finish and select **Download Configuration Archive**option.

> **Note:**
>
> The file is downloaded in .cfg format.

4.7 Managing Data Sources

This feature allows you to manage the Data Schemas registered with the application. The Data Source Summary window shows the list of Data schemas registered with the application. These Data schemas can be used either for workspace or for sourcing data. Click **Data Source Summary** to navigate to Data Source Summary window.

This window also allows you to manage these registered external sources.

The Data Source Summary is divided into two sections: Used Data Source and Unused Data Source.

To view the Data Source details., Click **Action** next to corresponding Workspace and select **View**

- **Used Data Source**: This shows the list of Data Sources registered with any workspace. Here, you can only view the Data Source details. The count of Used Data Sources also displayed at the top of Data Source Summary page.
- **Unused Data Source**: This shows the list of Data Sources those are not registered with any workspace. Here, you can only view, edit, or delete the Data Source details. The count of Unused Data Sources also displayed at the top of Data Source Summary page.

To add a Data Source from Workspace Summary window, follow these steps:

1. Navigate to **Workspace Summary** window.
2. Click **Data Source Summary** to navigate to **Data Source Summary** window and click **Add Datasource**
   The Add Data Source window is displayed.

To add a Data Source from Data Source Summary window, follow these steps:

1. Navigate to **Data Source Summary** window.
2. Click **Add Datasource** .

For more information, see the Create a Workspace section.

4.8 File Manager

A new feature File Manager has been added in Workspace Summary screen. For it to be visible in UI, File Read, File Write, and File Delete role should be mapped.

File manager is a ftpshare home path. It has option to Upload a new file., create a new folder, preview the file, delete and download options.

Document Gateway related feature added.

The following are the configurations added in config.sh:

- **MMG Gateway Configuration**
  - export MMG_GATEWAY_ENABLED=##MMG_GATEWAY_ENABLED##
  - export MMG_GATEWAY_PORT=##MMG_GATEWAY_PORT##

  If Gateway is enabled, the following property can be set to control the pages where MMG can be embedded:
  - Set to 'self' to allow embedding only from the same origin (recommended for most setups).
  - Set to 'all' or '*' to allow embedding from any origin. (less secure).
  - Set to a comma-separated list of origins to allow embedding from those specified origins and from the same origin. Export MMG_CSP_FRAME_ANCESTORS=##MMG_CSP_FRAME_ANCESTORS## MMG_CSP_FRAME_ANCESTORS should be configured to all or the AAI or EST origin when MMG has to be embedded from EST or AAI. By default MMG pages cannot be embedded if Gateway is enabled. This is to prevent clickjacking vulnerability.
- If the Gateway is enabled, this property can be set to control the pages where Data Studio can be embedded:
- Set to '*' to allow embedding from any origin (less secure). # Set to a comma-separated list of origins to allow embedding from those specified origins and from the same origin.
- By default, this is set to MMG Gateway URL.
- If a load balancer or an external gateway is configured for MMG Gateway, the URL must be included in the list of origins. export DATASTUDIO_CSP_FRAME_ANCESTORS=##DATASTUDIO_CSP_FRAME_ANCESTORS##

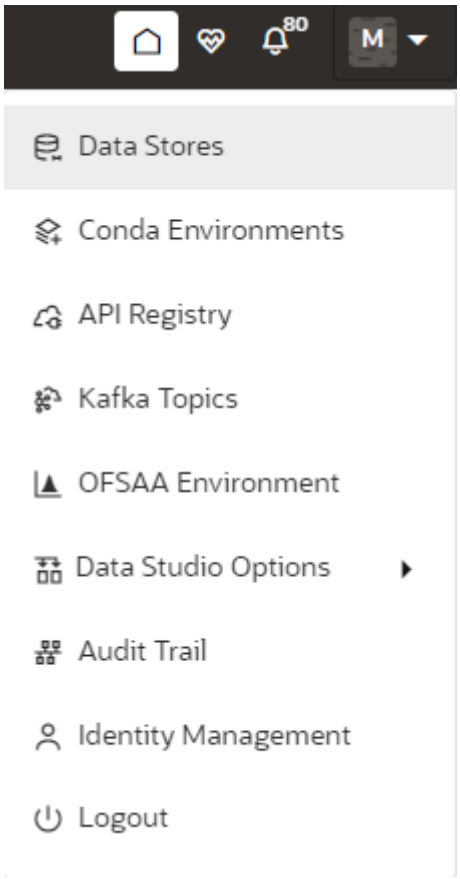
SAML changes:

- export SAML_SP_ENTITY=[https://ofss-mum-2476.snbomprshared1.gbucdsint02bom.oraclevcn.com:7001/cs]
- export SAML_SRV_URL=[https://ofss-mum-2476.snbomprshared1.gbucdsint02bom.oraclevcn.com:7001/cs/home]export SAML_SP_ENTITY=https://ofss-mum-2476.snbomprshared1.gbucdsint02bom.oraclevcn.com:<GATEWAYPORT>/cs
- export SAML_SRV_URL=https://ofss-mum-2476.snbomprshared1.gbucdsint02bom.oraclevcn.com:<GATEWAYPORT>/cs/home

5
Administration

The Header displays icons, buttons, and text for generic information and access to the OFSAA application's features. The following user-interface elements are displayed.

Figure 5-1 Administration menu



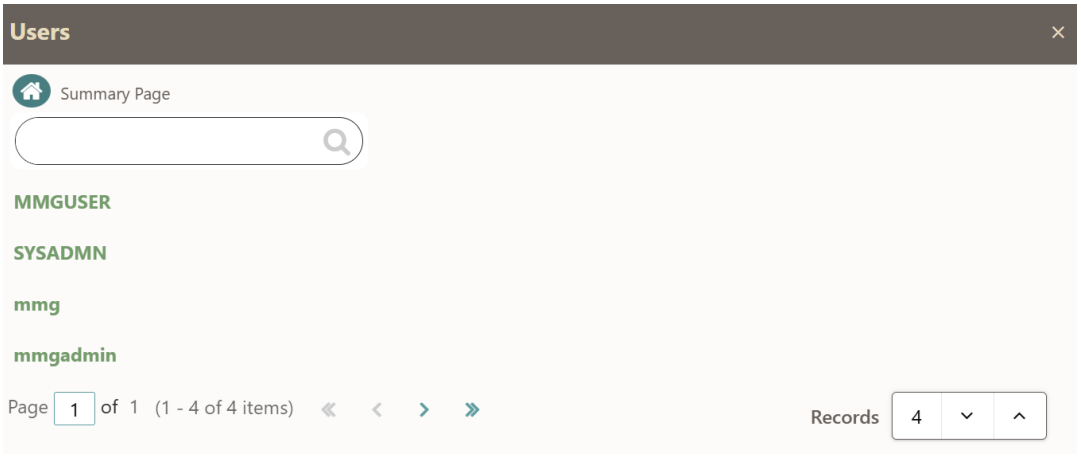
- **User Name**: Displays the first letter of the logged-in user name. Hover over to view the complete name of the user. Click on this icon to select from the following options in the drop-down list:
- **Data Stores**: Click to navigate to Data Store Summary page. You can add and manage data stores from this page.
- **Conda Environments**: Click to navigate to Environment Summary page. You can configure and manage conda environments from this page.
- **API Registry**: Click to navigate to API Summary page. You can configure and manage APIs from this page.
- **Kafka Topics**: Click to navigate to Kafka Topics page. You can configure and manage kafka topics and clusters from this page.

- **OFSAA Environment**: Click to navigate to OFSAA Environment page. You can register OFSAA environments (Simulation and Production) from this page.
- **Data Studio Options**: Contains an underlying Notebook Server which has the following configurable options:
  - Interpreters
  - Tasks
  - Permissions
  - Credentials
  - Templates

  Clicking on the option will navigate you to the specific page. For more details, see Data Studio Options section.
- **Admin**: Navigates to IDCS page.
- **Audit Trail**: Click to navigate to Audit Trail page The details such as start and stop of UI and Service, add or delete of datasource, API Registry, Kafka topics, Conda Environments and so on are displayed. You can view the sequence of actions performed in table view or timeline view (graphical representation) by clicking on Audit Table View, and Audit Timeline View options respectively. You can search the model by entering the name in the search box.
- **Identity Management**: Click to navigate to IDCS page. You can manage users, groups, roles, and functions from this page.

Figure 5-2 Identity Management Summary page

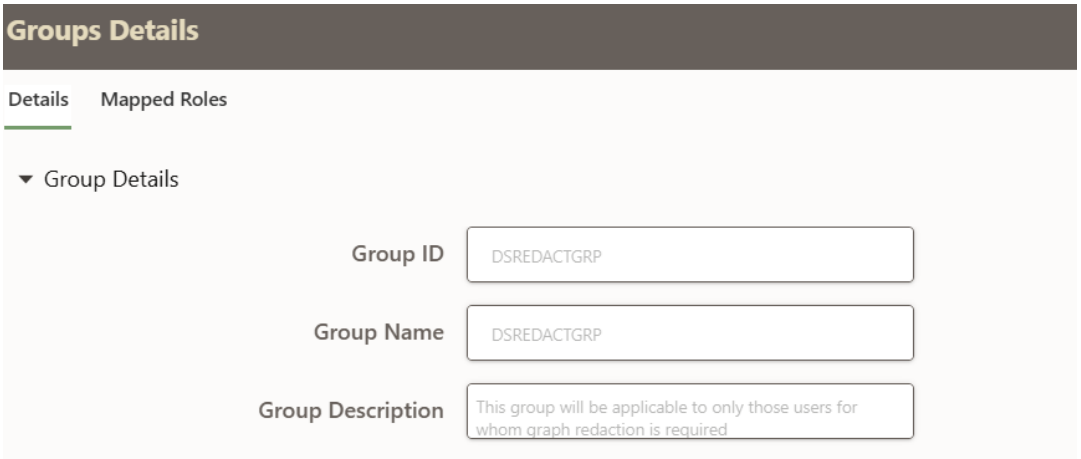
Figure 5-3 Identity Management Users



The users who are mapped to the application are listed here.

Figure 5-4 Identity Management Groups

Figure 5-5 Group Details



The MMG Admin users has groups. The above image displays the groups assigned to a particular user.

Figure 5-6 Identity Management Roles

Figure 5-7 New Mapping and Unmap

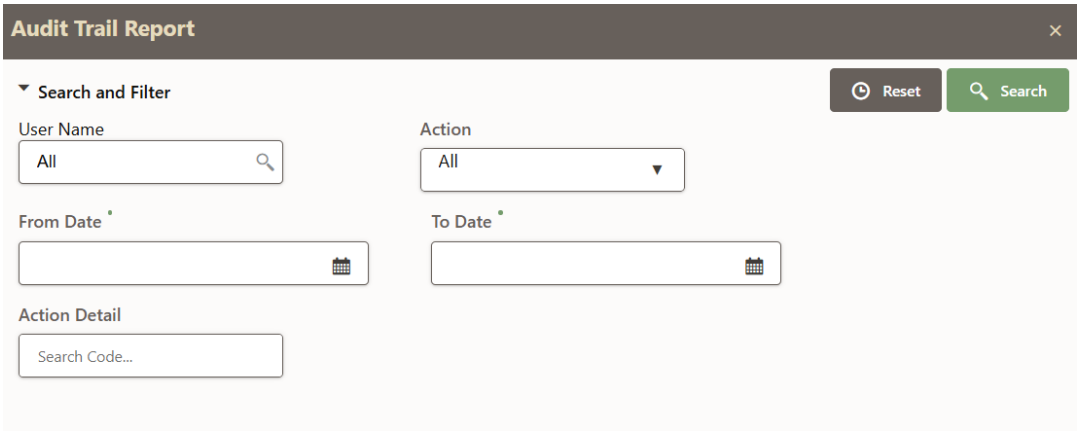
All the groups have specific roles mapped to it.

- **New Mapping:** The user can click on New Mapping for a new additional role.
- **Unmapping:** The user can remove any role by clicking unmapping.

> **Note:**
>
> However, if an individual has mapped in with a particular user credential, they can unmapp only other groups, and not the user group they have logged in with.

Figure 5-8 Identity Management Functions

Figure 5-9 Audit Trail Report



The Audit Trail Report reflects the activities the user performs during a particular period.

- **Reset:** You can click the Reset button to reset the data in the Audit Trail Report.
- **Search:** You can click the Search button to track a particular activity.

- **Logout**: Click to log out of the application.

5.1 Data Stores

This feature allows you to manage the Data Schemas registered with the application. The Data Store Summary window shows the list of Data schemas registered with the application. These Data schemas can be used either for workspace or for sourcing data. Click Data Stores to navigate to Data Source Summary window.

This window also allows you to manage these registered external sources.

Figure 5-10 Data Source Summary window

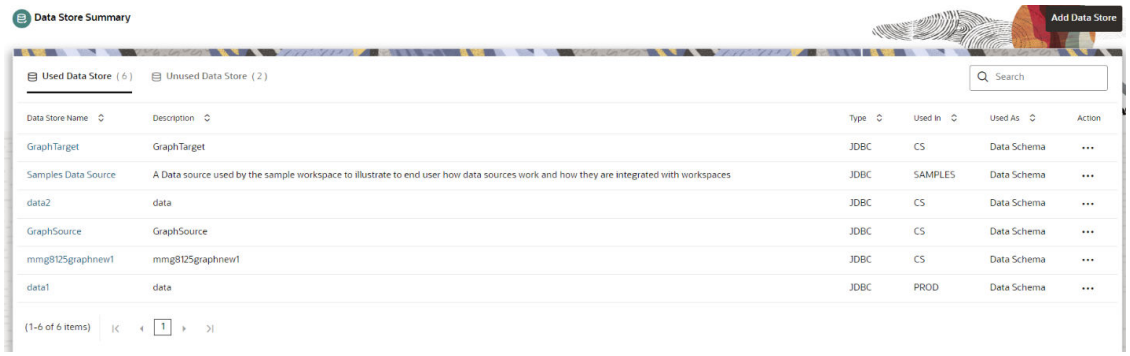
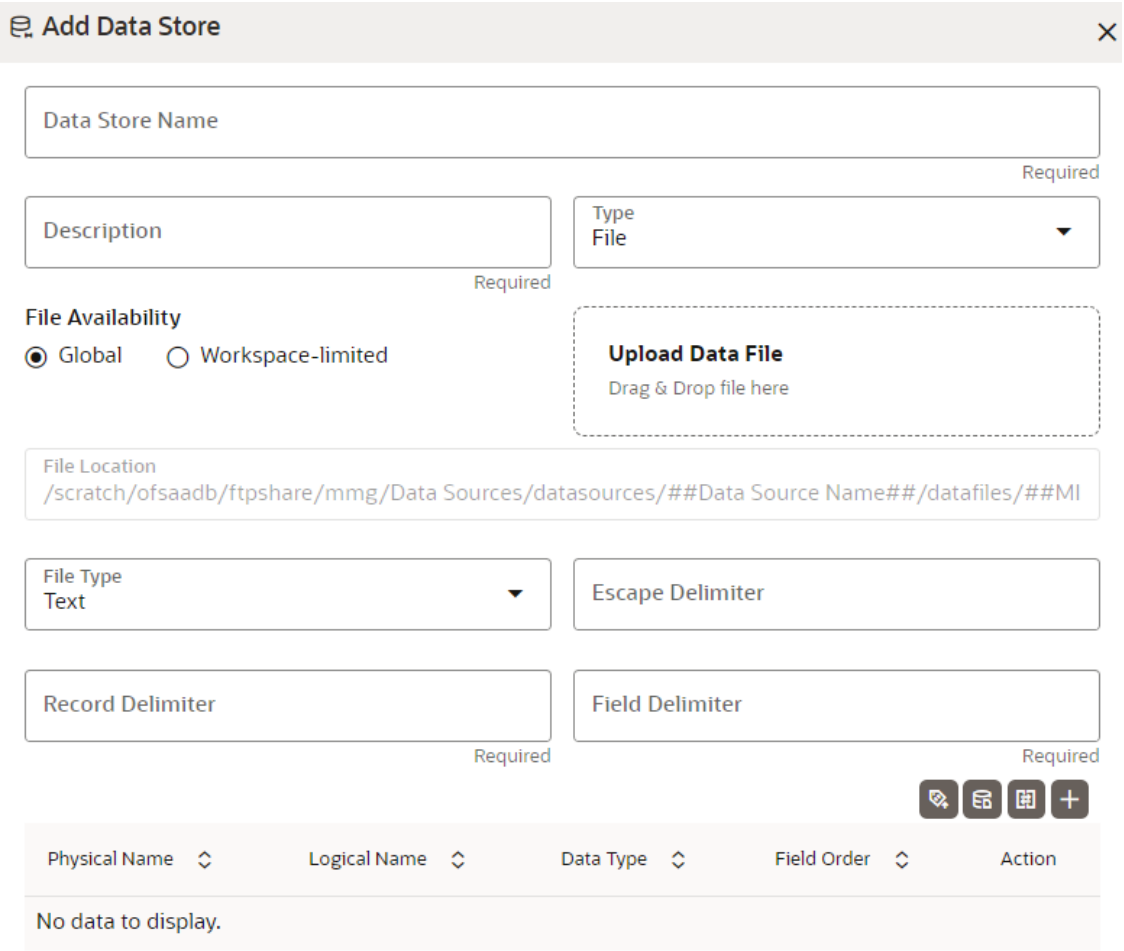


Figure 5-11 Add Data Store



The Data Store Summary consists of two sections: Used Data Source and Unused Data Source.

- **Used Data Source:**
- This shows the list of Data Sources registered with any workspace. Here, you can only view the Data Source details. The count of Used Data Sources also displayed at the top of Data Source Summary page.
- **Unused Data Source:**
- This shows the list of Data Sources those are not registered with any workspace. Here, you can only view, edit, or delete the Data Source details. The count of Unused Data Sources also displayed at the top of Data Source Summary page.

> **Note:**
>
> Data Store name should not be allowed to be created with spaces at the start and end of Data Store name and with multiple spaces in between.

5.1.1 Adding a Data Store

To add a Data store, follow these steps:

1. Navigate to **Workspace Summary** window.
2. Click **Profile** icon and select **Data Stores**.

   The **Data Store Summary** page is displayed.
3. Click **Add Store**.

   The **Add Data Store** window is displayed.

Figure 5-12 Add Data Store with Type File

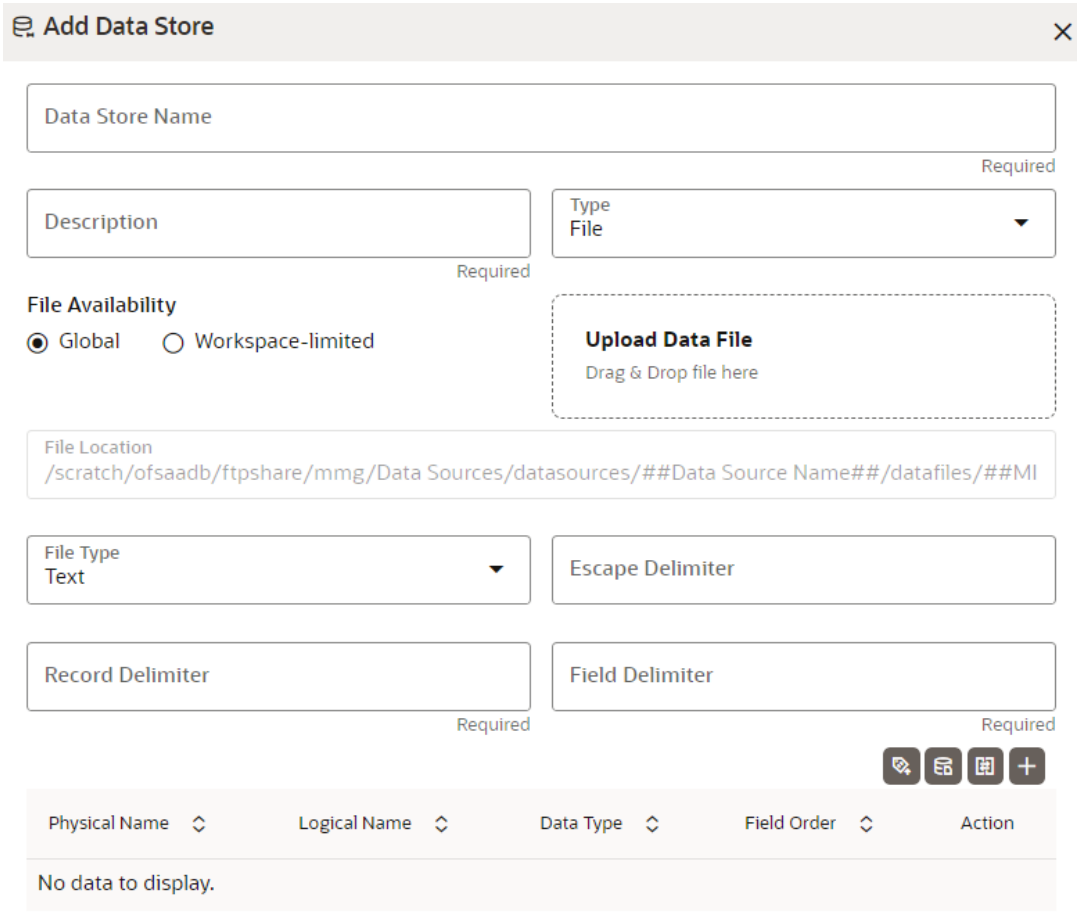


Figure 5-13 Add Data Store with Oracle Database

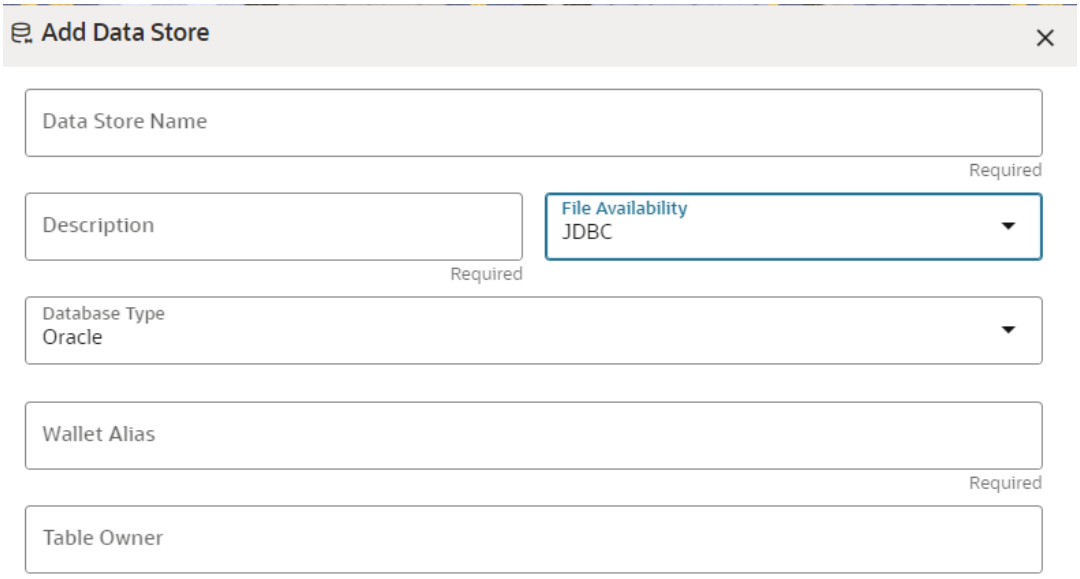
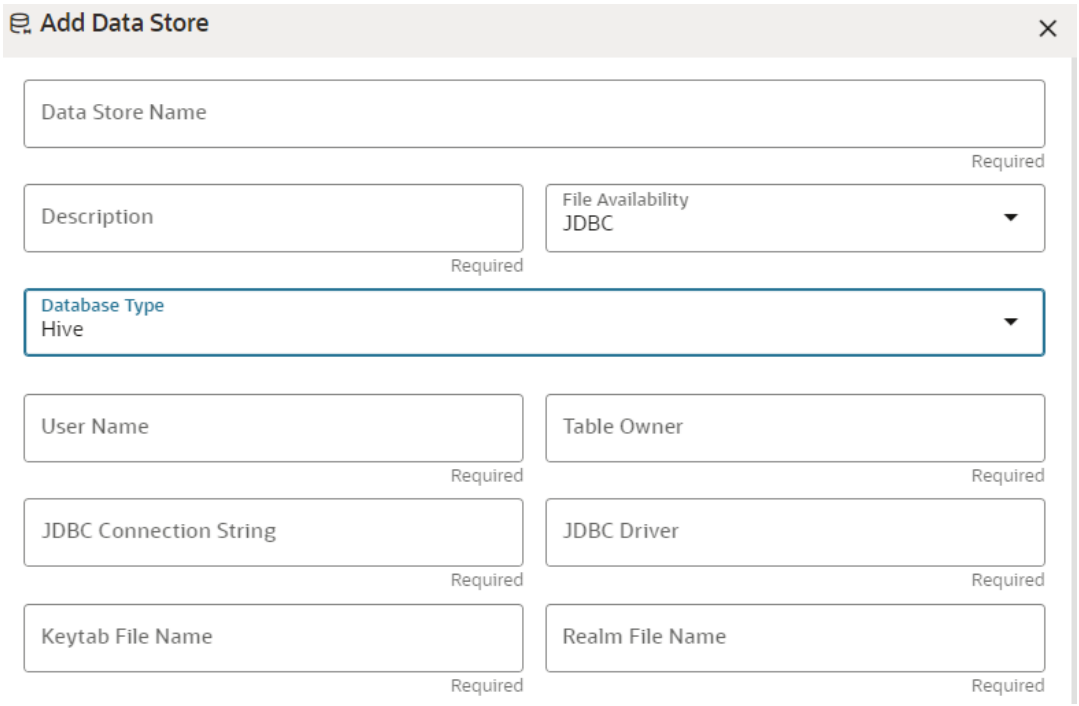


Figure 5-14 Add Data Source with Hive Database



4. Enter the required details as shown in the following table.

Table 5-1 Add Data Source

Field	Description
Data Store Name	Enter the connection URL to the database for the data schema.

Table 5-1 (Cont.) Add Data Source

Field	Description
Description	Enter the description of database connection.

Table 5-1 (Cont.) Add Data Source

Field	Description
Type	Enter the type of the database connection. The supported types are JDBC and File. • **JDBC** : If selected, the Database type options Oracle and Hive are displayed. • **File**: If selected, the following options are displayed. **Global**: Select this option if you want to fetch the global level datasource details. You need to place the file with the datasource details in JSON format in the following location: <installed path>/ftpshare/mmg/Data Sources/datasources/##Data Source Name##/datafiles/##MISDATE##/* **Workspace-limited**: Select this option if you want to fetch the workspace level datasource details. You need to place the file with the datasource details in JSON format in the following location: <installed path>/ftpshare/mmg/##Workspace##/datasources/##Data Source Name##/datafiles/##MISDATE##/* **Upload Data File**: Click to upload data file in text or csv format. Any pre-filled values will be overridden with the selected data file. Once the file is uploaded, you can derive the schema which will auto populate the values based on the file. **File Type** : Select the Type as text or csv format. **Escape Delimiter**: An escape delimiter is simply a sequence that is recognized and ignored during parsing. Its purpose is to allow the use of escape sequences to embed byte sequences in data that would otherwise be seen as delimiter occurrences. This option differentiates multiple records in one single column from second column. Escape delimiter (eg. "), added in from of multiple records will treat it as multiple records for one single column. For example, if there is a normal delimiter “+” at a given level, and you define an escape delimiter “\+” as shown in the following figure, then aaa+b\+c+ddd will parse as three fields: aaa, b\+c, and ddd. If the escape delimiter were not defined, the sequence would then parse as four fields: aaa, b\, c, and ddd. **Record Delimiter**: There is a separation of the records using a delimiter character like a comma, semicolon, hyphen, and so on for the rows. Enter the delimiter in the Record Delimiter field. This is a mandatory field and limited to two characters. **Field Delimiter**: There is a separation of the records using a delimiter character for the columns. Enter the delimiter in the Field Delimiter field. This is a mandatory field. **Derive Schema**: It is an optional feature that populates the data grid by analyzing schema details from any data file stored in the above

Table 5-1 (Cont.) Add Data Source

Field	Description
	specified directory. This is applicable when you upload data file in previous steps. You can either add the file details using data template or manually. Click Data File Template icon to select the Data Source entities and click Save. **OR** Click Add icon to add the details such as Physical Name, Logical Name, Data Type, and Field Order manually and click Save. Click Reorder Grid option to reorder the field orders.

Table 5-1 (Cont.) Add Data Source

Field	Description
Database Type	If the Type field is selected as JDBC, the following options are displayed. Select the Database Type as Oracle or Hive. NOTE: Selected tables during Hive sourcing should be preexisting in the RDBMS data schema before the workspace population. If you select Database Type as Oracle (see Figure 10), then following additional fields are displayed to enter details: • Wallet Alias : Enter the Wallet Alias. This value should be same as configured using Oracle Wallet. For more information, see the Oracle Financial Services Model Management and Governance Installation Guide • Table Owner: Enter the Oracle Database schema name. If you select Database Type as Hive (see Figure 11), then following additional fields are displayed to enter details: User Name: User Name / Principal is used for Kerberos authentication. Example: mmg/hostname@ORACLE.COM. Table Owner: Enter the Hive schema. JDBC Connection String: Enter the JDBC Connection String. Example: jdbc:hive2://hostname:10000/default;principal=hive/hive-service-hostname@ORACLE.COM. JDBC Driver: Supports org.apache.hive.jdbc.HiveDriver and com.cloudera.hive.jdbc4.HS2Driver. Keytab File Name: Enter the Name of the keytab file present in conf directory. Example: mmg.keytab Realm File Name: Enter the Name of the configuration file present in conf directory. Example: krb5.conf NOTE: • Schema population for Hive as target is not supported. • This is applicable only for Sandbox Workspace. For more information on setting the environment, see the Oracle Financial Services Model Management and Governance Installation Guide.

5. Click **Create**.

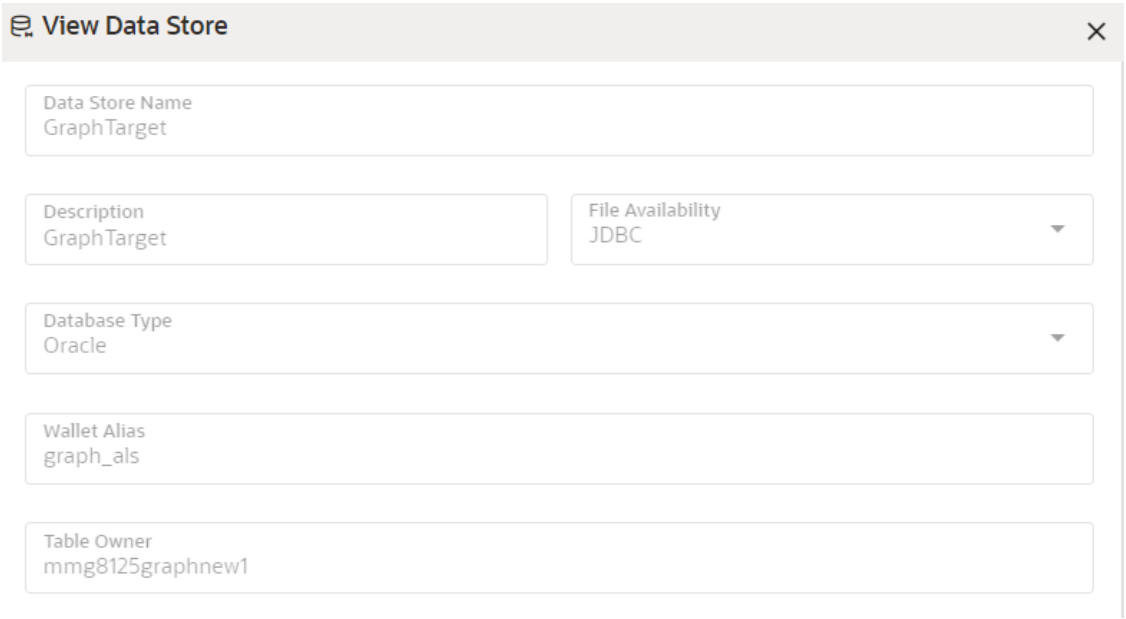
   The Data Store is created.

> **Note:**
>
> Click **Test Connection** to check the connection. A success message is displayed.

5.1.2 Viewing the Data Store Details

To view the Data Store details, click the **Action** icon next to corresponding Workspace and select **View**.

Figure 5-15 View Data Store



5.2 Conda Environments

Conda as a package manager helps you to find and install packages. With the capability of environment manager, you can set up a totally separate environment to run different versions of Python. In addition, you can continue to run your usual version of Python in your normal environment. You can configure the channels required for the Conda environment creation in the .condarc file.

> **Note:**
>
> The supported version of Conda is 3.9.x.

You must add the miniconda/bin directory in the following path variable as shown below:

PATH=/scratch/ofsaadb/miniconda3/bin:$PATH

For proxy support, you need to manually configure the .condarc file present in the root directory.

Limitation: Conda support currently does not include linking externally created environments to the application.

To configure:

Add the below in the .condarc file.

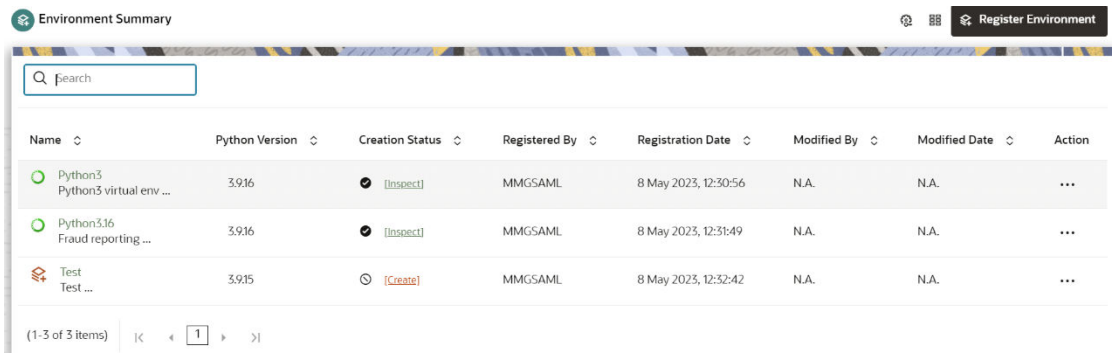
proxy_servers:

- `http: http://www-proxy.idc.oracle.com:80`
- `https: http://www-proxy.idc.oracle.com:80`

> **Note:**
>
> - Conda package search is restricted only to the package name and does not include versions.
> - Python version displayed in the Register Conda UI.

Click **Conda Environments** to navigate to **Environment Summary** page from any other window in the application. You can register and manage Conda environments from this page.

Figure 5-16 Environment Summary page



The following table provides descriptions for the fields and icons on the **Environment Summary**page.

Table 5-2 Fields and icons on the Sandbox Summary page

Field or Icon	Description
Search	The field to search for a Conda environment. Enter specific terms in the field for which you want to search, and press Enter on the keyboard to display the results. **NOTE**: Conda package search is restricted only to the package name and does not include versions.
Name	The name of the Conda environment.
Python Version	The python library version of the Conda environment.

Table 5-2 (Cont.) Fields and icons on the Sandbox Summary page

Field or Icon	Description
Creation Status	The status of the Conda environment. The status can be either Inspect or Create. • **Inspect**: The environment is registered and created. You can perform functions such as Edit/View/Inspect/Clone/Export/Remove for the Inspect status environments. • **Create**: The environment is registered but not created. You can perform functions such as Create/Edit/View/Deregister for the Create status environments
Registered By	The User Id of the User who registered the Conda environment.
Registration Date	The date on which the Conda environment has been registered.
Modified By	The ID of the Last Modified by user who has modified the Conda environment.
Modified Date	The date on which the Conda environment was modified.
Register Environment	Click to register a new Conda environment.
Conda Settings	Click to add or modify the Conda components settings.
Conda Properties	Click to view the Conda properties.
Action	Click the three dots to perform Edit/View/Inspect/Clone/Export/Remove functions on selected Conda environment. The functions might vary based on the selected Conda environment.

5.2.1 Register a Conda Environment

To register a Conda environment:

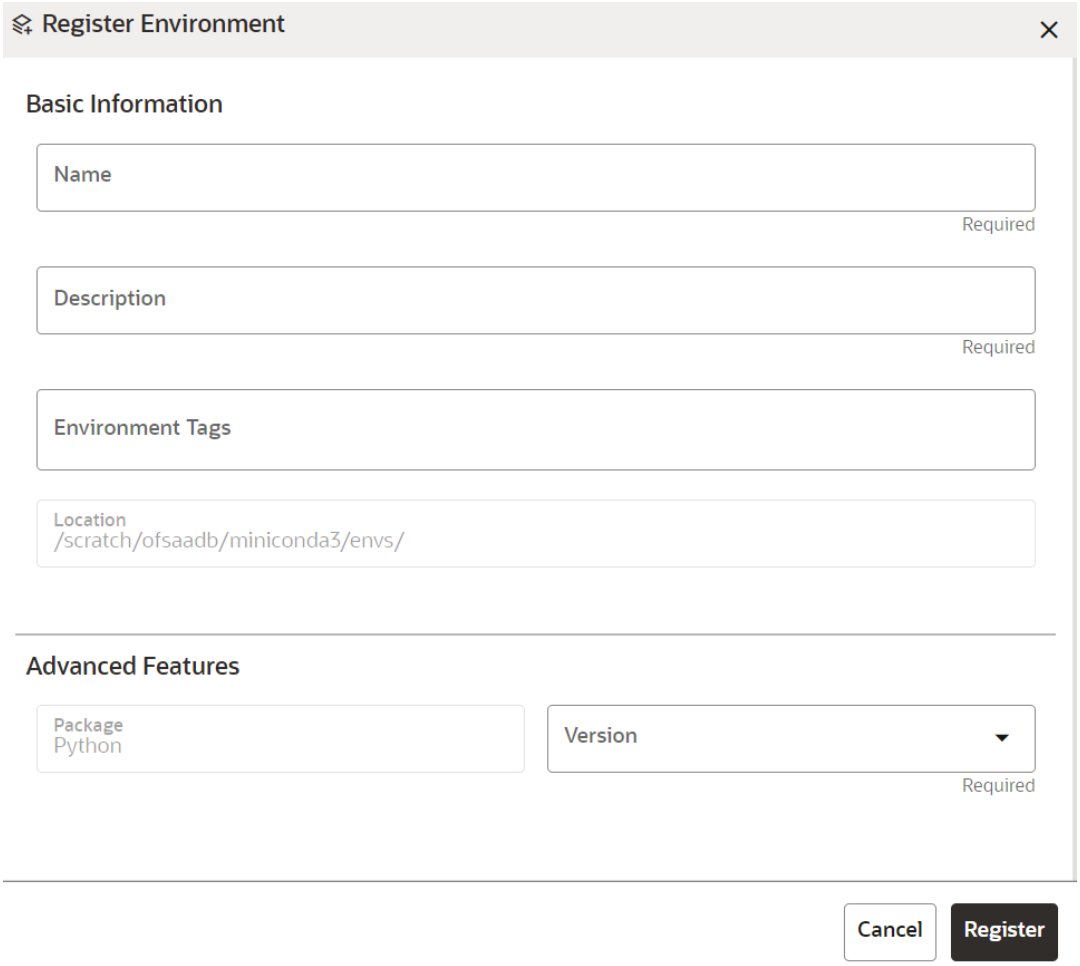
1. Click **Conda Environments**to navigate to **Environment Summary** page from any other window in the application.

   This page displays the Conda environments in a table.

2. In the **Environment Summary** screen, click **Register Environment**.

   The **Register Environment** page is displayed.

Figure 5-17 Register Environment page



3. Enter the **Name** for the environment. This field is mandatory.
4. Enter the **Description** for the environment. This field is mandatory.
5. Enter a tag in the **Environment Tags** field. You can add multiple tags.

> **Note:**
>
> The **Location** of the Conda environment is displayed. This field is greyed out and cannot be modified.

6. In the **Advanced Features** drop-down list, select the required python package for the Conda environment. Python as a package is selected by default.
7. Click **Register**.

   The Conda environment is registered and displayed in the Environment Summary page.

   After the environment registration is complete, you must create the Conda environment. For more details, see the Create a Conda Environment section.

5.2.2 Create a Conda Environment

To create a Conda environment:

1. Click **Conda Environments** to navigate to **Environment Summary** page from any other window in the application.

   This page displays the Conda environments in a table.

2. Click **Action**next to corresponding environment and select **Create** to create the Conda environment.

   > **Note:**
   >
   > The **Create** option is not displayed for the Conda environment with the creation status **Inspect**.

   A confirmation dialog is displayed.

3. Click **Create**.

   The Conda environment is created.

   These environments can be selected while creating a Workspace. For more details, see the Workspace Schema section.

   In addition, these environments can be selected during the Model pipeline execution.

5.2.3 Edit a Conda Environment

To edit a Conda environment:

1. Click **Conda Environments**to navigate to **Environment Summary**page from any other window in the application.

   This page displays the Conda environments in a table.

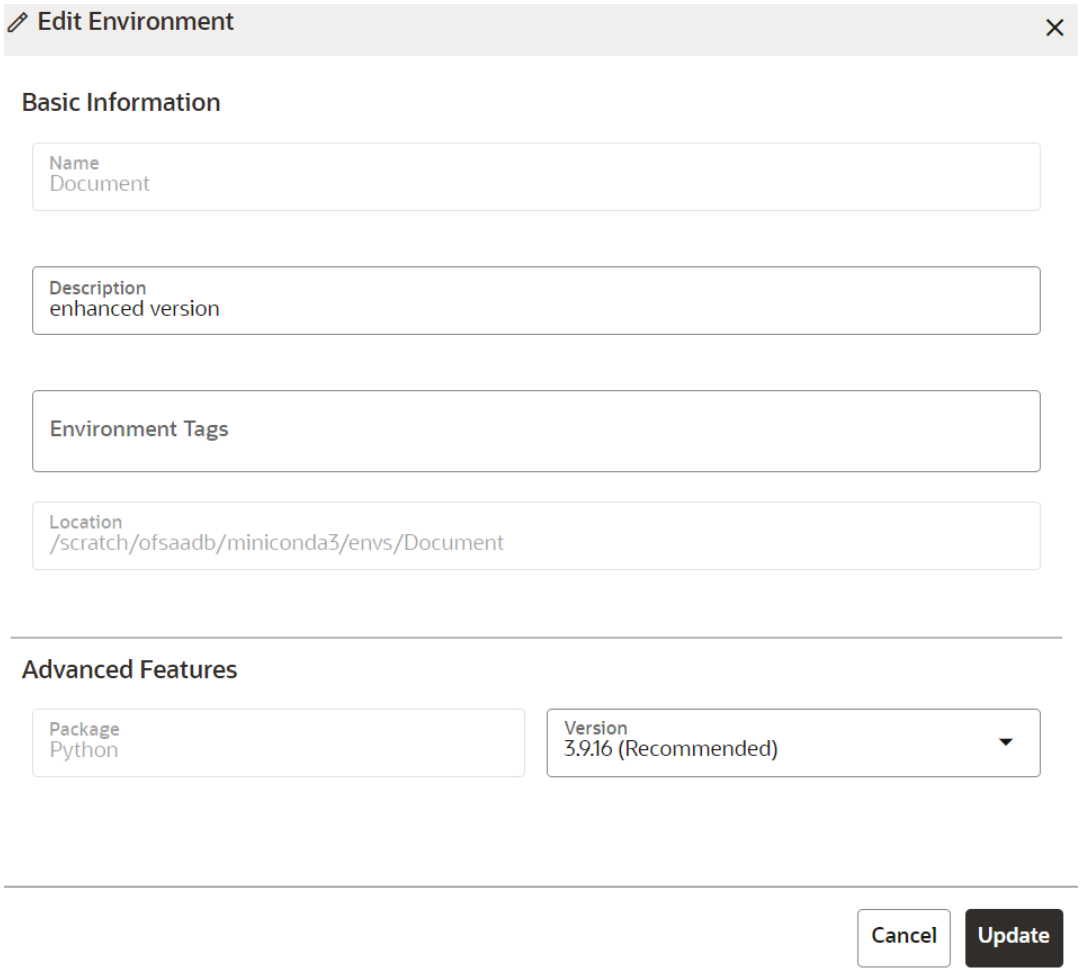
2. Click **Action**next to corresponding environment and select **Edit** to edit the details.

   The **Edit Environment** page is displayed.

   For more details on the Edit environment fields, see Register a Conda Environment section.

   > **Note:**
   >
   > The **Name**of the Conda environment is greyed out and cannot be modified.

Figure 5-18 Edit Environment page



3. Click **Update**.

   The details of the Conda environment is updated.

5.2.4 View a Conda Environment

To view a Conda environment details, follow these steps:

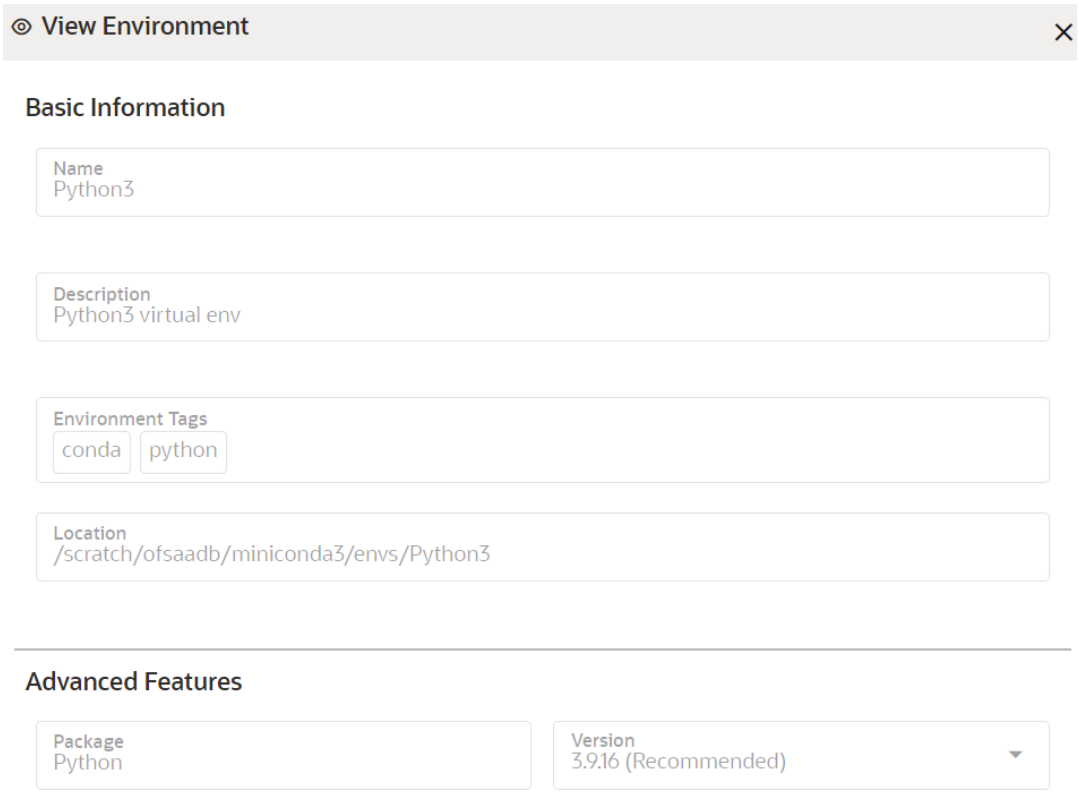
1. Click **Conda Environments** to navigate to **Environment Summary** page from any other window in the application.

   This page displays the Conda environments in a tabular format.

2. Click **Action**next to corresponding environment and select **View** to view the details.

   The details of the Conda environment are displayed.

Figure 5-19 View a Conda Environment



5.2.5 Deregister a Conda Environment

To Deregister a Conda environment:

1. Click **Conda Environments**to navigate to **Environment Summary** page from any other window in the application.

   This page displays the Conda environments in a table.

2. Click **Action**next to corresponding environment and select **Deregister** to deregister the Conda environment.

> **Note:**
>
> The **Deregister** option is not displayed for the Conda environment with the creation status **Inspect**. A confirmation dialog is displayed.

3. Click **Deregister**.

   The registered Conda environment is deregistered and removed from the Environment Summary page.

5.2.6 Inspect a Conda Environment

To inspect a Conda environment:

1. Click **Conda Environments** to navigate to **Environment Summary** page from any other window in the application.

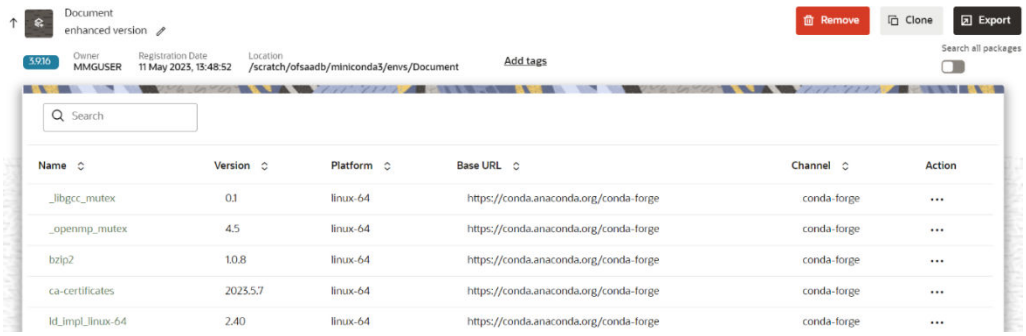
This page displays the Conda environments in a table.

2. Click **Action**next to corresponding environment and select **Inspect** to edit the details.

> **Note:**
>
> The **Inspect** option is not displayed for the Conda environment whose creation status is **Create**.

The Conda environment details are displayed.

Figure 5-20 Conda environment details



You can perform the following in the above page:

- Remove Environment
- Clone Environment
- Export Environment
- Edit a Conda Environment Description
- Add or Remove the Environment Tags
- Search All Packages
- View Package Properties
- Update Package
- Remove Package

5.2.6.1 Remove Environment

To remove a Conda environment:

1. In the Conda details page, click **Remove**.

   A confirmation message is displayed.

2. Click **Remove**.

   The Conda environment is deregistered and removed from the **Environment Summary** page.

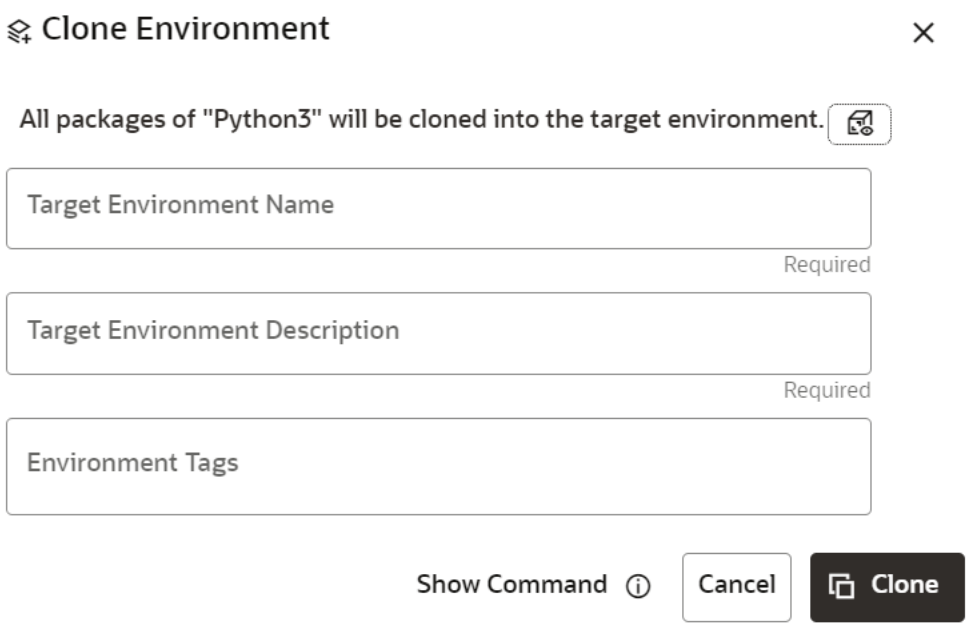
5.2.6.2 Clone Environment

To clone a Conda environment:

1. In the Conda details page, click **Clone**.

   The **Clone Environment** page is displayed.

Figure 5-21 Clone Environment



> **Note:**
>
> You can preview the packages of the current Conda environment by clicking the **Preview** icon.

2. Enter the **Target Environment Name**to which the current environment needs to be cloned.
3. Enter the **Target Environment Description**.
4. Enter a tag in the **Environment Tags**field. You can add multiple tags.
5. Click **Clone**.

   The Conda environment is cloned to the targeted environment.

5.2.6.3 Export Environment

To export a Conda environment:

- In the Conda details page, click **Export**.

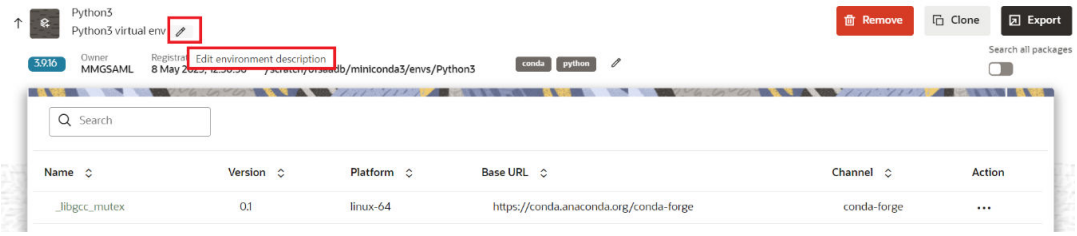
  The Conda environment is exported in **yml** format.

5.2.6.4 Edit a Conda Environment Description

To edit a Conda environment description:

1. In the Conda details page, click **Edit** icon.

Figure 5-22 Conda details page



2. Modify the description and click **Save**.

> **Note:**
>
> The **Name** of the Conda environment cannot be modified.

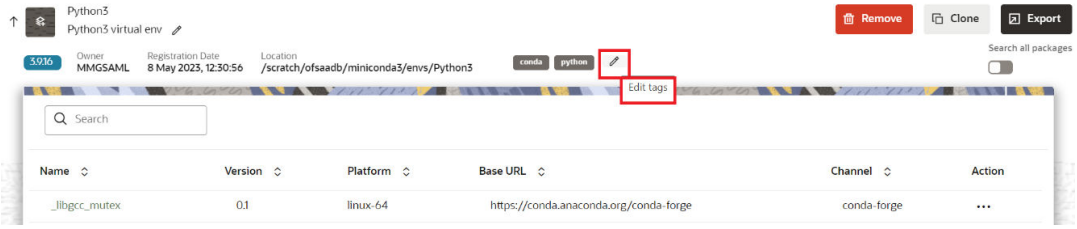
The details of the Conda environment is updated.

5.2.6.5 Add or Remove Environment Tags

To add or remove Environment Tags:

1. In the Conda details page, click **Edit**.

Figure 5-23 Conda details page



2. Add or remove the environment tags, and click **Save**.

   The environment tag details of the Conda environment are updated.

5.2.6.6 Search All Packages

To search all the packages:

1. In the Conda details page, enable the **Search all packages** option.
2. If you want to search the packages in cache, enable the Look up in cache option.

   This option is set to true by default for faster processing. The searched package is looked up in the cache and the result published, if found. If not, a new search in the conda-forge

channel is performed. You can toggle this option if you expect more recent versions to be available.

3. Enter the package name in the search bar and press **Enter**.

   All the packages meeting your search criteria are displayed.

   Click the **Action** icon to view and install packages.

5.2.6.7 View Package Properties

To view package properties of the Conda environment, on the Conda Details page, click **Action** and select **View**.

The details of the package are displayed.

Figure 5-24 Package Properties

5.2.6.8 Update Package

To update the package of the Conda environment:

1. In the Conda details page, click **Action** and select **Update**.

   A confirmation message is displayed.

2. Click **Update**.

   The selected package is updated.

5.2.6.9 Remove Package

To remove a package from the Conda environment:

1. In the Conda details page, click **Action** and click **Remove**.

   A confirmation message is displayed.

2. Click **Remove**.

   The selected package is removed.

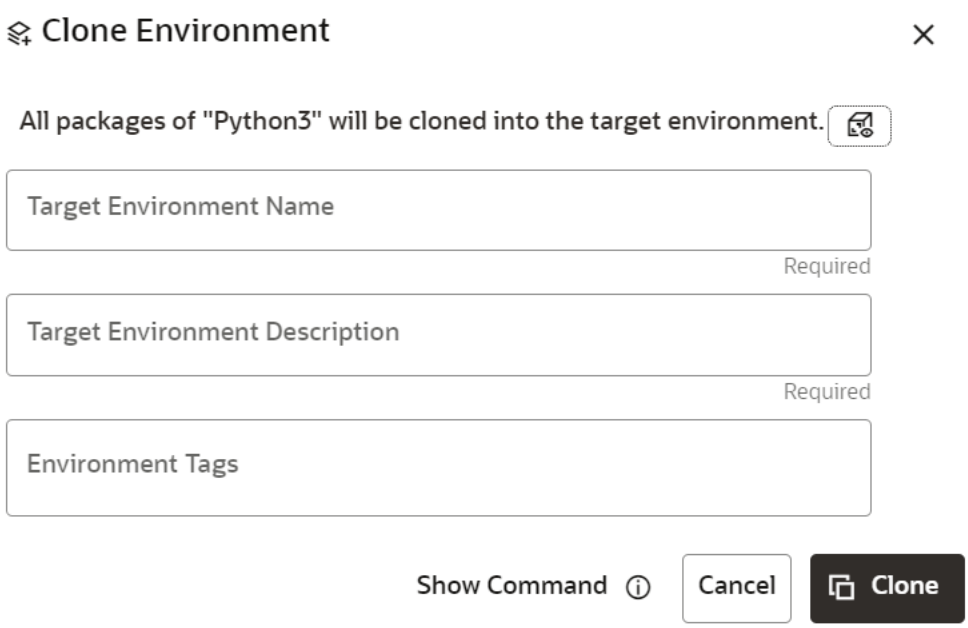
5.2.7 Clone Environment

To clone a Conda environment:

1. In the Conda details page, click **Clone**.

   The **Clone Environment** page is displayed.

Figure 5-25 Clone Environment page



> **Note:**
>
> You can preview the packages of the current Conda environment by clicking the Preview icon.

2. Enter the **Target Environment Name** to which the current environment needs to be cloned.
3. Enter the **Target Environment** Description.
4. Enter a tag in the **Environment Tags** field. You can add multiple tags.
5. Click **Clone**.

   The Conda environment is cloned to the targeted environment.

5.2.8 Export Environment

To export a Conda environment:

1. Click the **Conda Environments** icon to navigate to the **Environment Summary** page from any other window in the application.

   This page displays the Conda environments in a tabular format.

2. Click **Action** and select **Export**.

   The Conda environment is exported in **yml** format.

5.2.9 Modify Conda Settings

To modify the Conda Settings:

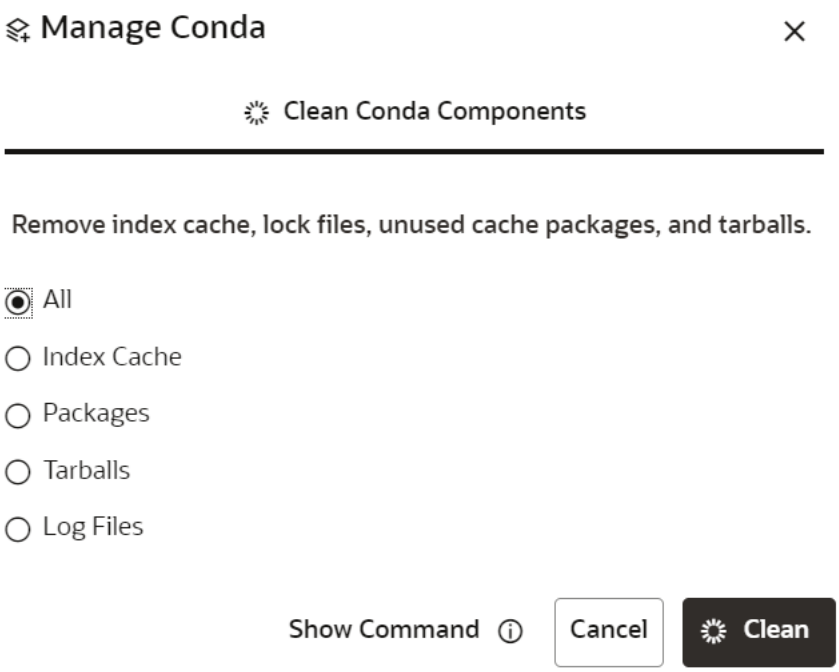
1. Click the **Conda Environments** icon to navigate to the **Environment Summary** page from any other window in the application.

   This page displays the Conda environments in a tabular format.

2. Click the **Conda Settings** icon.

   The **Manage Conda** screen is displayed.

Figure 5-26 Manage Conda screen



3. Select the required option and click **Clean**.

   - **All**: Removes index cache, lock files, unused cache packages, and tarballs.
   - **Index Cache**: Removes index cache.
   - **Packages**: Removes unused packages from writable package caches. WARNING: This does not check for packages installed using symlinks back to the package cache.
   - **Tarballs**: Removes cached package tarballs.
   - **Log Files**: Removes log files.

   This will clean up the Conda components based on the selection.

5.2.10 View Conda Properties

To view Conda Properties:

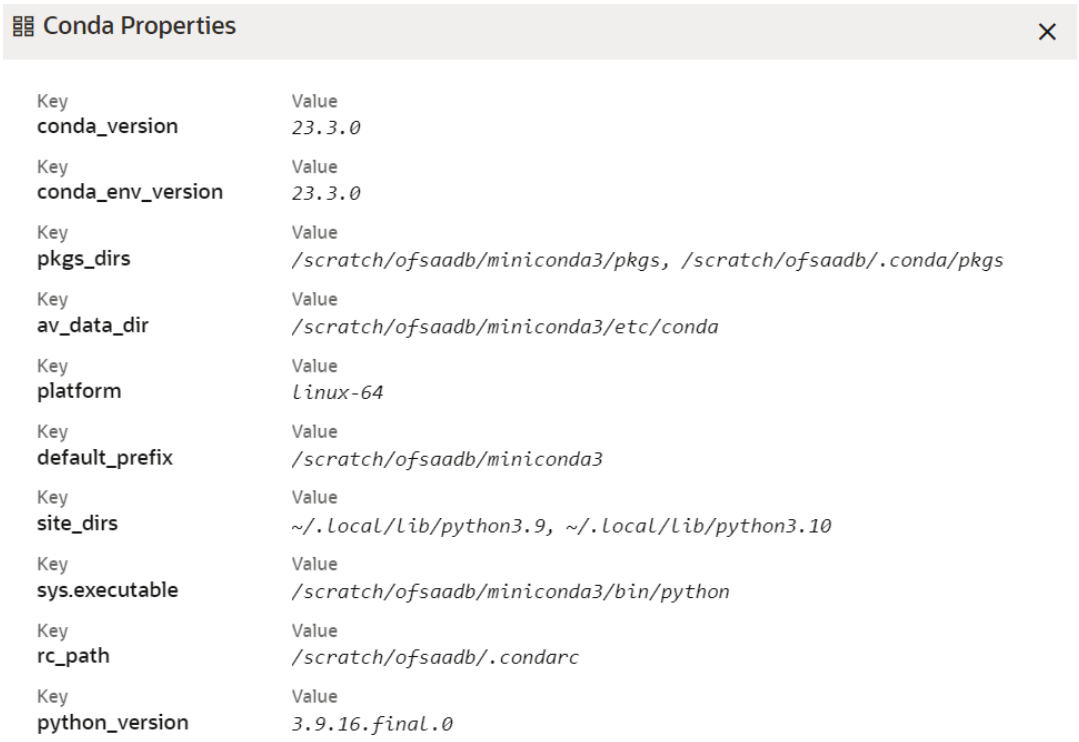
1. Click the **Conda Environments** icon to navigate to the **Environment Summary** page from any other window in the application.

   This page displays the Conda Environments in a tabular format.

2. Click the **Conda Properties** icon.

   The Conda properties are displayed in a key-value format.

Figure 5-27 Conda Properties



5.2.11 Using the .condarc Conda Configuration File

Overview

Overview The Conda Configuration file, .condarc, is an optional runtime configuration file that allows advanced users to configure various aspects of Conda, such as which channels it searches for packages, proxy settings, and environment directories.

> **Note:**
>
> A .condarc file can also be used in an administrator-controlled installation to override the users' configuration. See Administering a multi-user Conda Installation.

The .condarc file can change many parameters, including:

- Where Conda looks for packages.
- If and how Conda uses a proxy server.
- Where Conda lists known environments.
- Whether to update the Bash prompt with the currently activated environment name.
- Whether user-built packages should be uploaded to Anaconda.org.
- What default packages or features to include in new environments.

Creating and Editing

The .condarc file is not included by default, but it is automatically created in your home directory the first time you run the Conda Config command. To create or modify a .condarc file, open a terminal and enter the conda config command. The .condarc configuration file follows simple YAML syntax.

Example:

```
conda config --add channels conda-forge
```

Alternatively, you can open a text editor such as Notepad on Windows, TextEdit on MacOS, or VS Code. Name the new file `.condarc` and save it to your user home directory or root directory. To edit the `.condarc` file, open it from your home or root directory and make edits in the same way you would with any other text file. If the `.condarc` file is in the root environment, it will override any in the Home Directory.

You can find information about your file by typing Conda info in your terminal. This will give you information about your `.condarc` file, including where it is located.

You can also download a sample .condarc file to edit in your editor and save to your user Home Directory or Root Directory.

Example:

To set the auto_update_conda option to False, run:

```
conda config --set auto_update_conda False
```

For a complete list of Conda Config commands, see the command reference. The same list is available at the terminal by running `conda config --help`. You can also see the Conda Channel Configuration for more information.

Conda supports a wide range of configuration options. This page gives a non-exhaustive list of the most frequently used options and their usage. For a complete list of all available options for your version of conda, use the `conda config --describe` command.

Searching for .condarc

Conda looks in the following locations for a `.condarc file`:

`XDG_CONFIG_HOME` is the path to where user-specific configuration files should be stored defined following The XDG Base Directory Specification (XDGBDS). Default to `$HOME/.config` should be used. `CONDA_ROOT` is the path for your base conda install. `CONDA_PREFIX` is the path to the current active environment. `CONDARC` must be a path to a file named `.condarc`, `condarc`, or end with a YAML suffix (`.yml` or `.yaml`).

> **Note:**
>
> Any condarc files that exist in any of these special search path directories need to end in a valid yaml extension (".yml" or ".yaml").

Conflict Merging Strategy

When conflicts between configurations arise, the following strategies are employed:

- Lists-merge
- Dictionaries-merge
- Primitive-clobber

Precedence

The precedence by which the Conda Configuration is built out is shown below. Each new arrow takes precedence over the ones before it. For example, config files (by parse order) will be superseded by any of the other configuration options. Configuration environment variables (formatted like CONDA_<CONFIG NAME>) will always take precedence over the other 3.

Obtaining Information from the .condarc file

You can use the following commands to get the effective settings for conda. The effective settings are those that have merged settings from all the sources mentioned above.

To get all keys and their values:

```
conda config --get
```

To get the value of a specific key, such as channels:

```
conda config --get channels
```

To show all the configuration file sources and their contents:

```
conda config --show-sources
```

Saving Settings to your .condarc file

The .condarc file can also be modified via conda commands. Below are several examples of how to do this.

To add a new value, such as http://conda.anaconda.org/mutirri, to a specific key, such as channels:

```
conda config --add channels http://conda.anaconda.org/mutirri
```

To remove an existing value, such as http://conda.anaconda.org/mutirri from a specific key, such as channels:

```
conda config --remove channels http://conda.anaconda.org/mutirri
```

To remove a key, such as channels, and all of its values:

```
conda config --remove-key channels
```

To configure channels and their priority for a single environment, make a `.condarc` file in the root directory of that environment.

Sample .condarc file

Because the `.condarc` file is just a YAML file, it means that it can be edited directly. Below is an example `.condarc` file:

5.3 API Registry

Using API Registry, you will be able to register APIs above workspace level. Configured APIs can be mapped to multiple workspaces and the mapped APIs will be available for use in model pipelines.

Currently, only token-based and noAuth authentications are supported. Additionally, body parameters can only be provided with text data. All values that needs to be replaced during run time can be given in the format '<<placeholder variable name>>'.

Configured API can be viewed, edited (API code is not editable) and deleted. APIs that are not mapped to any workspace can only be deleted.

Integrating API with Model Pipelines

Mapped APIs will be available for use in model pipelines for the workspace. By selecting a specified API from the dropdown, you can link a node with an API. Additionally, you can provide an output variable name which will be assigned with the response from the API when that node is executed.

> **Note:**
>
> API Registry is available only for %python interpreters. Output variable will be a python variable accessible only for paragraphs with %python interpreters.

In pipeline execute screen, values can be entered for all the placeholder variables during run time.

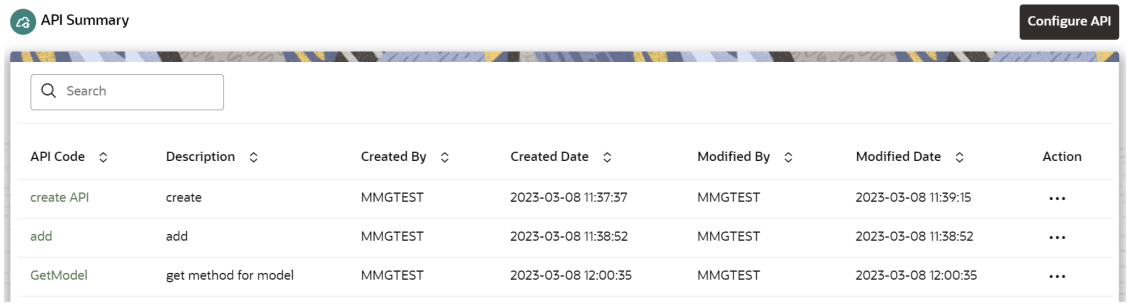
> **Note:**
>
> Since variables are fetched during run time, it will be available in notebook or execute screen only after the first execution. This is the default behavior of all programmatic dynamic forms in Datastudio.

On executing the node/pipeline, result of the API executed can be accessed from the variable given under 'Output_variable_name'.

When the pipeline is imported to another workspace, it will seamlessly use the API that is mapped with the target workspace in the same order as that of the API mapped in source workspace.

Click **API Registry** to navigate to **API Summary** page from any other window in the application. You can configure and manage APIs from this page.

Figure 5-28 API Summary page



The following table provides descriptions for the fields and icons on the **API Summary**page.

Table 5-3 Fields and icons on the Sandbox Summary page

Field or Icon	Description
Search	The field to search for an API. Enter specific terms in the field for which you want to search, and press Enter on the keyboard to display the results.
API Code	The code of the API.
Description	The description for the API.
Created By	The User Id of the User who created the API.
Created Date	The date on which the API was created.
Modified By	The ID of the Last Modified by user who has modified the API.
Modified Date	The date on which the API was modified.
Configure API	Click Configure API to create a new API.
Action	Click the three dots to perform View/Edit/Delete functions on selected API.

5.3.1 Configure an API

To configure an API:

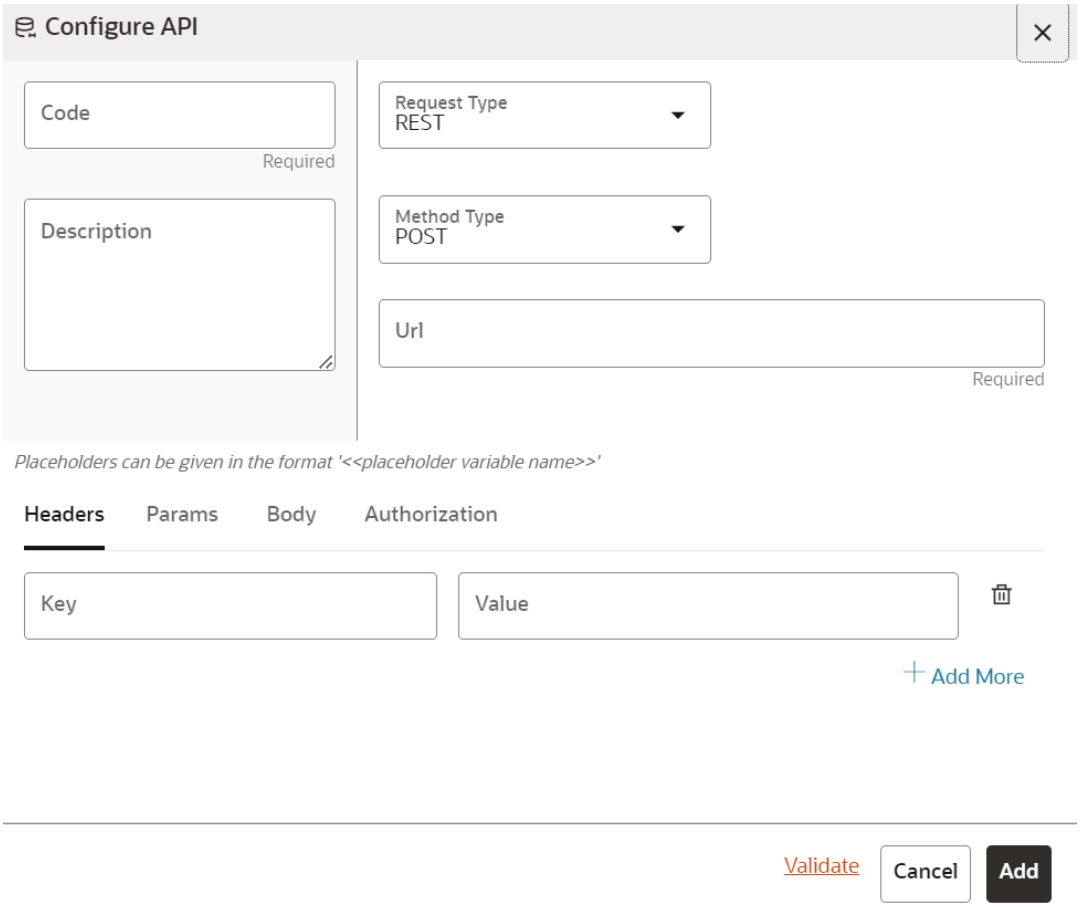
1. Click **API Registry**to navigate to **API Summary** page from any other window in the application.

   This page displays the API records in a tabular format.

2. Click the **Configure API** button.

   The **Configure API**page is displayed.

Figure 5-29 Configure API page



3. Enter the **Code**for the API. This field is mandatory.
4. Enter the **Description**for the API.
5. In the **Request Type**drop-down list, select the required option. Currently, REST type is supported.
6. In the **Method Type** drop-down list, select the required option. The supported types are GET, POST, PUT, and DELETE.
7. Enter the **URL** for the API. This field is mandatory.
8. You can provide the details such as Headers, Params, Body, and Authorization under the respective fields.
9. Click **Validate** to validate the API details.

   You are prompted to add the values for all the placeholder variables. After entering the values for placeholders, click the **Validate** button to view the API responses.
10. Click **Add**.

    The API is added and displayed in the API Summary page. These API's can be selected while creating a Workspace. For more details, see the Workspace Schema section.

5.3.2 View an API Registry

To view an API Registry:

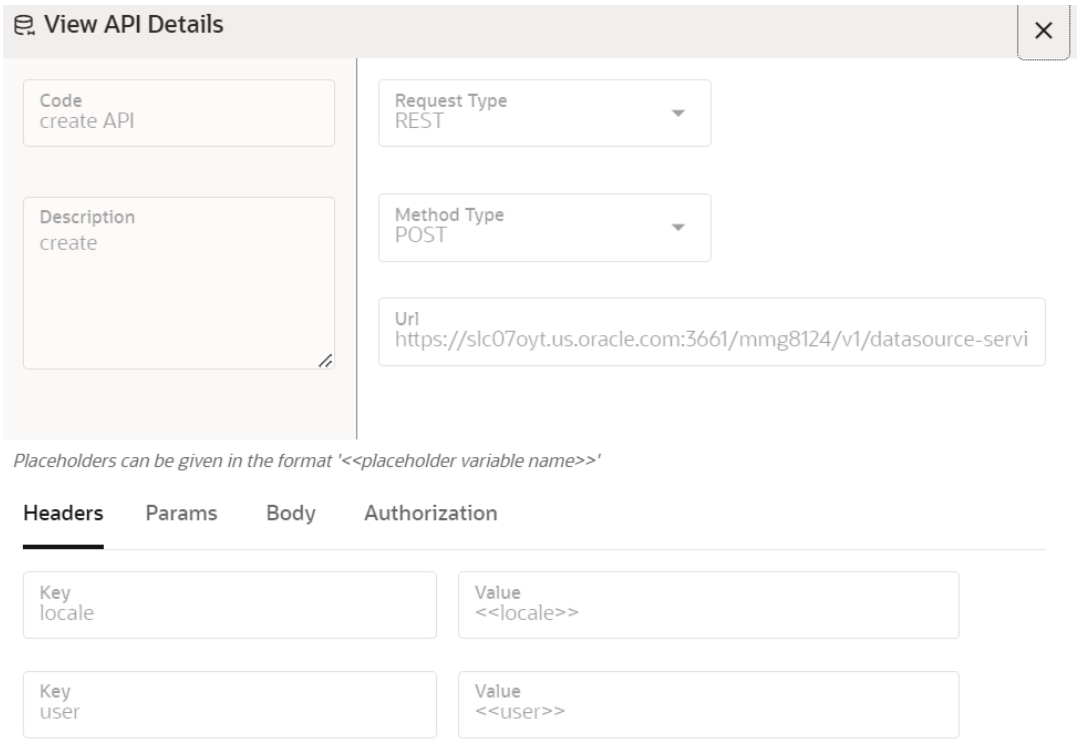
1. Click **API Registry**to navigate to **API Summary** page from any other window in the application.

   This page displays the API records in a tabular format.

2. Click **Action**and select **View** to view the API.

   The details of the API is displayed.

Figure 5-30 View API Details



3. Navigate to the Headers, Params, Body, and Authorization to view the respective details.

5.3.3 Edit an API Registry

To edit an API configuration:

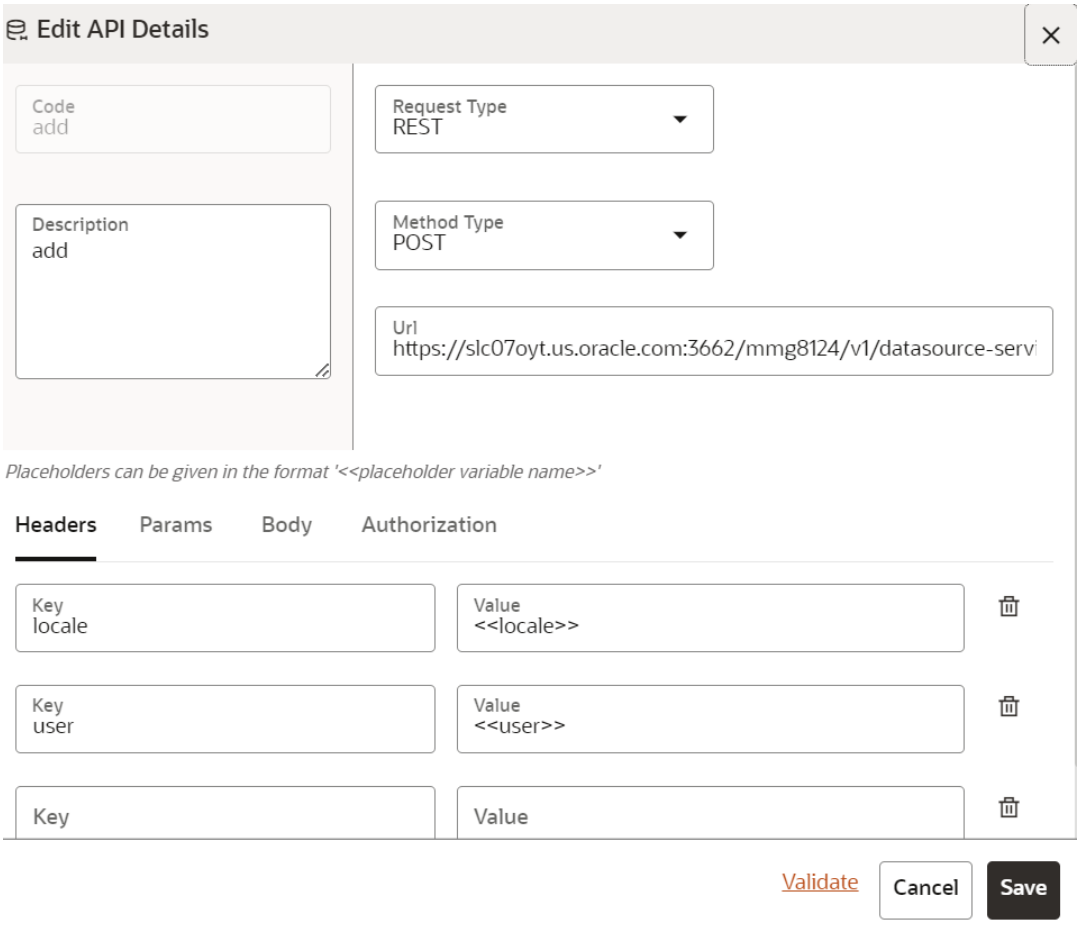
1. Click **API Registry** to navigate to **API Summary page** from any other window in the application.

   This page displays the API records in a table.

2. Click Action next to corresponding API and select **Edit** to edit the API.

   The **Edit API Details**page is displayed.

Figure 5-31 Edit API Details page



For more details on the Edit API Details fields, see Configure an API section.

3. Click **Save**.

5.3.4 Delete an API Registry

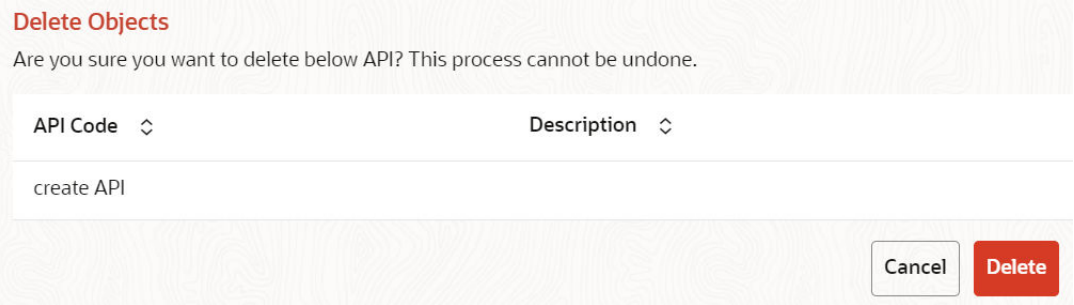
To delete an API Registry:

1. Click **API Registry**to navigate to **API Summary** page from any other window in the application.

   This page displays the API records in a tabular format.

2. Click Action and select **Delete**.

Figure 5-32 Delete Objects screen



3. Click **Delete**.

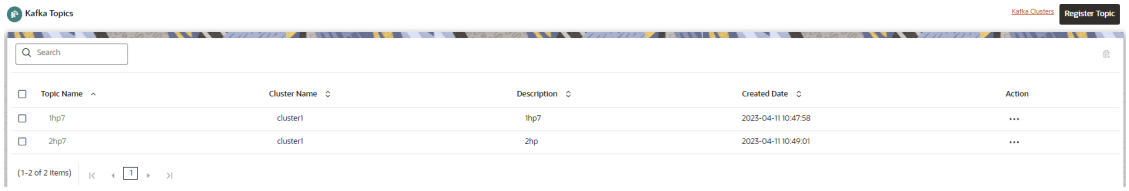
   The API Registry is deleted.

5.4 Kafka Topics

Kafka is a distributed data source which is used to create real-time data pipelines and allows you to decouple data streams and systems.

Click **Kafka Topics** icon to navigate to **Kafka Topics** page from any other window in the application. You can configure and manage Kafka topics and clusters from this page.

Figure 5-33 API Summary page



The following table provides descriptions for the fields and icons on the **Kafka Topics**page.

Table 5-4 Fields and icons on the Sandbox Summary page

Field or Icon	Description
Search	The field to search for a Kafka topics and clusters. Enter specific terms in the field for which you want to search, and press Enter on the keyboard to display the results.
Topic Name	The topic name of the Kafka.
Cluster Name	The cluster name of the Kafka.
Description	The description of the Kafka.
Created Date	The date on which the API was created.
Kafka Clusters	Click Kafka Clusters to view/create/delete Kafka Clusters.
Register API	Click Register API to create a new Kafka topic.
Action	Click the three dots to perform View/Edit/Delete functions on selected Kafka topic.

5.4.1 Create a Kafka Cluster

To create a Kafka Cluster:

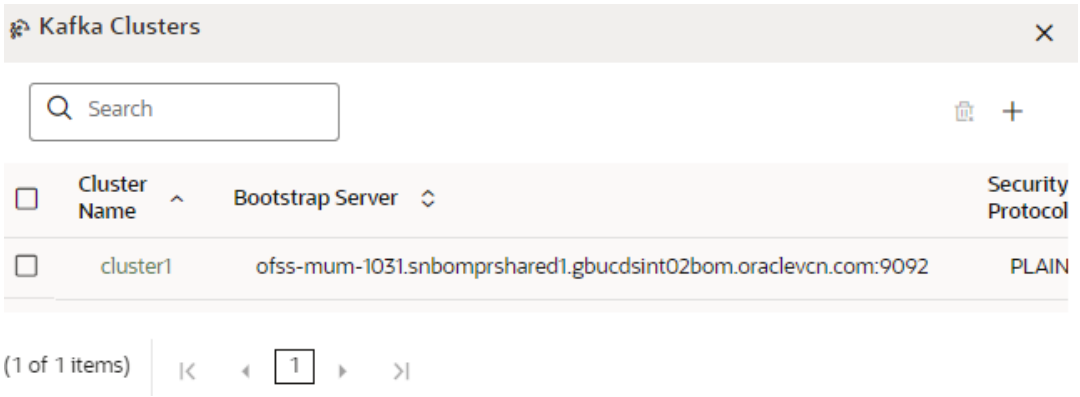
1. Click **Kafka Topics** to navigate to **Kafka Topics** page from any other window in the application.

   This page displays the Kafka Topics in a tabular format.

2. Click **Kafka Clusters**option.

   The **Kafka Clusters**page is displayed.

Figure 5-34 Kafka Clusters page



3. Click **Add** .

   The Add Cluster window is displayed.

4. Enter the **Cluster Name**. This field is mandatory.
5. Enter the **Bootstrap server** details. This field is mandatory.
6. Select the **Security Protocol** as **Plain Text**.

> **Note:**
>
> In the current release:
>
> - The Kafka feature is not supported in Solaris OS.

7. Click **Save**.

   The Kafka cluster is created, and this will be displayed while registering the Kafka Topic.

5.4.2 Delete a Kafka Cluster

To delete a Kafka Cluster:

1. Click **Kafka Topics** to navigate to **Kafka Topics** page from any other window in the application.

   This page displays the Kafka Topics in a tabular format.

2. Click **Kafka Clusters**option.

The **Kafka Clusters**page is displayed.

3. Select the Kafka Cluster which you want to delete and click **Delete**.

   A confirmation message is displayed.

4. Click **Delete**.

   The Kafka Cluster is deleted.

5.4.3 Register a Kafka Topic

To register a Kafka Topic:

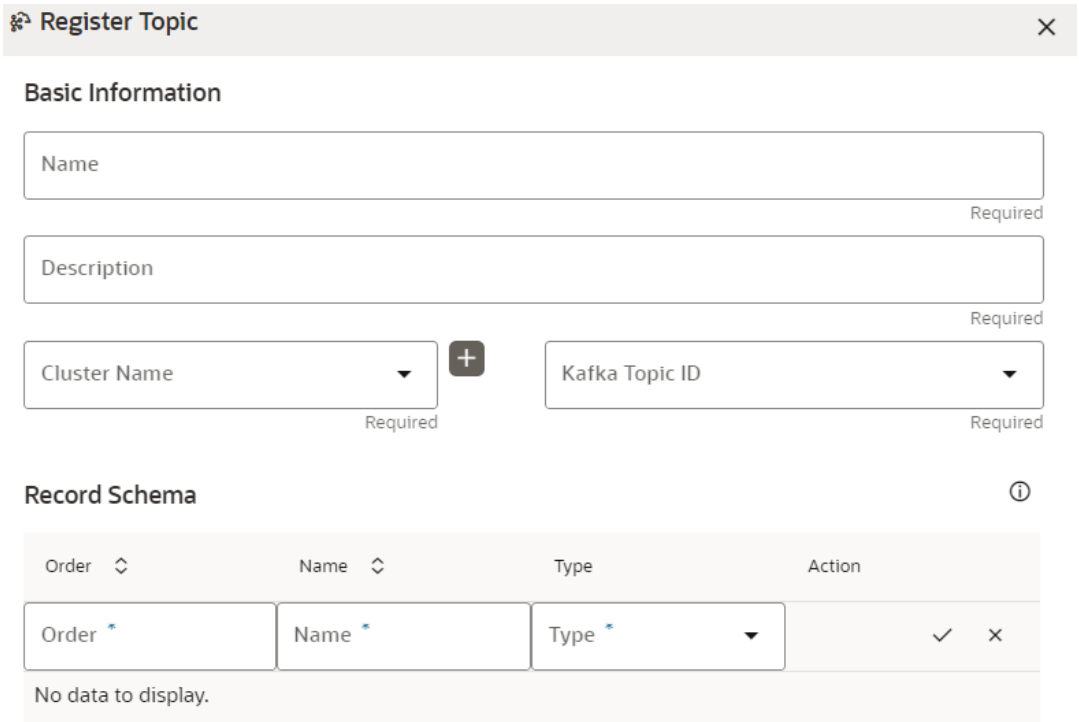
1. Click **Kafka Topics**to navigate to **Kafka Topics**page from any other window in the application.

   This page displays the Kafka Topics in a tabular format.

2. Click **Register Topic** button.

   The Register Topic page is displayed.

Figure 5-35 Register Topic page



3. Enter the **Kafka Topic Name**. This field is mandatory.
4. Enter the **Description** for the Kafka Topic. This field is mandatory.
5. In the **Cluster Name** drop-down list, select the required option and enter a unique Kafka Topic ID. This field is mandatory.
6. In the **Record Schema**field, enter the order, name, and the data type for the Kafka topic.

> **Note:**
>
> Currently, **PyFlink, Python,** and **Java Schema** types are supported. For more details, click the icon next to Record Schema field.

7. Click **Register**.

   The Kafka topic is added and displayed in the Kafka Topics page. These topics can be selected while creating a Workspace.

5.4.4 View a Kafka Topic

To view a Kafka Topic:

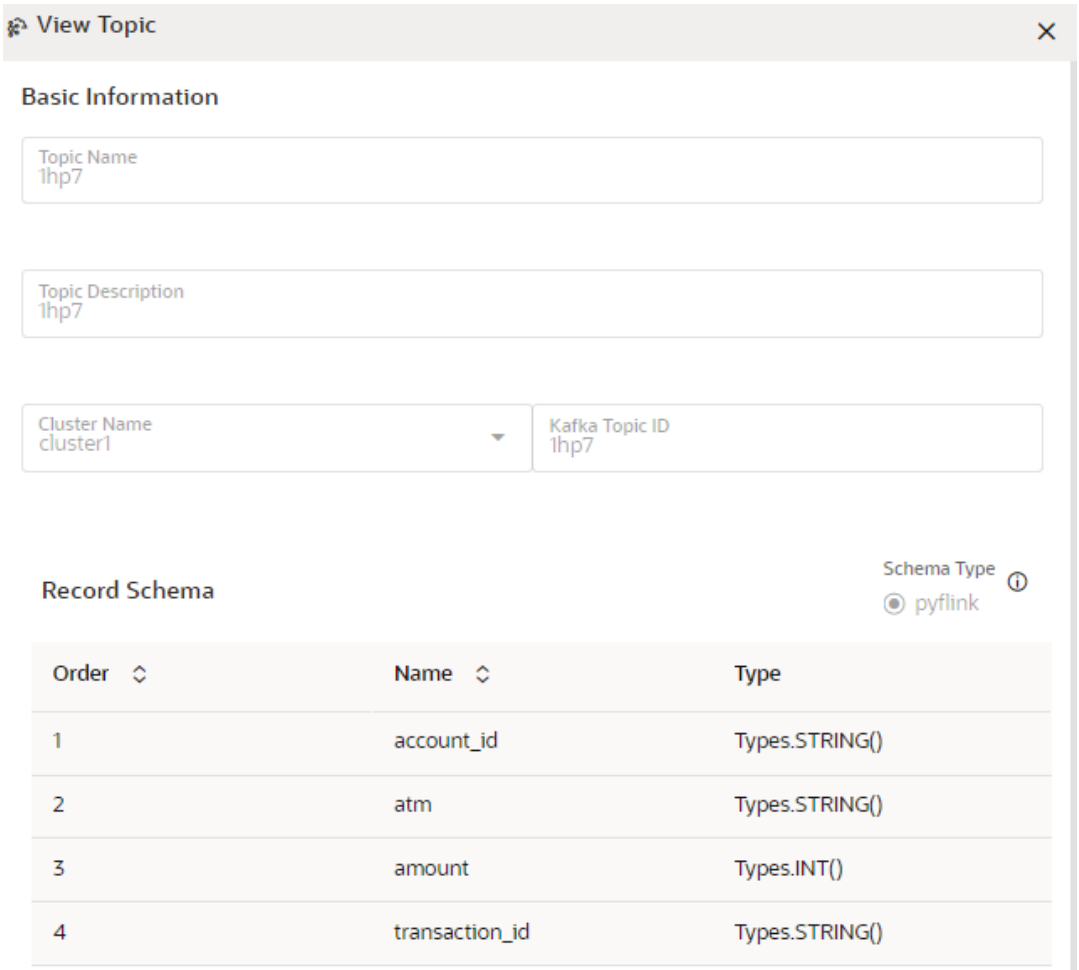
1. Click **Kafka Topics** to navigate to **Kafka Topics** page from any other window in the application.

   This page displays the Kafka Topics in a tabular format.

2. Click **Action**and select **View**.

   The details of the Kafka Topic is displayed.

Figure 5-36 Kafka Topics page



5.4.5 Edit a Kafka Topic

To edit a Kafka Topic:

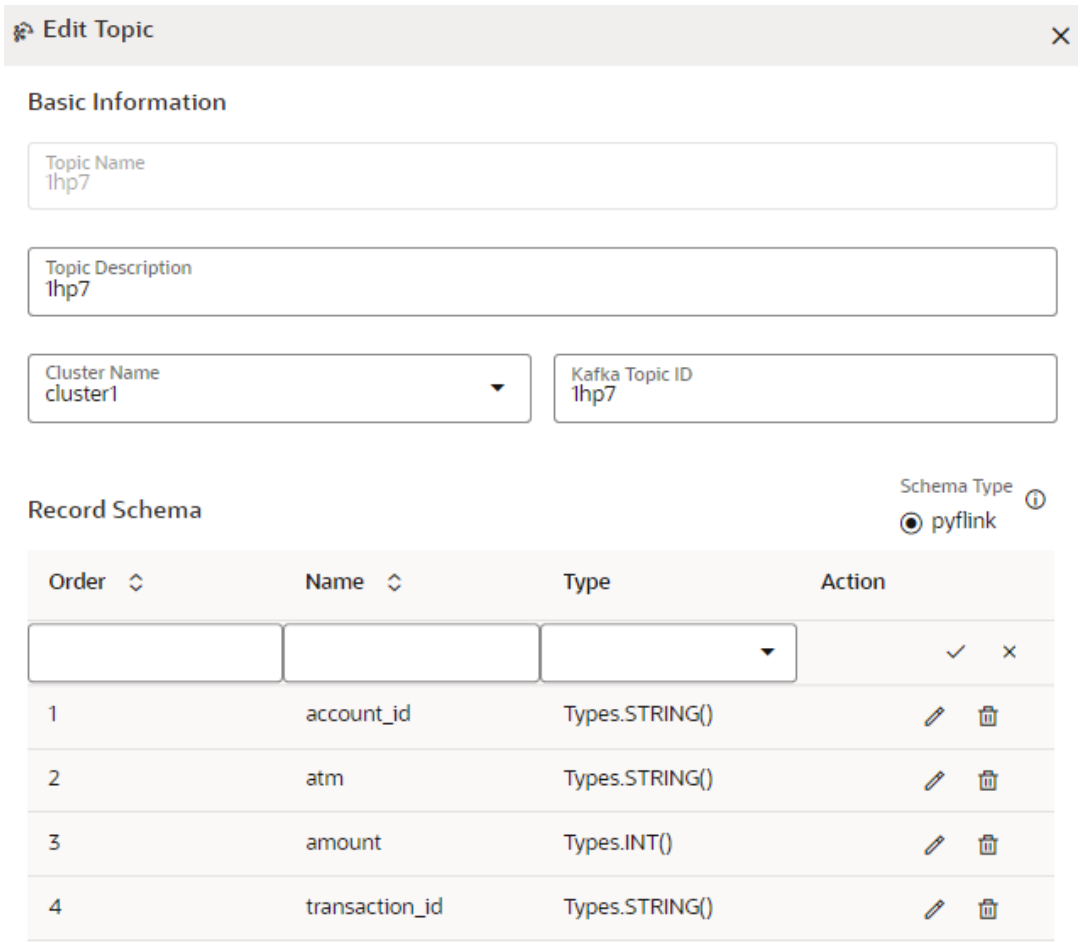
1. Click **Kafka Topics**to navigate to **Kafka Topics**page from any other window in the application.

   This page displays the Kafka Topics in a tabular format.

2. Click **Action** next to corresponding Kafka Topic and select **Edit**.

   The **Edit Topic** page is displayed.

   Figure 5-37 Edit Topic page



   For more details on the Kafka Topic fields, see Register a Kafka Topic section.

3. Click **Update**.

5.4.6 Delete a Kafka Topic

To delete a Kafka Topic:

1. Click **Kafka Topics** icon to navigate to **Kafka Topics** page from any other window in the application.

This page displays the Kafka Topics in a tabular format.

2. Click **Action** and select **Delete**.

   A confirmation message is displayed.

> **Note:**
>
> You cannot delete a Kafka Topic mapped to a workspace.

Click **Delete**. The Kafka Topic is deleted.

5.5 OFSAA Environment

You can register OFSAA environment (Simulation and Production) from this page.

> **Note:**
>
> The user must be mapped to ENVACCESS role to view the OFSAA Environment option.

To register a OFSAA environment:

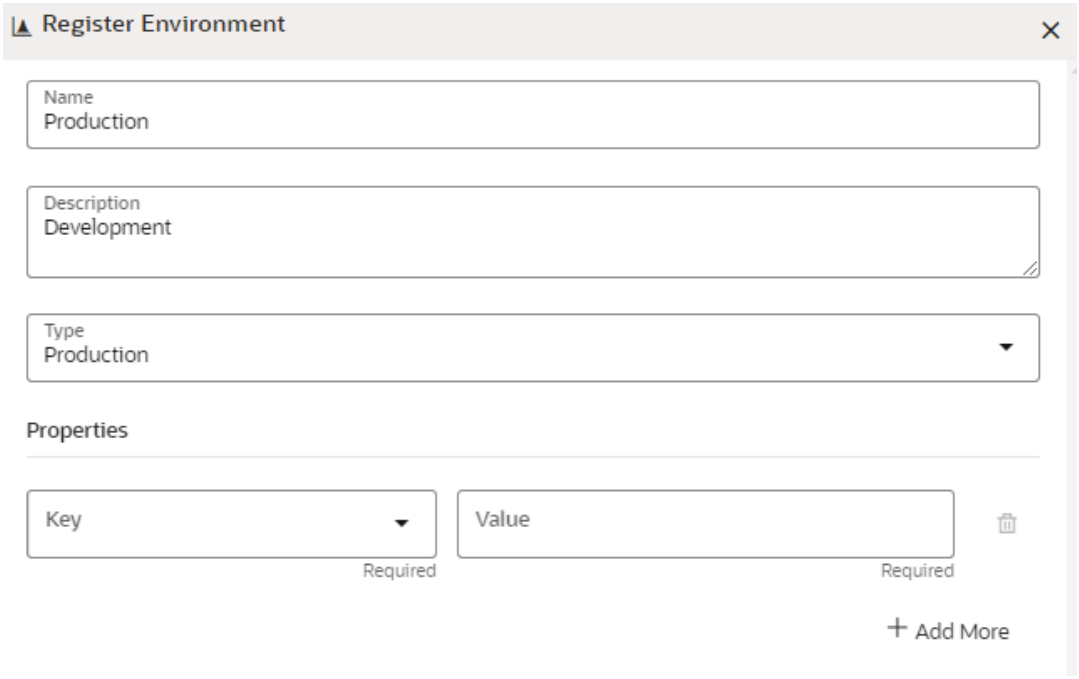
1. Navigate to the OFSAA environment page.

   This page displays the OFSAA environments in a tabular format.

2. Click **Register Environment**.

   The Register Environment page is displayed.

Figure 5-38 Register a OFSAA Environment



3. Enter the environment **Name**. This field is mandatory.
4. Enter the **Description** for the environment. This field is mandatory.
5. In the **Type** field, select the environment as either Production or Simulation.
6. In the **Key** field, select the required option from the drop-down and enter value for the selected property. You can add more properties by clicking **Add More** icon.
7. Click **Create**.

   The OFSAA Environment is added and displayed in the OFSAA Environment page.

   You can view, edit, and delete OFSAA environments by clicking Actions icon.

5.6 Data Studio Options

The application's Notebook Server has the following configurable options:

- Interpreters
- Tasks
- Permissions
- Credentials
- Templates

To access the Data Studio Options page:

1. Click **Launch Workspace** to launch the workspace to display the **Dashboard** window with the application configuration and model creation menu.
2. Hover your mouse over the Data Studio Options widget . The following options are available:
   - Interpreters
   - Tasks
   - Permissions
   - Credentials
   - Templates

> **Note:**
>
> DS 23.4.10 has introduced 'Select-Multiple' forms in Data Studio. The same should now be handled in MMG (parameters sets, dashboard, execute pipeline drawer, simulations, and so on).

5.6.1 Configure Interpreters

An Interpreter is a program that directly executes instructions written in a programming or scripting language without requiring them previously to be compiled into a machine language program. Interpreters are plug-ins that enable users to use a specific language to process data in the backend. Examples of Interpreters are jdbc-interpreter, spark-interpreters, python-interpreters, etc. Interpreters allow you to define customized drivers, URLs, passwords, connections, SQL results to display, etc.

In OFS Compliance Studio, Interpreters are used in Notebooks to execute code in different languages. Each Interpreter has a set of properties that are adjusted and applied across all notebooks. For example, using the python-interpreter makes it possible to change between versions, whereas the jdbc-interpreter offers to customize the URL, schema, or credentials. You can either use a default interpreter variant or create a new variant for an interpreter. You can create more than one variant for an interpreter. The benefit of creating multiple variants for an Interpreter is to connect different versions of interpreters (Python ver:3, Python ver:2, etc.). This helps to connect a different set of users and database schema. For example, Compliance Studio schema, BD schema, etc. Compliance Studio provides secure and safe credential management such as Oracle Wallet (jdbc wallet), Password (jdbc password), or KeyStores to link to interpreter variants to access secured data.

The application has ready-to-use interpreters and you can configure them based on the use case. Additional variants of interpreters are created as multiple users might require different settings to access the database securely. The jdbc Interpreters use the credentials to enable secure data access.

> **Note:**
>
> Pyspark, spark, and ore are a few other available interpreters.

Interpreters are configured when you want to modify URL, data location, drivers, enable or disable connections, etc.

To configure ready-to-use interpreters:

1. Click the Interpreter that you want to view from the list displayed on the LHS. The default configured interpreter variant is displayed.
2. Modify the values in the fields as per requirement. For example, to modify a parameter's limit, connect to a different schema, PGX server, and so on.
3. Click **Update**. The modified values are updated in the Interpreter.

5.6.2 Manage Tasks

Tasks are created when notebooks or paragraphs are executed by the Notebook users. It is important to know the status of the execution, whether the tasks are created, rejected, canceled, etc. The Tasks page allows you to view the status of the task and associated notebooks, paragraphs, interpreters, etc. By default, all the tasks are listed on the Task page. You can view the specific task using filters such as task status, date of creation, and notebook name.

5.6.3 Manage Permissions

You can view the logged-in users and view, add, or modify ready-to-use permissions granted to the users, roles, or groups. You can create groups, roles, and permission templates (actions).

> **Note:**
>
> You can only view users and their details.

See the *User Access and Permissioning Management* section in the OFS Compliance Studio Administration and Configuration Guide.

5.6.4 Manage Credentials

Compliance Studio provides secure and safe credential management. Examples of credentials are passwords, Oracle Wallets, or KeyStores. This section links credentials (a wallet and a password) to the jdbc interpreter variant to enable secure data access. This linking enables the jdbc interpreter to securely connect to the specified Oracle Database. You can also create new credentials based on your requirement to connect to the new interpreter variants.

For more information on Credentials, see the Link Credentials section in the OFS Compliance Studio Administration and Configuration Guide.

> **Note:**
>
> You can link credentials only to the jdbc interpreter. The Credentials section is enabled if an Interpreter variant can accept credentials.

5.6.5 Configure Templates

To create, import, and set the default FCCM template in the Template Dashboard:

1. Click CS Launch Workspace to display the CS Production Workspace window.
2. Click Templates from Design Studio Options.

   Compliance Studio offers different formats to view the result after a paragraph's execution. Templates enable you to define parameters to customize the result formats. You can customize the visualization of the result by defining parameters in a template and then applying that template to a notebook.

   > **Note:**
   >
   > - Compliance Studio comes with a default template, but users can customize this at the template level but can also override any global template settings in each notebook paragraph.
   > - It is recommended to use the template that is available from out-of-the-box Compliance Studio.

   For more information, see the *Interpreter Configuration and Connectivity* section in the OFS Compliance Studio Administration and Configuration Guide.

6
Modeling

Modeling covers the following:

- **Datasets**: Explore data, engineer features, and monitor data drifts.
- **Model Libraries**: Manage python libraries to create and use models.
- **Model Techniques**: Manage techniques and algorithms from registered model libraries.
- **Model Catalog**: A centralized hub for all your machine learning models.
- **Pipelines**: Design, deploy and govern end-to-end machine learning pipelines for seamless model development.
- **Graphs**: Analyze complex relationships within your data through interactive graph representations.

6.1 Dataset

You can define a metadata about how you want to create a dataset and also take a snapshot of real time data and store it, which you can later use in the pipeline.

This is similar to T2Ts, where you can select data from Data Source, such as file, table or another dataset and so on. The data can be imported from one column of one table, and another column from another table, or any file. After extracting the data from tables or files, provide the name to Target Dataset.

The Dataset is a trail-based UI that allows you to configure the Dataset details.

6.1.1 Access the Dataset Summary

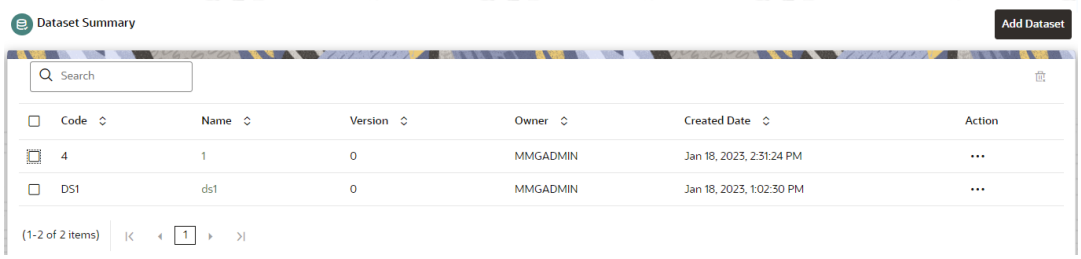
The Dataset Summary page gives access to the various Dataset functions such as create, view, and delete.

To access the Dataset Summary page, follow these steps:

1. Click **Launch Workspace**next to corresponding Workspace to Launch Workspace to display the **Dashboard** window with application configuration and model creation menu.
2. In the LHS menu, click **Dataset** to display the **Dataset Summary** window.

   This window displays the dataset records in a table.

Figure 6-1 Dataset Summary page



The following table provides descriptions for the fields and icons on the **Dataset Summary** page.

Table 6-1 Fields and icons on the Sandbox Summary page

Field or Icon	Description
Search	The field to search for Dataset. Enter specific terms in the field for which you want to search, and press Enter on the keyboard to display the results.
Code	The code of the dataset.
Name	The name of the dataset.
Description	The description for the Workspace
Version	Version of dataset
Created Date	The date on which the Dataset was created.
Owner	The owner of the dataset.
Add Dataset	Click Add Dataset to create a new Dataset.
Delete	Click Delete to delete multiple Datasets.
Action	Click the three dots to perform View/Edit/Delete/ Profile functions on selected dataset.

6.1.2 Create a Dataset

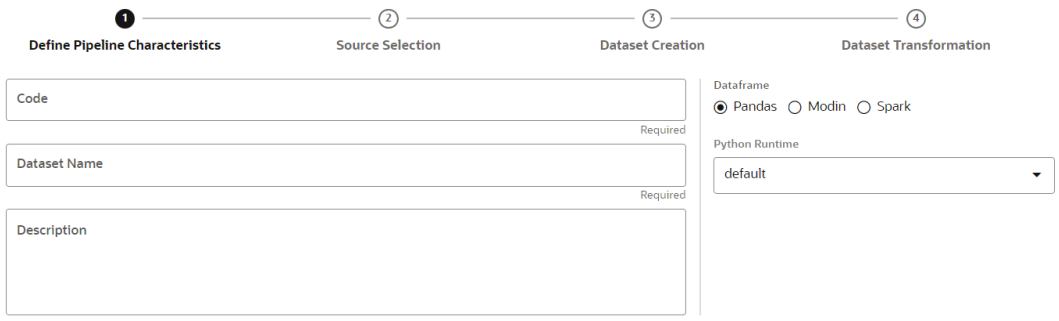
The Dataset creation requires entry of the source of dataset and validation. These datasets are required for schema creation.

6.1.2.1 Creating a Dataset

To create a Dataset:

- On the Dataset Summary page, click **Add Dataset**.

Figure 6-2 Dataset Creation window



Provide the following information to create a Dataset:

- Define Pipeline Characteristics
- Source Selection
- Dataset Creation

- Dataset Transformation

6.1.2.1.1 Define Pipeline Characteristics

Enter basic configuration details in this window.

To configure the basic details for the dataset:

1. Enter the required details in the **Define Pipeline Characteristics** window as shown in the following table.

Table 6-2 Details for Basic Details pane

Field	Description
Code	Enter the identification code of the dataset. This field is limited to 30 alphanumeric characters.
Dataset Name	Enter the name of dataset. This field is limited to 30 alphanumeric characters. Space not exceeding 30 characters. You cannot keep this field blank.
Description	Enter the purpose of the creation of the dataset. This field is limited to 150 alphanumeric characters. Space not exceeding 150 characters.

2. Select the data library from the options: **Pandas**, **Modin**, or **Spark** and select **Python Runtime** from the drop-down and click **Close**.
   - **Pandas**: An open-source data manipulation library for Python. It provides data structures such as Series (1-dimensional) and DataFrame (2-dimensional) that allow for easy manipulation and analysis of data. It also provides tools for reading and writing data to various file formats, including CSV, Excel, and SQL databases. **Pandas** is the default selection.
   - **Modin**: An open-source library that allows for faster operations on DataFrames using distributed computing which can lead to significant speed improvements, particularly for large datasets or computationally expensive operations.
   - **Spark** : Pyspark option for scaling dataset.
     If Spark library is selected, the Python Runtime drop-down option is not displayed.
3. Click **Next** to go to the next step.

6.1.2.1.2 Source Selection

This section allows you to define the source of data. Here, you can choose the data structures from an existing datasources. Click on the icon next to Data Source field to view the tables/ data source information associated with the data source.

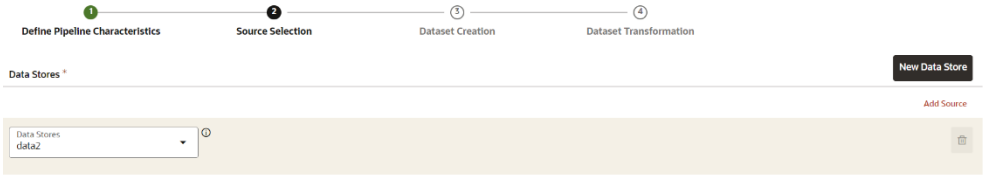
To configure the Source details for the dataset, follow these steps:

1. Select the required **Data Store** from the drop-down.

   OR

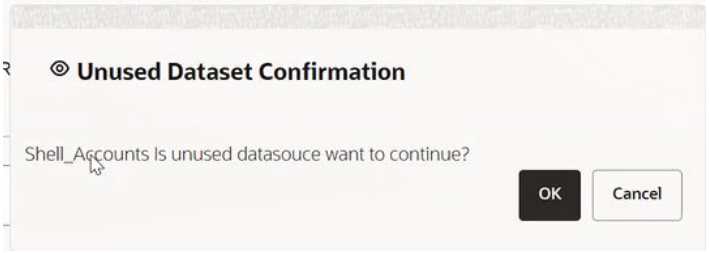
   Click **New Data Store** to create a data source.

Figure 6-3 Source Selection Details screen



> **Note:**
>
> The MISDATE feature is displayed only when the Data Source field is selected as a **File Type**.
> If enabled, you need to add the Physical and Feature name for the Data Source.

Table 6-3 Details for Basic Details pane

Field	Description
Data Stores	The Data Stores drop-down list shows all the data sources and the data dorms those are tagged with the corresponding workspace. In addition, this lists all the available unused data source, which can be used. You can search the data sources by clicking on **Data Store information** icon. A confirmation message is displayed to tag an unused datasource with workspace. **Figure 6-4 Unused Dataset Confirmation message**  Here, either use the unused or create a new datasource using **New Datasource**. When you create a new datasource, it becomes as unused. You can add more Data Sources using **Add More** link. Use **Delete**to delete added Data Source.

2. Click **Next** to go to the next step or click **Previous** to go back to previous step.

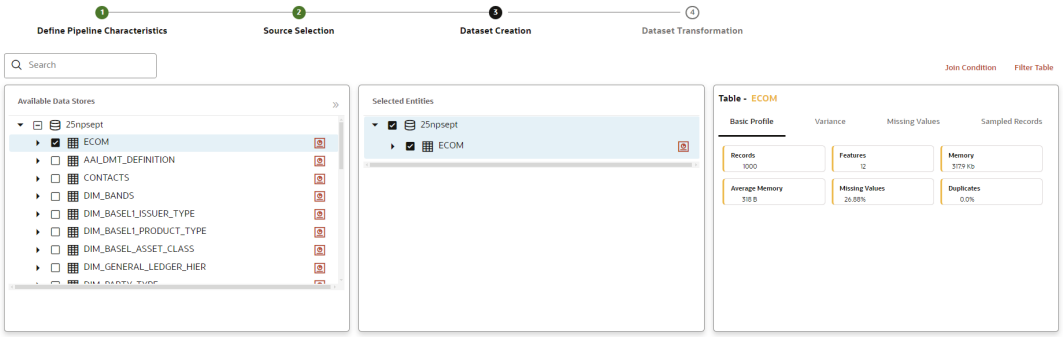
6.1.2.1.3 Create a Dataset

This window allows you to select entity part (for example, column of table) of the datasource. The data source selected in the previous step requires the definition of database objects to be used.

To configure the Dataset Creation:

1. Select the required data source from the **Available Data Sources** pane and click >> to move the selected data to **Selected Entities**pane. The **Available Data Sources** section shows the high-level data sources. You can select data from multiple data source entities. The selected Data Sources are displayed in **Selected Entities** pane. You can use to select or use to de-select the data in the **Selected Entities** pane. Additionally, clicking on **View Profile** icon displays the profile details in the **Basic Profile** tab.

Figure 6-5 Dataset creation page



2. For more information on Joins, see Joining Tables section.

   For more information on Filtering, see Filtering Tables section.

3. Click **Next** to go to the next step or click **Previous** to go back to previous step.

6.1.2.1.3.1 Joining Tables

Use Join Condition to combine the data source details. Each data sources have multiple tables and multiple columns, so you can join them using Join Condition window. Joins can be applied only when user has selected more than one table.

Figure 6-6 Join tables

You can join Tables either using drop-down or using the Joinscript option.

Using Drop-down

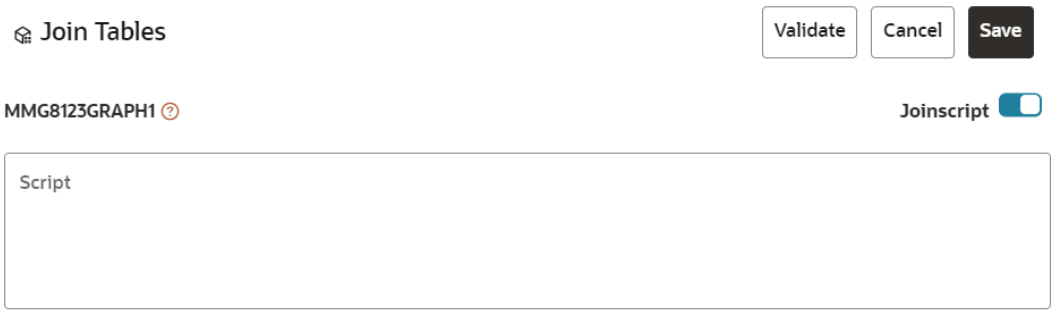
Select the Column name of Table from the drop-down list. You can use multiple join conditions by using icon.

> **Note:**
>
> In the current release, only INNER JOIN is present.

Using Script

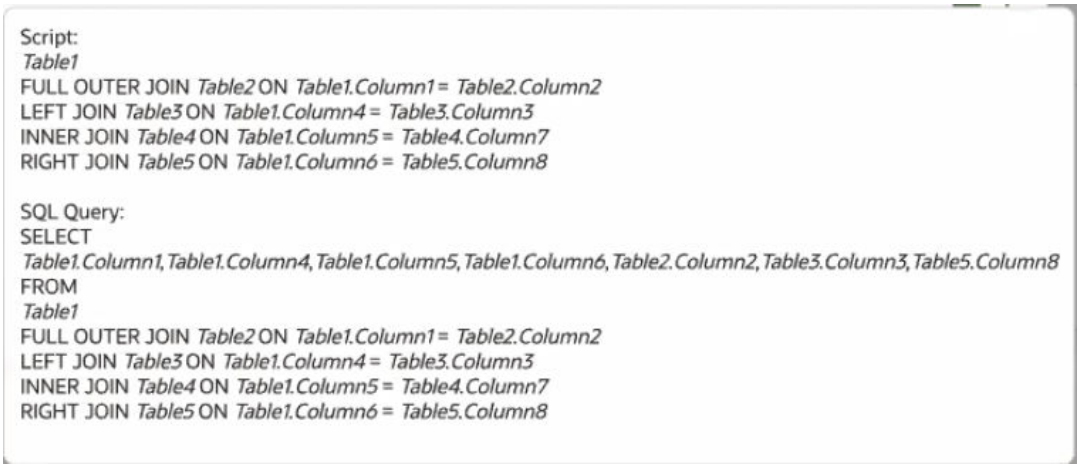
1. To use script, enable **Join Script** option.
   The Join Tables window is displayed to add scripts.

Figure 6-7 Join Tables window



You must provide the script in below format:

Figure 6-8 Script



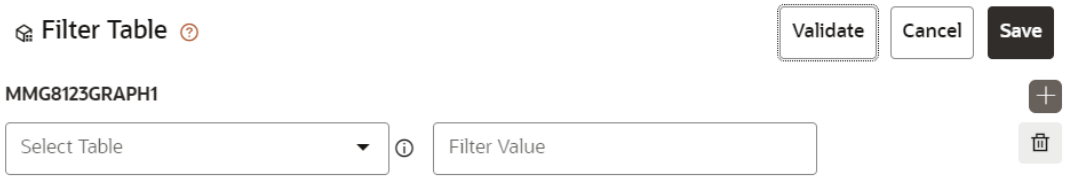
2. Click Validate to check the Join conditions.
3. Click Save.

6.1.2.1.3.2 Filtering Tables

You can also use the filters in dataframe creation from entity.

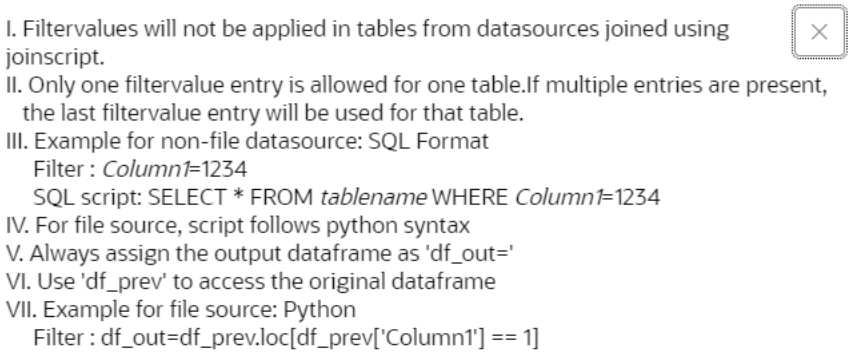
1. Click **Filter Table** to navigate to Filter tables window.

Figure 6-9 Filter Table



There are certain rules to be followed when adding filters.

Figure 6-10 Rules to be followed when adding filters

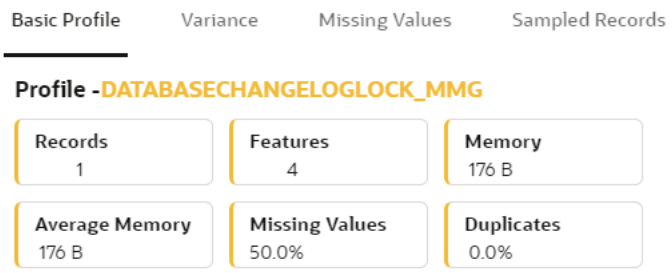


2. Multiple filters can be added by using **Add**.
3. Click **Validate** to check the Filter conditions.
4. Click **Save**.

6.1.2.1.3.3 Viewing Profile

The Profile section from RHS shows the report based on Selected Objects (complete dataset) section. For example, Missing Values report shows the details of missing data in selected table in Selected Objects section.

Figure 6-11 Selected Objects section



6.1.2.1.4 Dataset Transformation

This window allows you to transform the data source information. This complete grid displays the Table like structure, which helps you to make the data better using many methods, such as by remove all the missing values, performing scaling and so on.

This window shows the following columns:

- **Physical Name**: Shows the Physical Name of column.
- **Feature Name**: Shows the Feature Name of column along with its data type. You can edit the Feature Name.
- **Observations**: The Observations field shows the value of column. For example, if any column is missing any value, you can easily identify missing value columns. Use Transformation button to fix these values.
- **Create Feature**: To create a new feature/column by doing operations on the existing features.
- **Calculate Fairness**: You can calculate the fairness metrics of the dataset by choosing the required metric, target variable, and protected features. For more details, see Fairness section. Click Help icon for more details on how the fairness is calculated.
- **View Report**: To view the complete report of the sampled data.
- **Re-Validate**: To re validate all the transformations.
  You can also revalidate an individual transformation by clicking the More Action and select the Re-Validate option
- **Add Transformation**: To add new transformations

The profile button helps you to view profile the final data frame that you have selected. For more information, see the Viewing Profile section.

Here, you can perform the following actions:

- Re-ordering of transformations
- Insertion of transformation

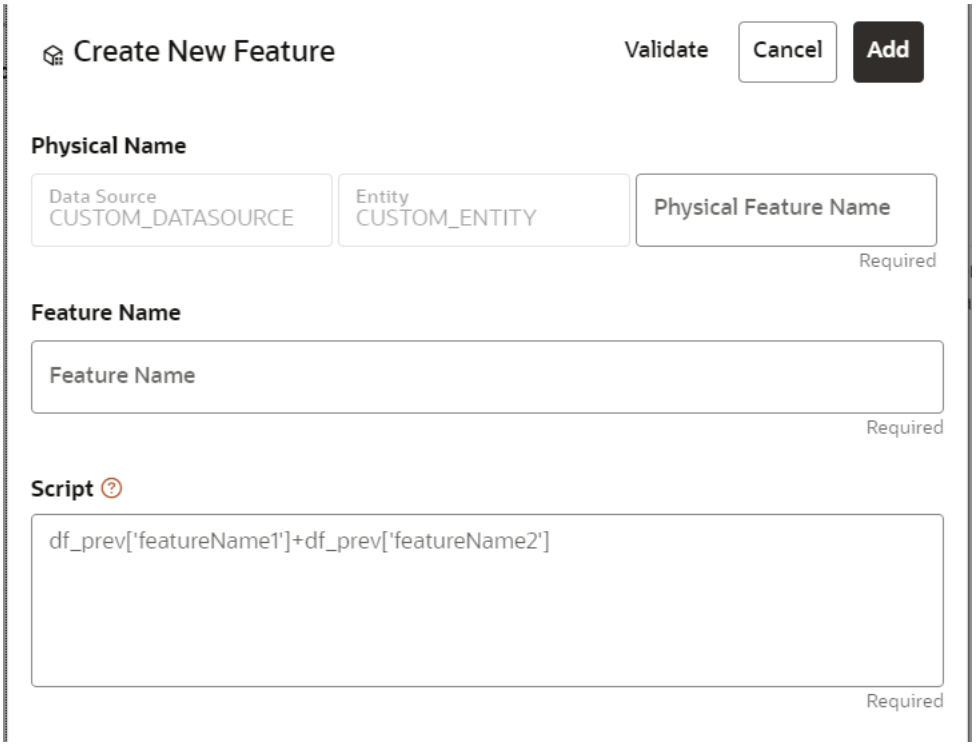
6.1.2.1.4.1 Create New Feature

The Create New Feature window allows you to create a new feature. This is used to add a new column to dataset. This is useful, if you want to add a new column based on derived T2Ts (from data source).

To create a New Feature, follow these steps:

1. Click Create Feature on Dataset Transformation of Dataset UI.

Figure 6-12 Create Feature on Dataset Transformation of Dataset UI



2. Enter the following details:
   - **Physical Feature**: Name of Physical feature
   - **Feature Name**: Logical Name of feature
   - **Script**: Update the script
3. Click **Add** after validating the feature.

   The new feature will be added at end of LHS section.
4. You can also edit or delete a newly added feature.

6.1.2.1.4.2 Add Data Transformation

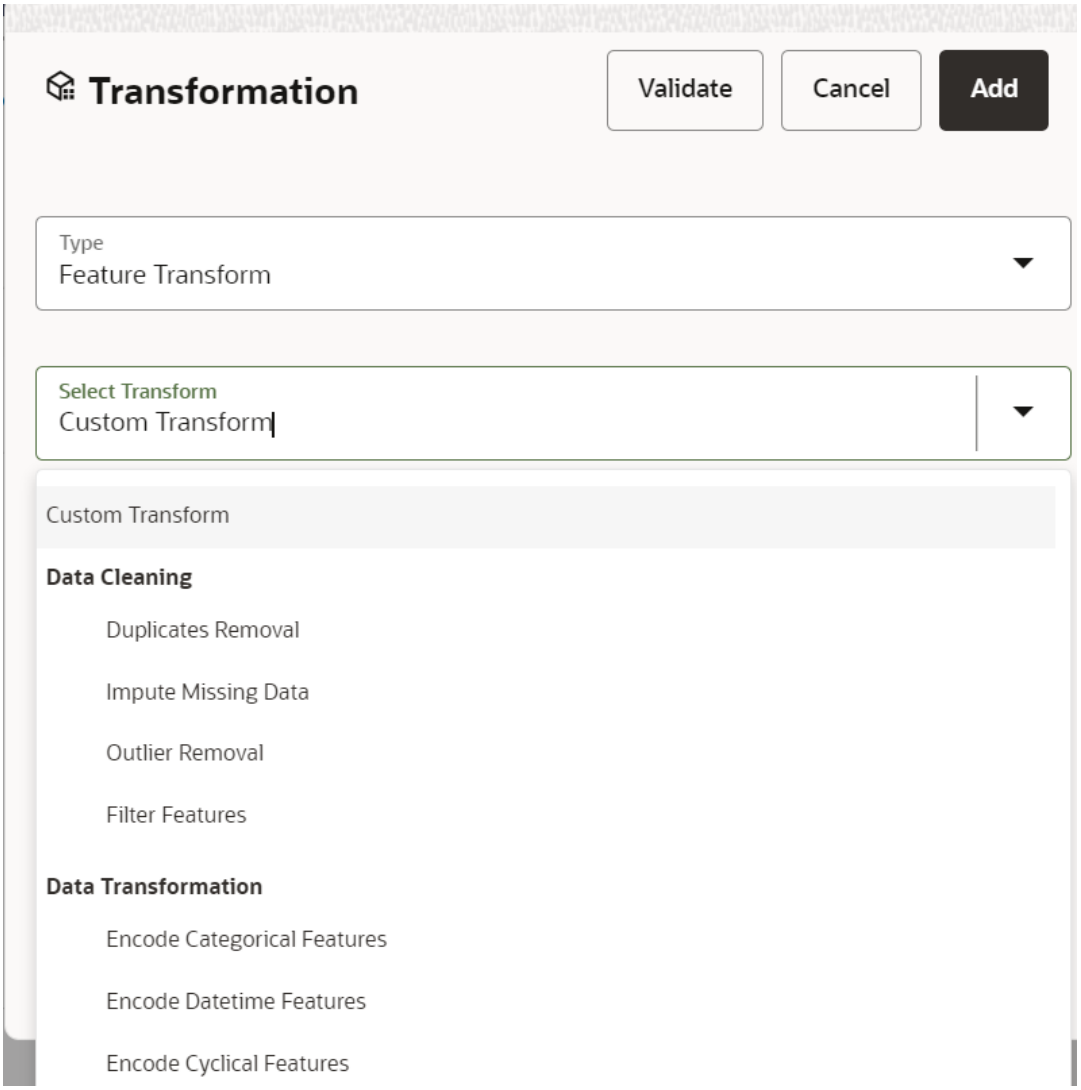
To add the data transformation, follow these steps:

1. Click **Add**.

   The Transform window is displayed.
2. Select the Transformation Type.

   Following types are available:
   - **Dataframe Transform**: This is used to transform entire dataframe. For example, if you want to remove all missing values from all the columns of entire sampled data.
   - **Feature Transform**: This is used to transform a particular column of dataframe.

Figure 6-13 Transformation screen



3. Select **Transform**: Here you either enter values using Method and Argument fields, or script.
4. Click **Validate** to validate the details.
5. Click **Add** to add the new transformation.

   The New Transform is created and displayed at RHS section.
6. Click **Finish** to navigate to Dataset Summary window.

   This saves the metadata of dataset.

   Example:

   If there is “missing value” for one of the columns, then perform following steps to add the transform.

   a. Click Add on Dataset Transformation window.
   b. Select Type as Feature Transform.
   c. Select Transform as Impute Missing Value.
   d. Select the Physical Name from Physical Name drop-down list.

e. Select the Method and enter parameter in Argument field. The Method is updated on selected Column type. For example, if the selected column type is numerical, then following methods will be available: Simple, Constant, KNN, and Mice
OR

Enter the script
Below are the Sample Custom Scripts:

- Directly pass input data frame to output: df_out=df_prev
- Drop first row of data frame: df_out=df_prev.drop(0,axis=0) if not df_prev.empty else df_prev
- Drop column from data frame: df_out = df_prev.drop('colname',axis=1)

6.1.2.1.4.3 Sample Custom Scripts

Below are few sample scripts which the users can refer to create transformations.

Table 6-4 Sample Custom Scripts

Function	Script	Comments
Directly pass input data frame to output	df_out=df_prev	
Drop first row of data frame	df_out=df_prev.drop(0,axis=0) if not df_prev.empty else df_prev	
Drop column from data frame	df_out= df_prev.drop('colname',axis=1)	FEATURE TRANSFORMATION

6.1.2.1.4.4 Transformations

These are the various transformations which can be done from the UI.

Table 6-5 Transformations

No.	Transformation	Function
1.	Add New Feature	A new feature can be added to the dataset which could be derived from the existing features using Script. Physical Feature Name and Feature Name are the names of the new feature. Script can be used to create a pandas Series for the new feature
2.	Encode Categorical Features	This function performs One Hot Encoding on a categorical feature and replaces it with multiple numerical features in the dataset.
3.	Encode Datetime Features	This function encodes a datetime feature and replaces it with multiple numerical features having the following information derived from the datetime feature - year, month, week, day, hour, minute, dayofweek.

Table 6-5 (Cont.) Transformations

No.	Transformation	Function
4.	Encode Cyclical Features	This function encodes a cyclical feature having hour, minute data, and so on and returns two features carrying the sine and cosine transformation of the cyclical data. 'fmax ' denotes the maximum possible value of the cyclical feature data.
5.	Impute Missing Data	This function imputes missing data within a feature. For numerical features, there are 4 methods for imputing missing data. simple - imputes with mean, median, most_frequent values based on chosen arg value using the SimpleImputer in sklearn. const - fills the missing values with the value given in the arg knn - imputes using the KNNImputer in sklearn with k value given in arg mice - imputes using the IterativeImputer in sklearn For non-numerical data, missing values can be imputed using the 'const' method by replacing all missing values with the value given in arg
6.	Feature Scaling	This function is used to scale multiple selected numerical features using the StandardScaler in sklearn
7.	Dimensionality Reduction	This function performs PCA on selected numerical features to reduce the dimensionality using sklearn.decomposition.PCA module. The number of output features can be specified using dim field. The names of the output features' names can be specified in the fields 'Physical Feature Name' and 'Feature Name'
8.	Outlier removal	This function is used to remove outliers present in a feature based on the specified zscore value. Non-numerical features are label encoded before removing the outliers.
9.	Duplicates Removal - Data Frame	This function removes all duplicate rows in the dataframe.
10.	Duplicates Removal - Feature	This function removes all duplicate rows within a specified subset of features and consequently removes those rows from the data frame

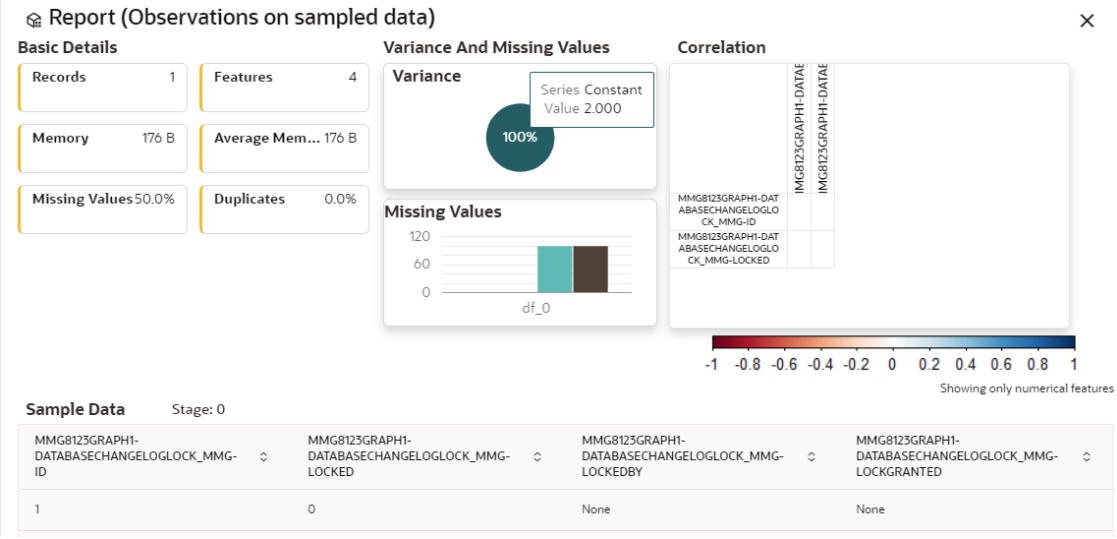
Table 6-5 (Cont.) Transformations

No.	Transformation	Function
11.	Filter Features	This function is used to filter the data frame based on conditions specified on features Operations allowed : >,>=,=,!=,<,<=,isin When the chosen operation is 'isin', the input to 'Filter Value' is a list of values that should be present in the output data frame

6.1.2.1.4.5 View Reports

This report shows the Observation on Sampled data, profiles, Variance, Correlation Matrix (correlation between columns), Sample Data. This report is based on sampled data.

Figure 6-14 Sampled Data Report



6.1.2.1.4.6 Re-ordering of Transformations

You can re-order the transformations using the drag-drop. During the transformation re-order, you can compare the profile of transformations. The transformation order gets adjusted accordingly.

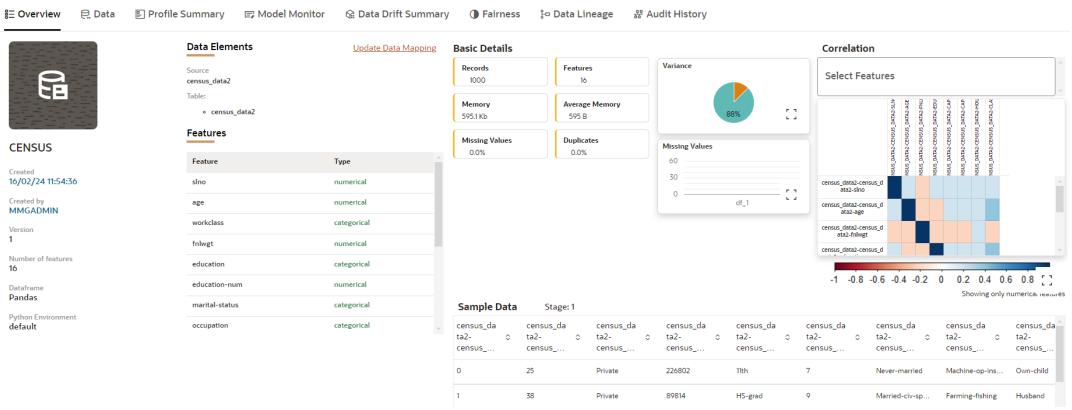
6.1.3 View a Dataset

To view a Dataset, follow these steps:

1. Navigate to Dataset Summary page.
2. Click **Action** next to corresponding Dataset and select **View**.

   The Profile Summary screen is displayed.

Figure 6-15 Dataset Details screen



6.1.3.1 Overview

The details such as Data Elements, Features, Basic Details, Correlation, and Sample Data are displayed in graphical format.

To view the Overview of a Dataset, follow these steps:

1. Navigate to **Dataset Summary** page.
2. Click the **Next** to corresponding Dataset and select **View**.

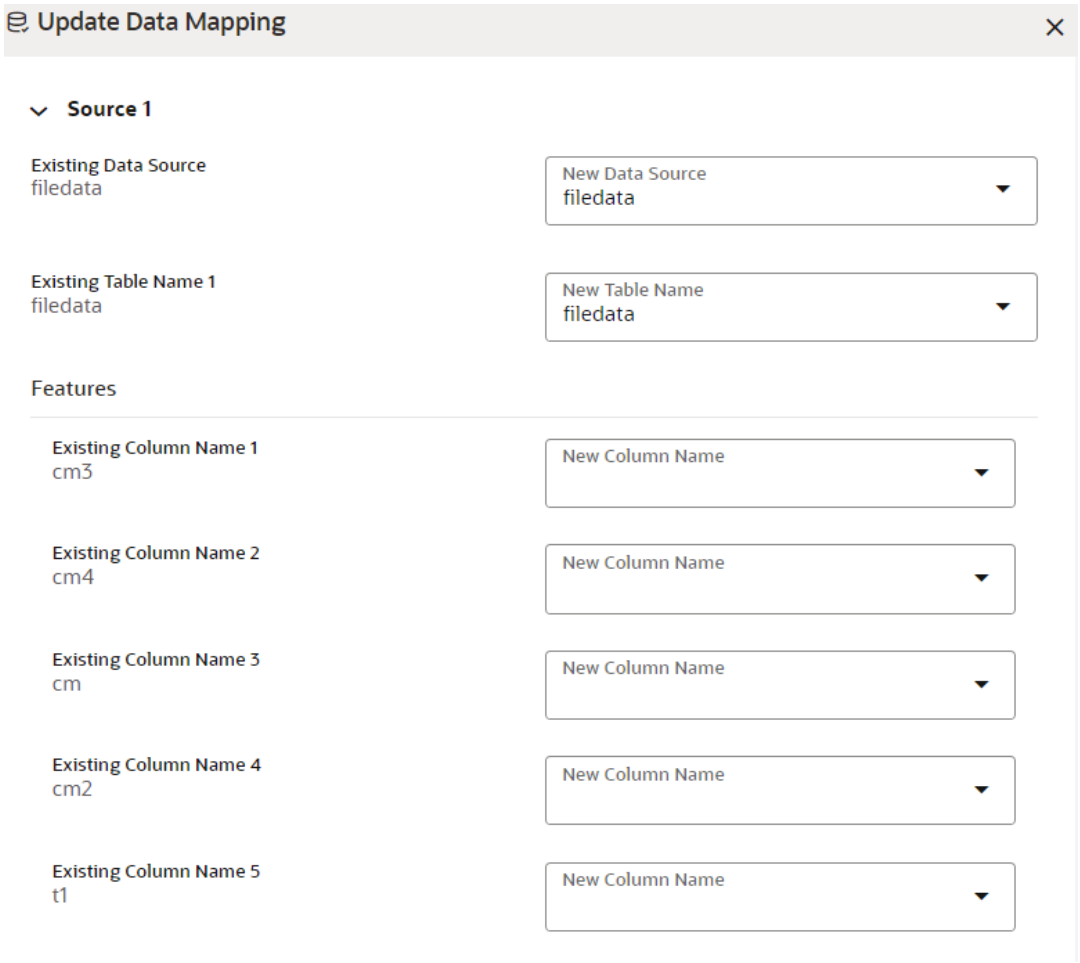
   OR

   Click on the Dataset **Name** column which you want to view.

   The Overview page is displayed.

   You can update the data mapping details by clicking on Edit icon in Data Elements group.

Figure 6-16 Update Data Mapping



6.1.3.2 Data

To view the data of a Dataset, follow these steps:

1. Navigate to the **Dataset Summary** page.
2. Click **Next** to the corresponding Dataset and select **View**.

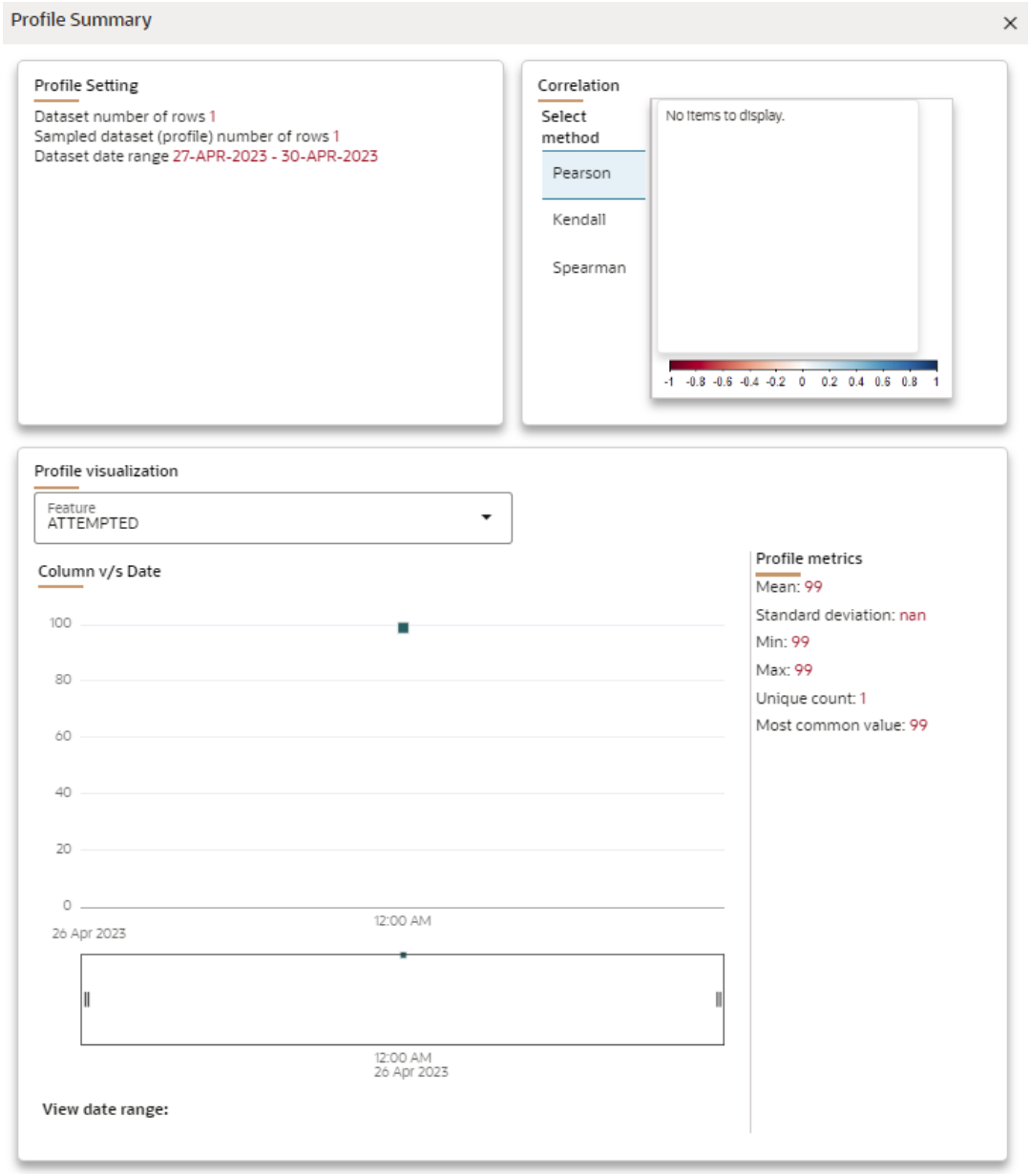
   The Overview page is displayed.
3. Navigate to **Data** tab.

6.1.3.3 Profile Summary

To view the Profile Summary, perform the following steps:

1. In the Profile Summary screen, click next to corresponding data profile and select View Summary.
   This window displays the following details in graphical format:
   - Profile Setting
   - Correlation
   - Profile visualization along with the metrics

Figure 6-17 Profile Summary



6.1.3.3.1 Start Profiling

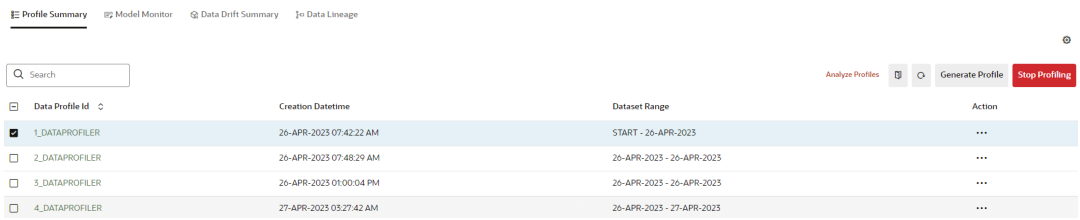
Prerequisite: You must create a Dataset with a datasource having columns with **DATA_TYPE** as **DATE**.

To start a Profile:

1. Navigate to Dataset Summary page.
2. Click **Action** next to corresponding Dataset and select **View**.

   This window displays the Dataset profiles in a table.

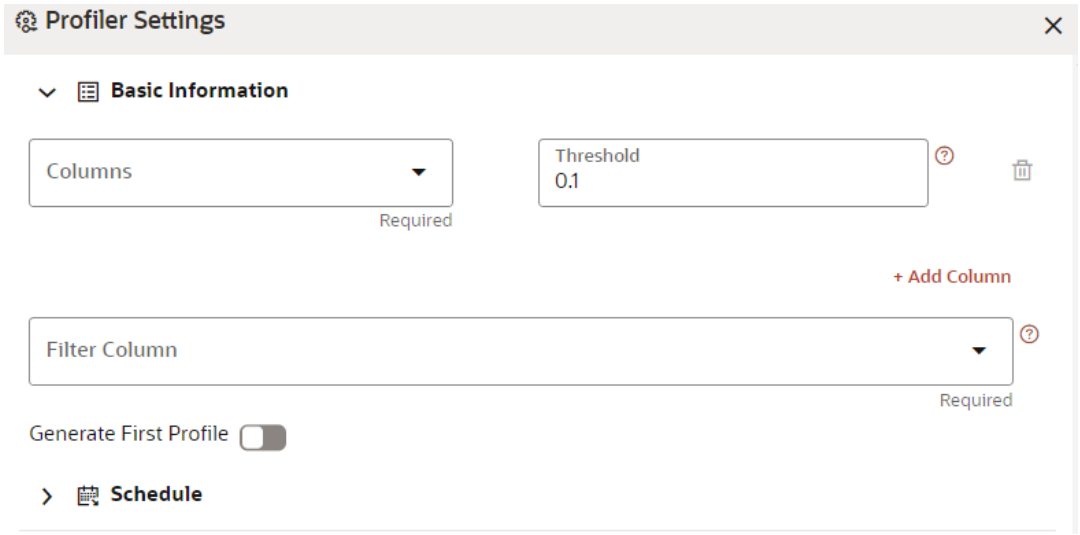
Figure 6-18 Dataset Profiles



3. Click **Start Profiling** or click **Settings** .

   The Profiler Settings page is displayed.

Figure 6-19 Profiler Settings page



4. Under Basic Information section, enter the following details:
   a. Select the required columns from the drop-down. All the numeric options are displayed under the Columns drop down. You can select multiple columns by clicking on + icon.
   b. Enter the threshold for the columns. Threshold defines the Drift detection threshold for each column and only numerical columns can be selected for profiling.
   c. Select the required filter column from the drop-down. Filter column contains date column of the dataset to filter.
   d. Enable **Generate First Profile** option if you want to generate a profile.
5. Under Schedule section, enter the following details:
   a. Enter a Schedule Name.
   b. Select the Schedule Type as required:
      - **Daily**: Select to run the schedule everyday.
      - **Weekly**: Select to run the schedule once in a week or the selected days in a week.
      - **Monthly**: Select to run the schedule once in a month or the selected days in a month.
   c. Click **Save**.

   The Data Profiling has been started and Stop Profiling option is displayed in the Profile Summary screen.

For more details on disabling dataset profile, see Stop Profiling section.

6. Click **Generate Profile**.

   The Data Profile is created.

   Schedule Daily

   To schedule the profiler to run daily, perform the following steps:

   a. In the **Profiler Settings** screen, select the Schedule Type as **Daily**.
   b. Select the start date on which you want to run the profiler.
   c. Select the end date on which you want to stop your schedule.
   d. Enter the time at which you want to run the profiler.
   e. Click **Save**.

   Schedule Weekly

   To schedule the profiler to run weekly, perform the following steps:

   a. In the **Profiler Settings** screen, select the Schedule Type as **Weekly**.
   b. Select the start date on which you want to run the profiler.
   c. Select the end date on which you want to stop your schedule.
   d. Enter the time at which you want to run the profiler.
   e. Select the day on which you want to run the profiler. You can select multiple days to run the profiler.
   f. Click **Save**.

   Schedule Monthly

   To schedule the profiler to run monthly, perform the following steps:

   a. In the **Profiler Settings** screen, select the Schedule Type as **Daily**.
   b. Select the start date on which you want to run the profiler.
   c. Select the end date on which you want to stop your schedule.
   d. Enter the time at which you want to run the profiler.
   e. Select the month on which you want to run the profiler. You can select multiple months to run the profiler.
   f. Select the date on which you want to run the profile of the selected month to run.
   g. Click **Save**.

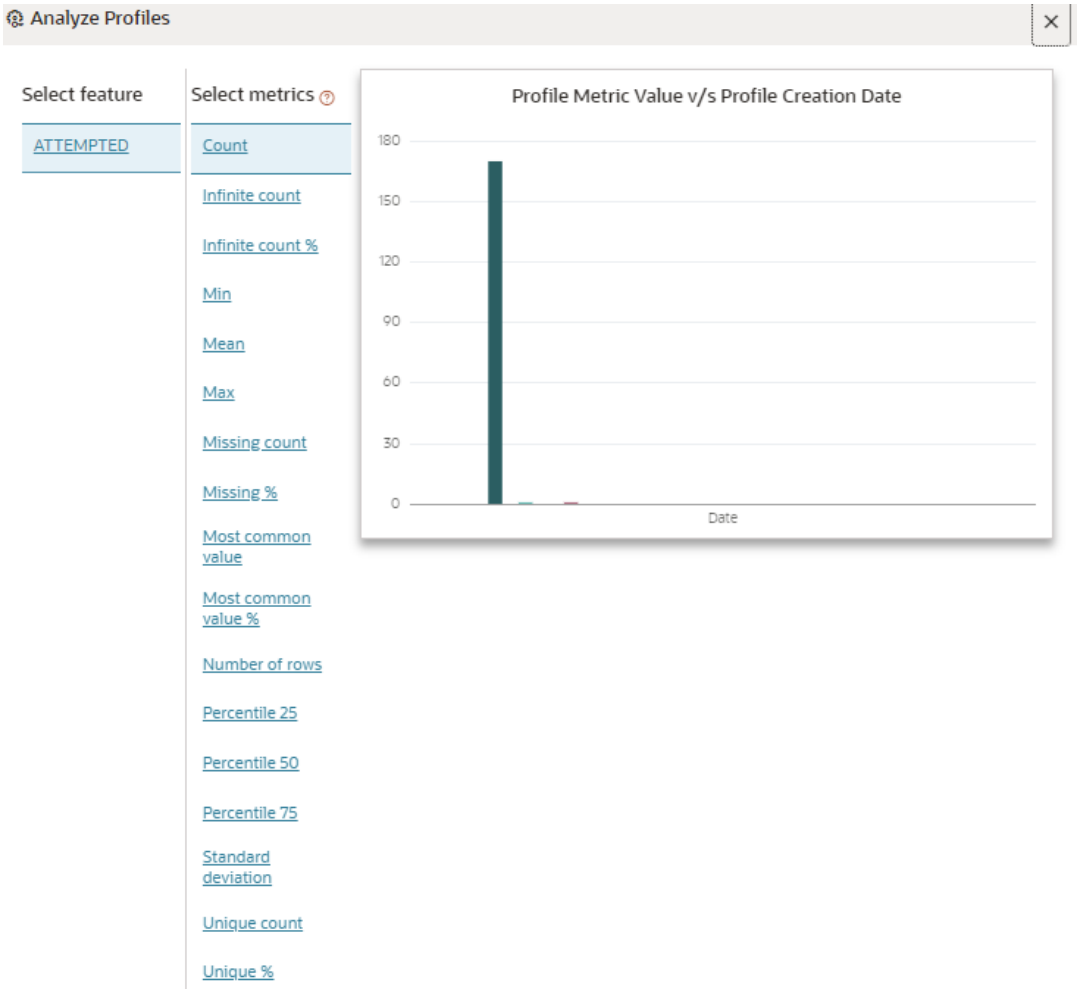
6.1.3.3.1.1 Analyze Profiles

To analyze multiple data profiles, perform the following steps:

1. In the Profiler Summary screen, select the data profiles that you want to analyze profiles. You can select multiple data profiles.
2. Click **Analyze Profiles**.

   A window is displayed with multiple metrics of the selected profiles. Click on each metric to view the data in graphical format.

Figure 6-20 Analyze Profiles screen



6.1.3.3.1.2 Compare Profiles

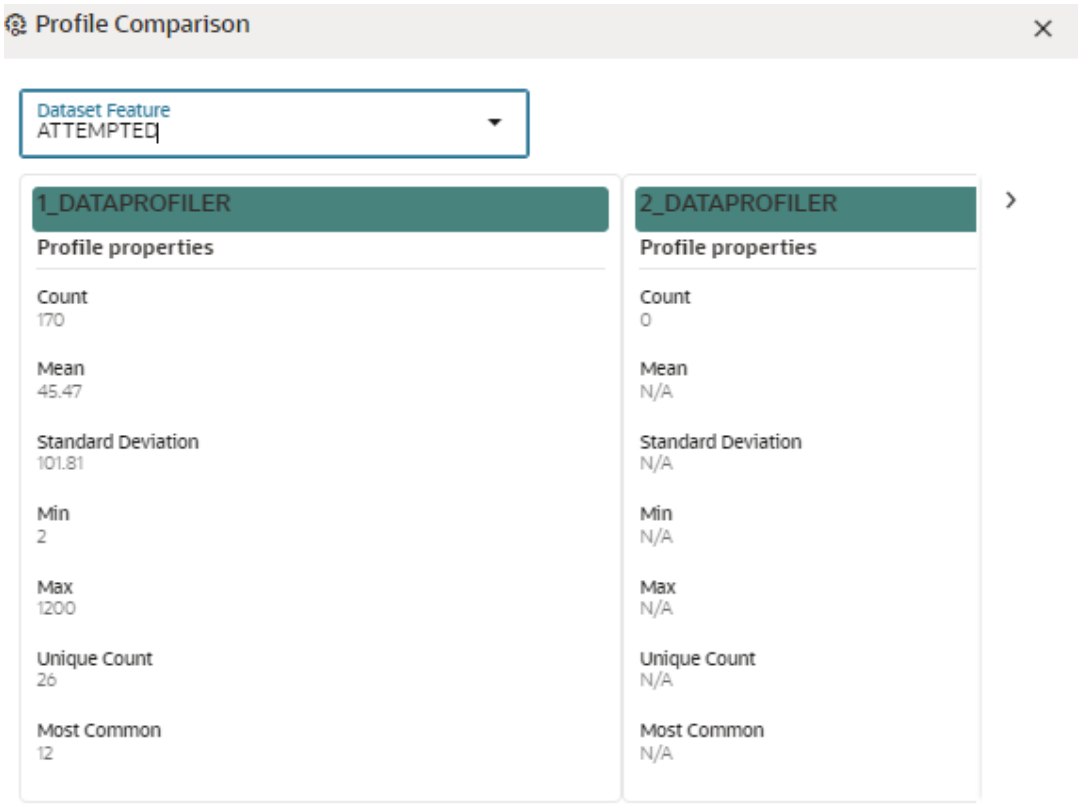
To compare the data between two data profiles, perform the following steps:

1. In the Profiler Summary screen, select two data profiles that you want to compare.
2. Click **Compare Profiles**.

   The Profile Comparison is displayed.
3. In the Dataset Feature drop-down, select the feature that you want to compare.

   The properties of both the profiles is compared and displayed in a tabular format.

Figure 6-21 Profile Comparison screen



6.1.3.3.1.3 Stop Profiling

To stop the profiling, perform the following steps:

1. In the Profiler Summary screen, select the data profiles that you want to stop profiling. You can select multiple data profiles.
2. Click **Stop Profiling**.

   A confirmation pop-up window is displayed.
3. Click **Yes**to disable the dataset profile.

   The dataset profiling has been disabled and Start Profiling option is displayed. For more information on profile, see Create a Profile section.

6.1.3.4 Model Monitor

Model Monitoring refers to the process of closely tracking the performance of models in production and helps you to frequently monitor the distribution of the model and provides alerts in case of any exceptions.

In addition, it enables you to identify and eliminate bad quality predictions and poor technical performance of the models.

The model's robustness depends not only on the training of the feature-engineered data but also on how well the model is monitored after deployment. Typically a model's performance degrades over time, and it essential to detect the cause of the decrease in performance of the model. The main cause of the decline in performance can be drift in the independent or/and dependent features which may violate the model’s assumption and distribution about the data. When models are built, the model builder will create a snapshot of the dataset used for training

the model and save it in a file which is later used for calculating drift with the current snapshot of the dataset.

To start the model monitoring, perform the following steps:

1. In the LHS menu, click **Dataset**to display the **Dataset Summary** window.

   This window displays the dataset records in a table.

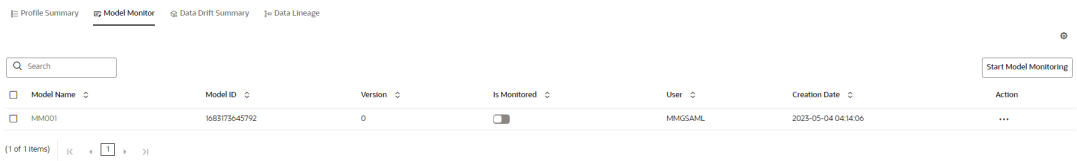
2. Click **Action**next to corresponding Dataset and select **View**.

   The Profile Summary screen is displayed.

3. Navigate to **Mode Monitor**page.

   All the models associated with the selected dataset will be displayed.

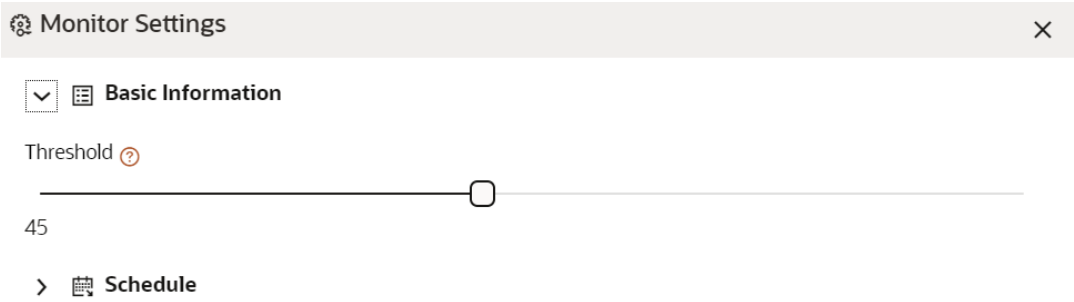
Figure 6-22 Mode Monitor page



4. Click **Start Model Monitoring**or **Monitor Settings** icon.

   The Monitor Settings page is displayed.

Figure 6-23 Monitor Settings page



5. Drag the slider and set the threshold settings. Threshold is the value in percentage, beyond which it is labelled as drifted.
6. Select the Schedule Type as required:
   - Daily: Select to run the schedule everyday.
   - Weekly: Select to run the schedule once in a week or the selected days in a week.
   - Monthly: Select to run the schedule once in a month or the selected days in a month.

   The monitoring for the model is started and Is Monitored option is enabled by default and you can disable it if you do not want to monitor.

6.1.3.4.1 Stop the Model Monitoring

To stop the model monitoring, perform the following steps:

1. Navigate to **Model Monitor** page.

   All the models associated with the selected dataset will be displayed.

2. Click **Stop Model Monitoring**.

   A confirmation pop-up window is displayed.

3. Click **Yes**.

   The Model Monitoring is stopped.

6.1.3.4.2 Viewing the Model Summary

To view the Model Summary, perform the following steps:

1. Navigate to **Model Monitor** page.

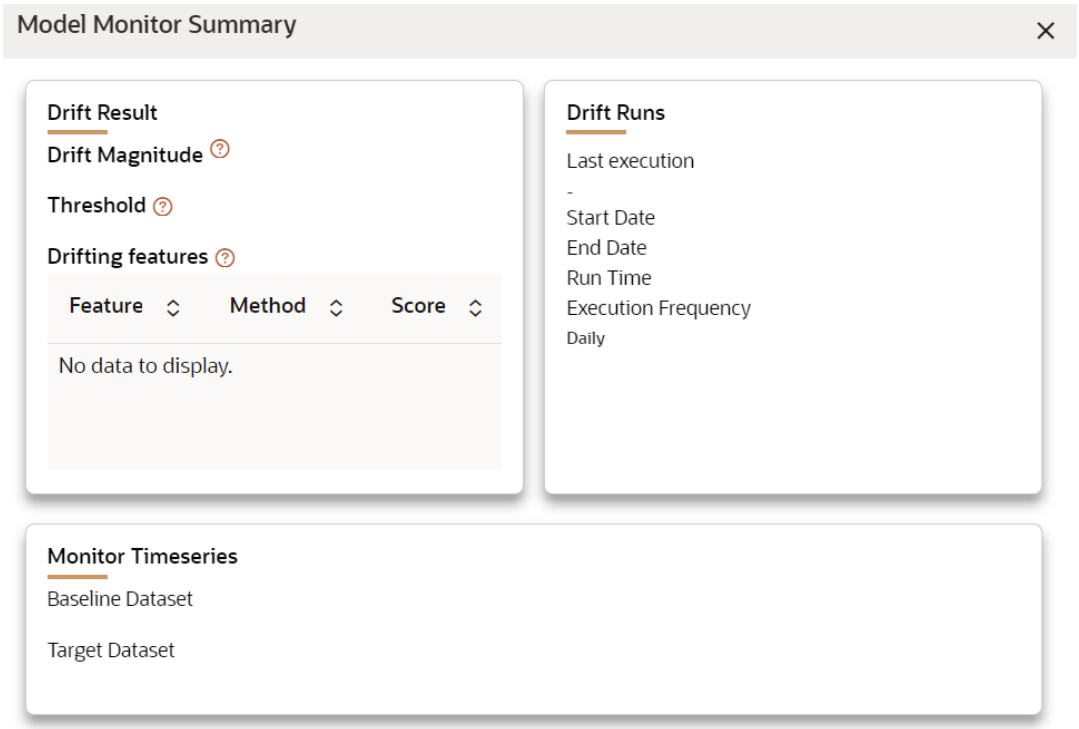
   All the models associated with the selected dataset will be displayed.

2. Click **Action** next to corresponding model and select **View Summary**.

   This window displays the following details in graphical format.

   - Drift Result
   - Drift Runs
   - Monitor Timeseries

Figure 6-24 Model Monitor Summary



6.1.3.4.3 Viewing the Model Latest Report

To view the Model Latest Report, perform the following steps:

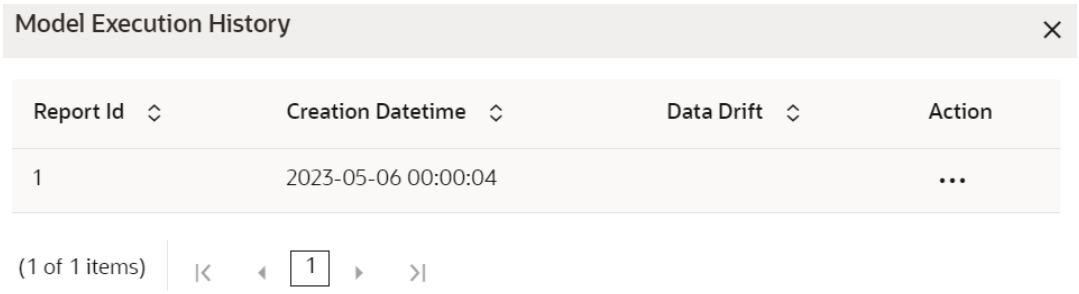
1. Navigate to **Model Monitor** page.

   All the models associated with the selected dataset will be displayed.

2. Click **Action** next to corresponding model and select **View Latest Report** .

   The execution details of the model are displayed.

Figure 6-25 Model Execution History



To view the report of each execution, click **Action** next to corresponding report and select **View Report**.

6.1.3.4.4 Viewing the Model Execution History

To view the Model Execution History, perform the following steps:

1. Navigate to **Model Monitor**page.

   All the models associated with the selected dataset will be displayed.

2. Click **Action**next to corresponding model and select **View Execution History**.

   This window displays the details such as Baseline Distribution, Target Distribution, Drift Method, Drift Score, and Data Drift of the models.

6.1.3.5 Data Drift Summary

A Data Drift Summary report summarizes the changes in a dataset over time. It typically includes information for any changes in the data distribution of the selected features over the selected baseline and target date ranges.

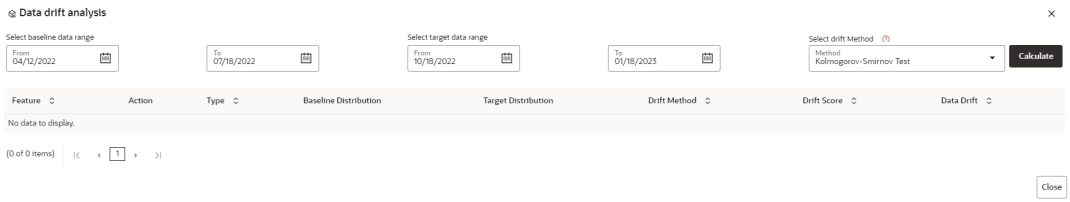
The report is used to help identify potential issues with data quality which helps the users to take necessary actions further.

To calculate the data drift, perform the following steps:

1. In the Data Drift Summary screen, click Calculate Drift.

   The Data Drift Analysis window is displayed.

Figure 6-26 Data Drift Summary screen



2. Select the data range (From and To) dates for the baseline.
3. Select the data range (From and To) dates for the target.
4. Select the drift method from the drop-down.

   The available drift methods are:

- Kolmogorov–Smirnov (K-S) test
  - Only for numerical features
  - Output: p_value, drift detected when p_value < threshold.
- Kullback-Leibler divergence
  - For numerical and categorical features
  - Output: divergence, drift detected when divergence >= threshold.
- Wasserstein distance (normed)
- Only for numerical features
- Output: distance, drift detected when distance >= threshold.
- Population Stability Index (PSI)
  - For numerical and categorical features
  - Output: psi_value, drift detected when psi_value >= threshold.
- Jensen-Shannon distance
  - For numerical and categorical features
  - Output: distance, drift detected when distance >= threshold.

5. Click **Calculate**.

   The drift analysis report is displayed.
6. Click **Save**.

   The reports are displayed under Data Drift Summary screen.

6.1.3.5.1 Disabling the Data Drift Monitoring

To stop the Data Drift monitoring, perform the following steps:

1. Navigate to **Data Drift Summary**page.

   The data drift summary reports are displayed in a table.
2. Click **Stop Data Monitoring**.

   A confirmation pop-up window is displayed.
3. Click **Yes**.

   The Data Drift Monitoring is disabled.

6.1.3.5.2 Start the Data Drift Monitoring

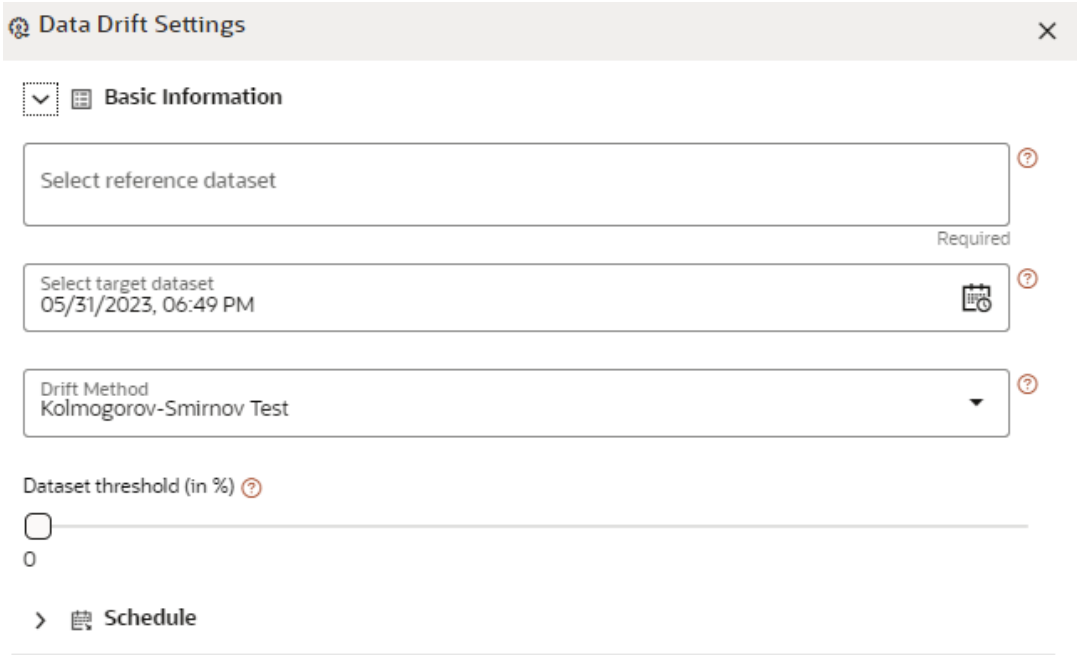
To start the Data Drift monitoring, perform the following steps:

1. Navigate to **Data Drift Summary**page.

   The data drift summary reports are displayed in a table.
2. Click **Start Data Monitoring**or **Data Drift Settings** icon.

   The Data Drift Settings window is displayed.

Figure 6-27 Data Drift Settings



3. Enter the basic information and schedule details and click **Save**.

   For more details, see Data Drift Summary section.

   The Data Drift Monitoring is started.

6.1.3.5.3 Viewing the Data Drift Summary Report

To view the Data Drift Summary Report, perform the following steps:

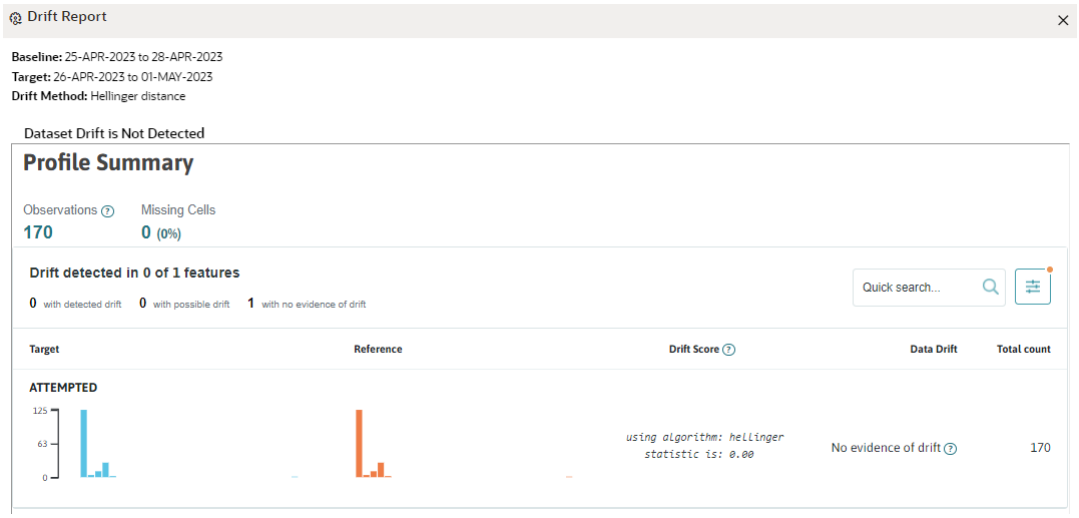
1. Navigate to **Data Drift Summary** page.

   The data drift summary reports are displayed in a table.

2. Click **Action** next to corresponding drift report Id and select **View**.

   The drift report is displayed in a graphical format.

Figure 6-28 Drift Report



6.1.3.5.4 Deleting the Data Drift Summary Report

To delete the Data Drift Summary Report, perform the following steps:

1. Navigate to **Data Drift Summary**page.

   The data drift summary reports are displayed in a table.

2. Click **Action** next to corresponding drift report Id and select **Delete** .

   OR

   If you wish to delete multiple data drift reports, select reports and click **Delete**.

   A confirmation pop-up window with report details is displayed.

3. Click **Delete**.

   The Data Drift Report is deleted.

6.1.3.6 Fairness

You can calculate the fairness metrics of the dataset by choosing the target variable and protected features. The module provides metrics dedicated to assessing and checking whether the model predictions and/or true labels in data comply with a particular fairness metric. For this example, the statistical parity metric also known as demographic parity, measures how much a protected group's outcome varies when compared to the rest of the population. Thus, such fairness metrics denote differences in error rates for different demographic groups/ protected attributes in data. Therefore, these metrics are to be *minimized*to decrease discrepancies in model predictions with respect to specific groups of people. Traditional classification metrics such as accuracy, on the other hand, are to be maximized.

You can perform the following functions:

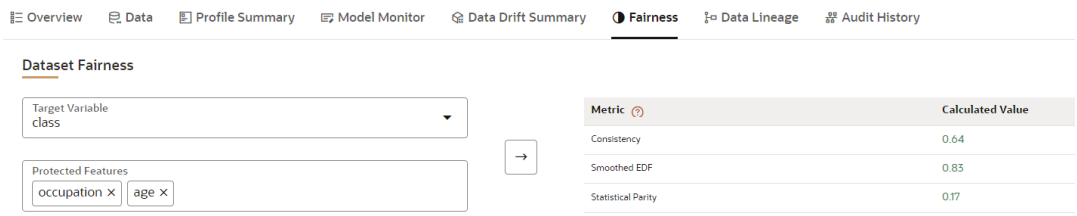
- Measure Fairness Metrics of Dataset
- Models Bias Mitigation of Models
- Privacy Estimation of Models

For Example: In model pipeline, a new widget can be added to calculate the fairness metrics of a trained model or an advanced option can be added to model training widget to calculate the fairness metric of the model after training.

To calculate the dataset fairness, perform the following steps:

1. In the Fairness screen, select the Target Variable from the drop-down.
2. In the Protected features, select the required features to understand the fairness of the dataset.

Figure 6-29 Fairness screen



The metric and calculated values are displayed. Click Help icon for more details on how the fairness is calculated.

6.1.3.7 Data Lineage

Data Lineage provides a summary view or visualization of Data Sources and its associated dataset and features in a graphical format. Also displays the complete data flow from start to finish.

To view the data lineage, perform the following steps:

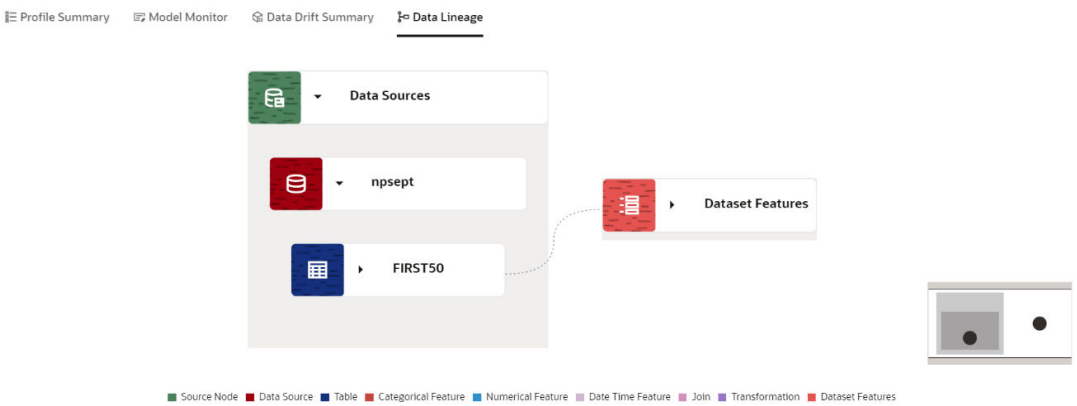
1. In the LHS menu, click **Dataset** to display the **Dataset Summary** window.

   This window displays the dataset records in a table.
2. Click **Action** next to corresponding Dataset and select **View**.

   The Profile Summary screen is displayed.
3. Navigate to **Data Lineage** page.

   A summary view of Data Sources and its associated dataset and features in a graphical format.

Figure 6-30 Data Lineage page



Click the drop-down of the respective features to view granular details and its workflow.

6.1.3.8 Audit History

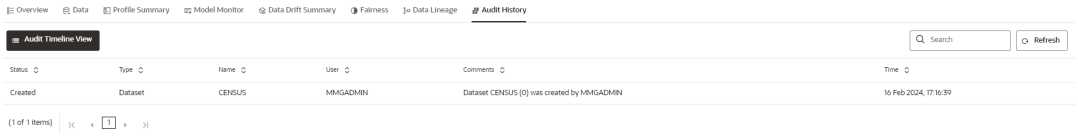
At any time, you can audit the datasets from the Audit History page. The Audit Trail window provides the complete details of datsets. The sequence of actions performed in the dataset lifecycle is listed in the table view and the timeline view (graphical representation).

To audit datasets:

1. Click **Launch Workspace** next to corresponding Workspace to Launch Workspace to display the **Dashboard** window with application configuration and model creation menu.
2. Navigate to **Dataset Summary** page.
3. Click More Options icon next to corresponding Dataset and select View.

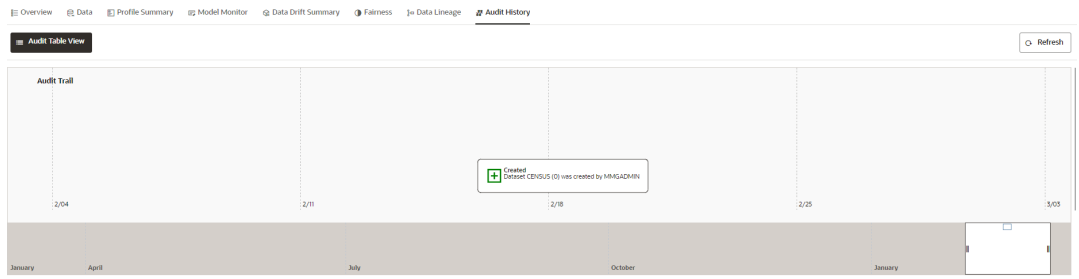
   The **Overview** screen is displayed.
4. Navigate to **Audit History** page.

Figure 6-31 Audit History



Click for **Audit Timeline View** for Timeline View and **Audit Table View** to switch to the regular view. Click to **Refresh**to refresh the Audit History. You can search the datasets by entering the name in the search box.

Figure 6-32 Audit History window Timeline View with Horizontal Time Axis



6.1.3.9 Create a Cache Snapshot

You can create a Cache Snapshot of the Dataset from this page.

To create a cache snapshot, follow these steps:

1. Navigate to Dataset Summary page.
2. Click next to corresponding Dataset and select View.

   The Overview page is displayed.
3. Click on the header and select Create a cache snapshot.

4. Enter the **Snapshot Name** having alphanumeric character, underscore and hyphen not exceeding 30 characters and click Create.

   This screen additionally allows users to cache a snapshot of the current state of data. The cached snapshots can be accessed later in Model Pipelines without any recompilation or re-read from data sources.

   To use the dataset in model pipeline or data pipeline, the actual data is fetched using the Cache option. For example, to take the data from dataset on As of Date, create the data frame and provide the name to cache. Only Cache pulls the data from dataset.

   This helps the things to work faster when you have millions of records, and you want to use intermediate data for use. For example, if you have 1 million records and want to use only 10,000 out of that, then perform the sampling for that 10, 000 entries. This increases the speed of processing, validation of information.

   - Once the metadata is created, the original data can be cached. A snapshot of the actual data in the dataset at the current time can be stored referenced by a tag name.
   - The location for caching is in the datastudio server location `$DS_HOME/work/ftpshare/mmg/workspace_name/dscode/tag`. The dataset will be saved as a parquet file with name dscode_tag.parquet
   - When executing all APIs from notebook, workspace has to be attached.
   - Caching can be performed in two ways:
     - From UI, immediately after saving the metadata.

   Or the users will have to fetch(create) a new snapshot/dataframe of the dataset using the API 'Fetch New Snapshot of dataset' and manually cache using the 'Caching Data Frame' API.

   The following table provides information on the error and the troubleshooting procedure in case of dataset failure.

Table 6-6 Information on the error and the troubleshooting procedure

Error	Troubleshooting procedure
`ModuleNotFoundError: No module named '_bz2'`	Install the package 'libbz2-dev' before building python.
`ModuleNotFoundError: No module named '_sqlite3'`	Install the package 'libsqlite3-dev' before building python.
`Python-env-health-check fails if pandas version is less than 1.4.1.` **NOTE:** `modin dataframe library is supported only from pandas version 1.4.1 and higher.`	You can switch between the options modin[dask] and pandas for the underlying dataframe library if pandas version 1.4.1 is installed.
"`Not a valid file`" errorwhile profiling the Hive Data sources.	Copy the following required files : kbank.keytab , krb5.conf , hive-jdbc-driver.jar into the path : $DS_HOME/conf folder of datastudio

6.1.4 Edit a Dataset

To edit a Dataset, follow these steps:

1. Navigate to **Dataset Summary** page.
2. Click on the Dataset Name column which you want to edit.

The Overview page is displayed.

3. Click **Edit**.

   You can edit the Dataset fields except Code and Dataset Name.

6.1.5 Delete a Dataset

To delete a Dataset, follow these steps:

1. Navigate to **Dataset Summary**page.
2. Click next to corresponding Dataset and select **Delete**.

   OR

   Click next to corresponding Dataset and select **View**. You can delete a Dataset from this page.

6.1.6 Python functions for accessing Dataset from Model Pipelines/ Notebooks

This section provides information on the Python functions for accessing Dataset from Model Pipelines/Notebooks.

6.1.6.1 List tags of Datasets Cached from UI

To get a list of all snapshots/tags of a dataset cached from UI, use the API below:

Table 6-7 Listing tags of Datasets Cached from UI

API	Notebook Script	Notes/Input
List Tags of a Dataset	from mmg.datasets import list_tags list_tags(dscode)	dscode = dataset code - string tag = tag with which the dataset was cached - string

6.1.6.2 Delete Cached Dataset

A Cached Dataset can be deleted by using the API below where tag refers to a particular snapshot/timestamp of the dataframe.

Table 6-8 Delete Cached Dataset

API	Notebook Script	Notes/Input
Delete Dataset	from mmg.datasets import delete_df delete_df(dscode, tag)	dscode = dataset code - string tag = tag with which the dataset was cached - string

6.1.6.3 List all Datasets with Metadata saved from UI

Table 6-9 List all Datasets with Metadata saved from UI

API	Notebook Script	Notes/Input
List Datasets	from mmg.datasets import list_datasets list_datasets()	

6.1.6.4 Fetch Dataset in Notebook

Dataset whose metadata has been saved from UI can be fetched in two ways:

1. Dataset that has already been cached from UI can be fetched using first API below by providing the dataset code and tag.(DataFrame will not be recalculated. It will be read from cache.)
2. A new data frame for a dataset can be calculated/fetched using data from the present time with the metadata saved for that dataset as given in second API.

Table 6-10 Fetching Dataset in Notebook

API	Notebook Script	Notes/Input
Fetch Dataset Cached from UI	from mmg.datasets import fetch_ds df = fetch_ds(dscode,tag)	dscode = dataset code - string tag - string
Fetch New Snapshot of Dataset	from mmg.datasets import fetch_ds df = fetch_ds(dscode)	dscode = dataset code - string

6.1.6.5 Cache User's Dataframe from Notebook

A user-made data frame can also be cached using the below API. It will be stored in the datastudio server in location : `$DS_HOME/work/ftpshare/mmg/workspace_name/cached/tag`. The dataset will be saved as a parquet file with name `cached_tag.parquet`.

> **Note:**
>
> This dataframe is not related to dataset created from UI. This is independent of dataset metadata.

Table 6-11 Cache User's Dataframe from Notebook

API	Notebook Script	Notes/Input
Caching Data Frame	from mmg.datasets import cache_df path=cache_df(df, tag) path	df = dataframe to be cached - pandas dataframe tag - string

6.1.6.6 Fetching Data Frame Cached from Notebook

The data frame cached from notebook can be fetched using the below API.

Table 6-12 Fetching Data Frame Cached from Notebook

API	Notebook Script	Notes/Input
Fetch dataset cached manually	from mmg.datasets import fetch_ds df = fetch_ds("cached",tag)	tag - string

6.1.6.7 List Tags of all Data Frames Cached from Notebook

Table 6-13 List Tags of all Data Frames Cached from Notebook

API	Notebook Script	Notes/Input
List Tags of Manually Cached Datasets	from mmg.datasets import list_tags list_tags()	

6.2 Model Libraries

The Model Library information is used to bind a particular technique and its details to one unique Model Library. The Starting point for the Model builder is to register a model library with the application after it is properly set up in the python environment to be used.

Currently, the following libraries are supported:

- keras
- ONNX
- scikit-learn
- xgboost

As an example, the following section provides information on how to set up and register the scikit-learn library.

Setting up the python library in the python environment:

Open the terminal and install the scikit-learn library using the following command:

```
python3 -m pip install scikit-learn
```

> **Note:**
>
> Proxy might be required to install the packages from pip.

Once the installation is complete, check for the package details using the following command:

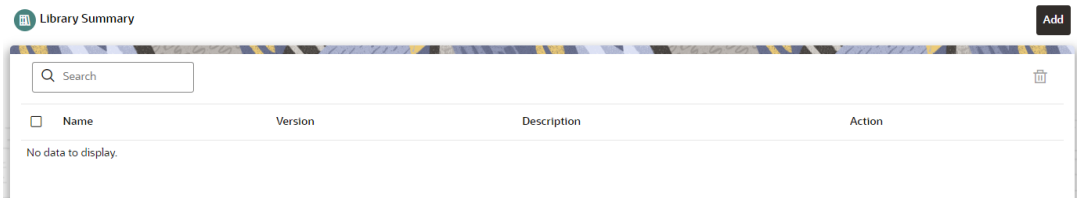
```
pip show scikit-learn
```

Once the installation is complete in python environment, you need to register the library into the application.

Registering the python library

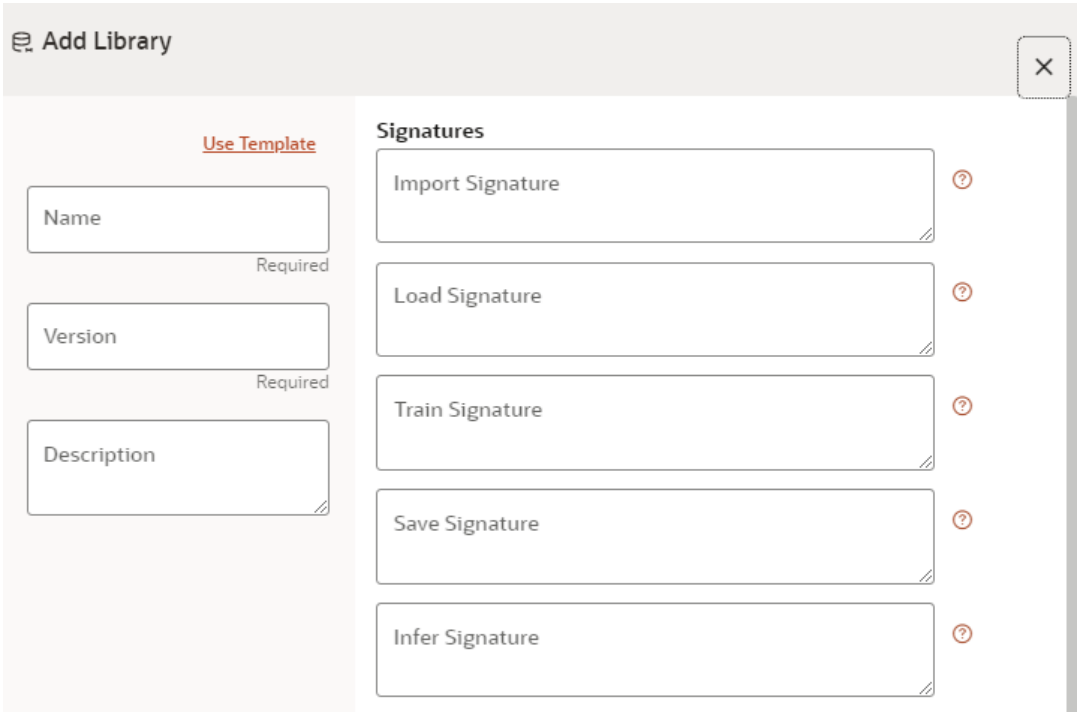
1. In the Mega menu, click Modeling > Model Libraries.
   The **Library Summary** page is displayed.

Figure 6-33 Library Summary page



2. In the **Library Summary** page, click **Add**.
   The **Add Library** page is displayed.

Figure 6-34 Add Library page



3. Enter the name of the library.
4. Enter the version and description of the library.
5. Enter the signatures such as Import, Load, Train, Save, and Infer. For more details on the signatures, click the respective help icons.
   Some of the signatures captured in library stage might not be standard across different algorithm/techniques provided by the library.
6. Click **Create**.
   The Model Objective is created and displayed in the Model Objective screen.

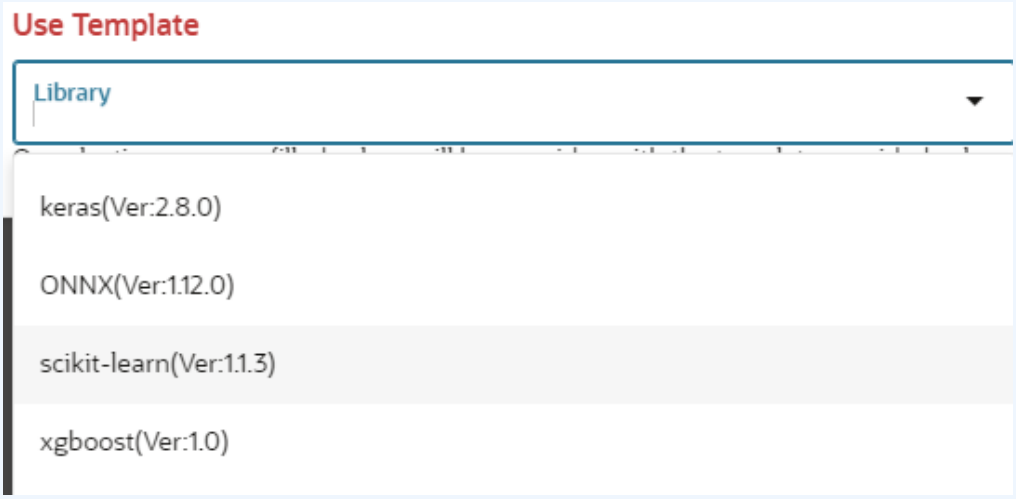
Note:

The User must check the details provided above from the home page of the library being captured.
For example: `https://scikit-learn.org/stable/`.

OR

You can select the **Use Template** option to pre-fill the entries from the seeded list of Libraries. Currently, the following libraries are present in the template.

Figure 6-35 Use Template



You can also edit the data based on your requirements.

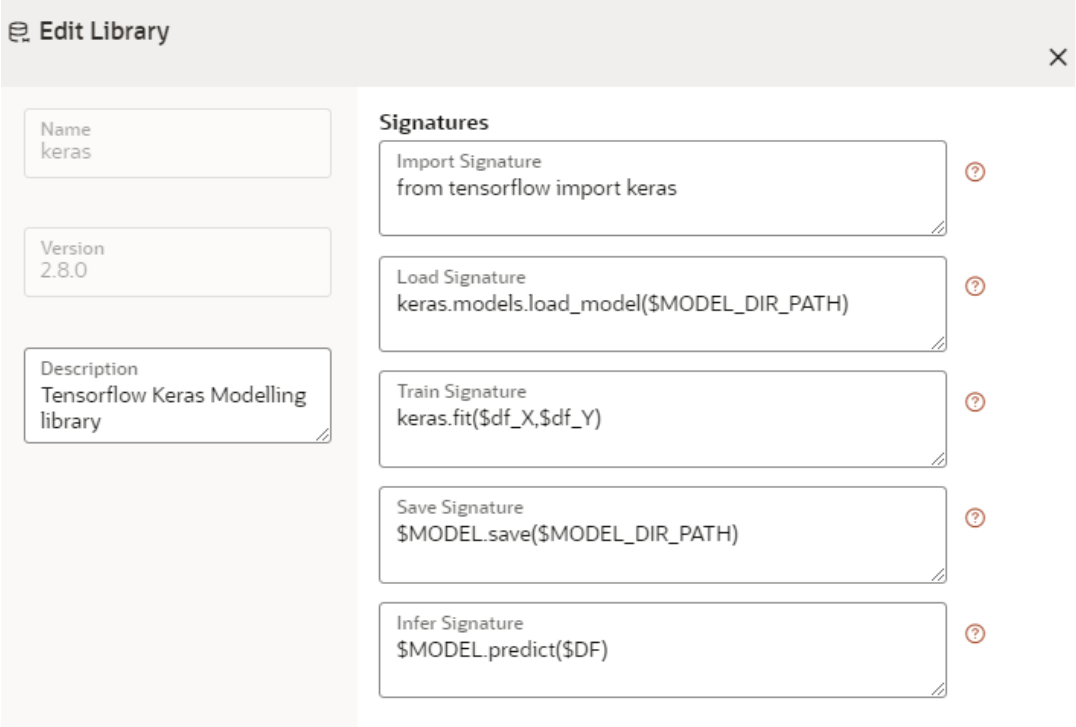
6.2.1 Edit a Model Library

To edit a Model Library:

1. Navigate to **Library Summary** page.
2. Click and select Edit.

   The **Edit Library** screen is displayed.

Figure 6-36 Edit Library screen



3. Make the necessary changes and click **Update**.

> **Note:**
>
> The Model Library name and version are unique and cannot be modified.

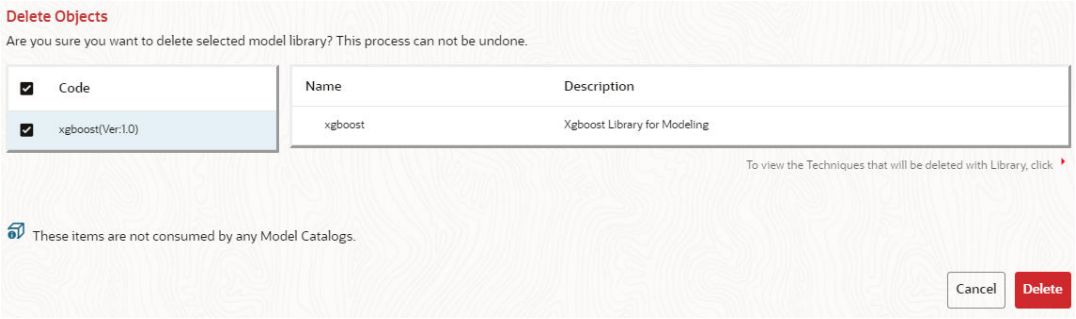
6.2.2 Delete a Model Library

To delete a Model Library:

1. Navigate to **Library Summary** page.
2. Click and select **Delete**.

   A confirmation message is displayed.

Figure 6-37 Confirmation Message



3. Click **Delete**.

   The selected model library and all the associated model techniques are deleted. The libraries that are deleted cannot be consumed by any model catalog.

   To delete multiple model libraries select all the model libraries that you want to delete and click the icon on the page header.

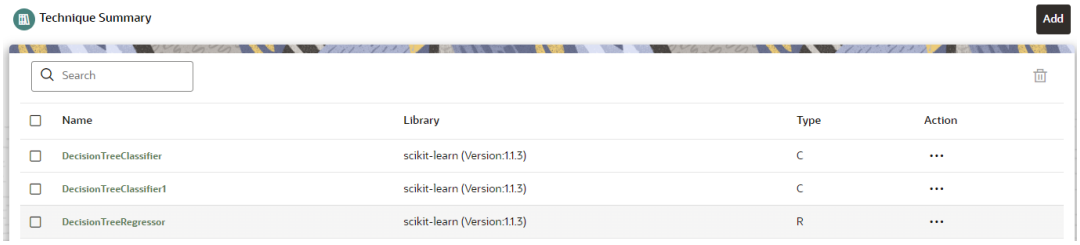
6.3 Model Technique

Model Technique is the algorithm/technique used to create python model using the library/ package which was created in the Model Library Screen. It is the actual information captured in the application that helps in training the model (Upload and Build).

1. In the LHS menu, click **Model Catalog > Model Technique**option.

   The **Technique Summary**page is displayed.

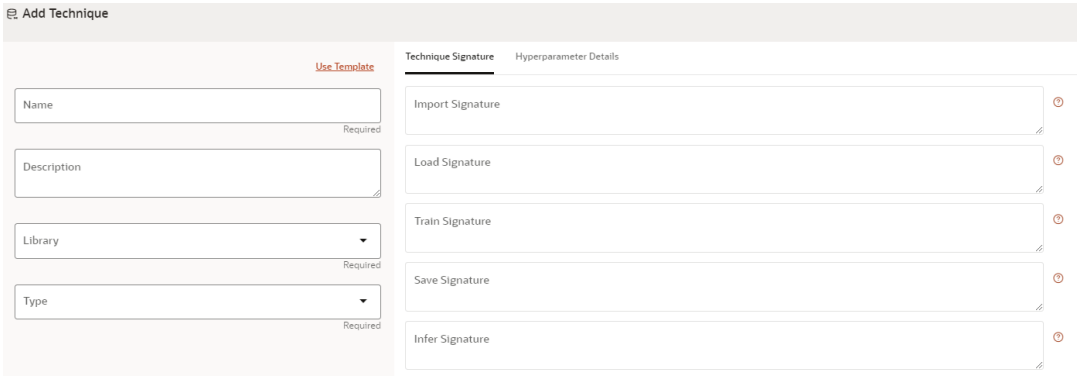
Figure 6-38 Technique Summary page



2. In the **Technique Summary** page, click **Add**.

   The **Add Technique** page is displayed.

Figure 6-39 Add Technique page



3. Enter the name of the technique.
4. Enter the description of the technique.
5. Select the library from the drop-down. Currently, the application supports the libraries such as keras, ONNX, scikit-learn, and xgboost.
6. Select the type as either as **Classification** or **Regression**.

   Classification: Classification is a process of finding a function which helps in dividing the dataset into classes based on different parameters. The task of the classification algorithm is to find the mapping function to map the input(x) to the discrete output(y).

Regression: Regression is a process of finding the correlations between dependent and independent variables. The task of the Regression algorithm is to find the mapping function to map the input variable(x) to the continuous output variable(y).

7. Enter the signatures such as Import, Load, Train, Save, and Infer. For more details on the signatures, click on the respective help icon.

   Some of the signatures captured in library stage might not be standard across different algorithm/technique provided by the library.

8. Navigate to **Hyperparameter Details** tab and click **Add** to add the parameters.

   You can add the type of parameters such as String, Integer, Float, and Boolean.

9. Click **Create**.

   The Model Technique is created and displayed in the **Technique Summary**screen.

> **Note:**
>
> The user needs to check the details provided above from the homepage of the library being captured.
> For example: `https://scikit-learn.org/stable/`.

OR

You can select the Use Template option to pre-fill the entries from the seeded list of Libraries.

You can also edit the data based on your requirements.

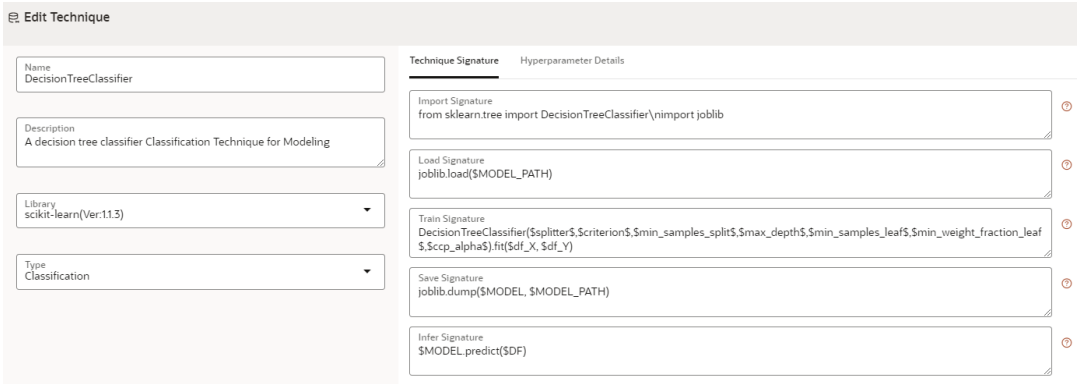
6.3.1 Edit a Model Technique

To edit a model technique, follow these steps:

1. Navigate to **Technique Summary** page.
2. Click **Action** next to corresponding model technique and select **Edit**.

   The **Edit Technique** screen is displayed.

Figure 6-40 Edit Technique screen



3. Make the necessary changes and click **Update**.

   For more details, see Model Technique section.

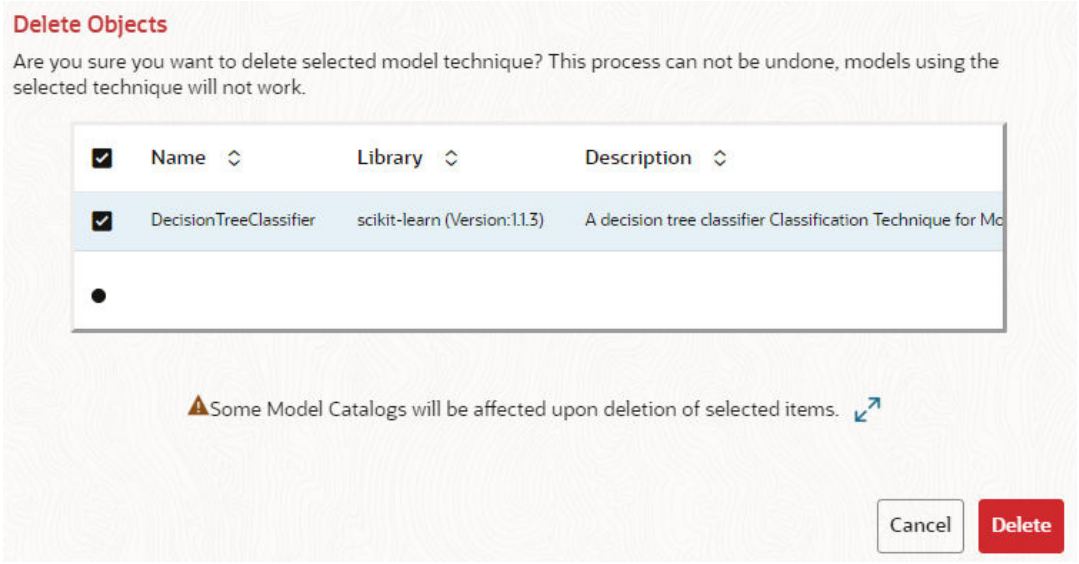
6.3.2 Delete a Model Technique

To delete a model technique, follow these steps:

1. Navigate to **Technique Summary** page.
2. Click **Action** next to corresponding model technique and select **Delete**.

   A confirmation message is displayed.

Figure 6-41 Confirmation Message



3. Click **Delete**.

   The selected model technique and all the model catalogs associated are deleted.

   To delete multiple model techniques, select all the model techniques that you want to delete and click the **Delete**icon on the page header.

6.4 Model Catalog

The Model Catalog page allows you to add and manage the Model Technique, Model Library, and Model Objectives. You can either import the models from external sources or create, train and label it as champion model.

6.4.1 Model Objective

Model Objective is the top level metadata where further models are created or trained which can be consumed in the upcoming steps.

> **Note:**
>
> Ensure Data Studio is up and running. A validation message is displayed when the Data Studio is down during creation of Model Objectives.

To add a Model Objective:

1. Click **Launch Workspace** next to the corresponding workspace to display the **Dashboard** window with application configuration and model creation menu.
2. In the LHS menu, click **Model Catalog** to display the **Model Objective Summary** window.

   This window displays the model objectives in a table.

   The following table provides descriptions for the fields and icons on the **Model Objective Summary** page.

Table 6-14 Fields and icons on the Model Objective Summary page

Field or Icon	Description
Search	The field to search for Model Objectives. Enter specific terms in the field for which you want to search, and press Enter on the keyboard to display the results.
Code	The unique identifier of the Model Objective.
Name	The name of the Model Objective.
Dataset	The dataset which is used to create the Model Objective.
Target Variable	Target Variable is the important parameter after the dataset. During the dataset selection, target variable is selected that constitutes of number of columns/features.
Type	The type which is selected during model objective creation. The type can be either Classification or Regression.
Action	Click the three dots to perform View/Delete functions on the selected model objective..

> **Note:**
>
> In the objective summary UI, a new button is introduced to show/hide empty objectives. By default "Show Empty Objectives" will be disabled (i.e. empty objectives wont be shown) User can enable this checkbox to view empty objectives. On successful addition of new objectives this checkbox is enabled automatically.

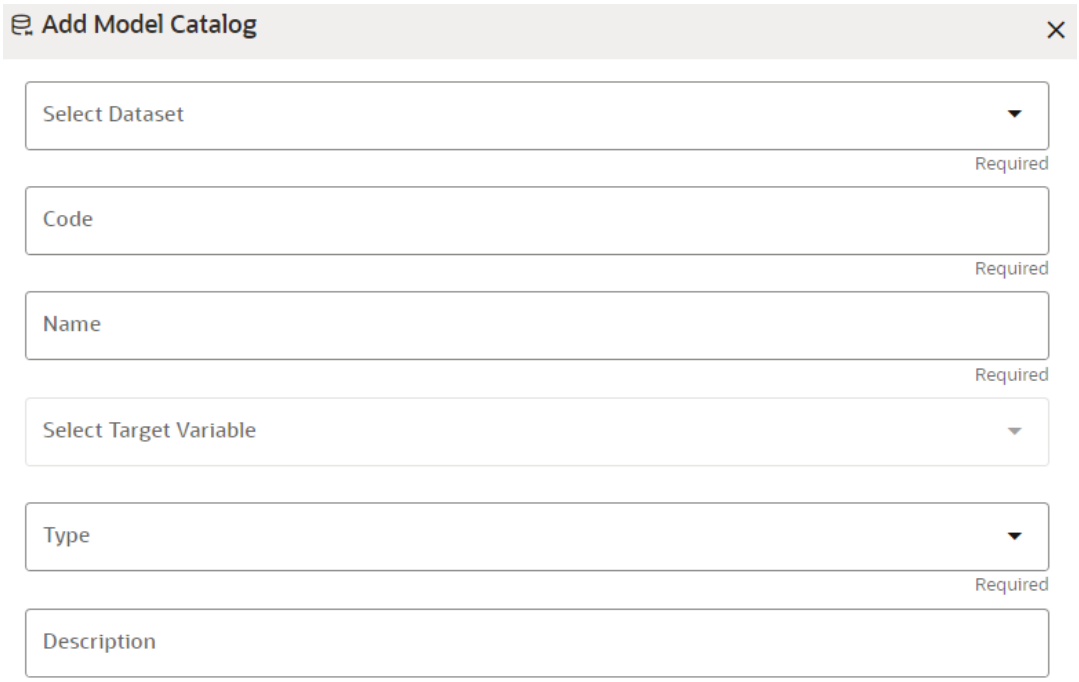
6.4.1.1 Add a Model Objective

To add a Model Objective:

1. In the **Model Objective Summary** screen, click **Add**.

   The **Add Model Objective** screen is displayed.

Figure 6-42 Add Model Objective screen



2. Select the required dataset from the drop-down.

   For more details on the dataset, see Dataset section.

3. Enter the code for the Model Objective.

   The code should be unique for each model objective.

4. Enter the name and description for the Model Objective.

5. Select one of the column/feature as the target variable for the models which needs to be trained/uploaded.

6. Select the Type from options such as **Classification** , **Regression**, or **Others**.

   Classification: Classification is a process of finding a function which helps in dividing the dataset into classes based on different parameters. The task of the classification algorithm is to find the mapping function to map the input(x) to the discrete output(y).

   Regression: Regression is a process of finding the correlations between dependent and independent variables. The task of the Regression algorithm is to find the mapping function to map the input variable(x) to the continuous output variable(y).

   Others: If the Type is not Classification or Regression, select this option.

7. Enter the description for the Model Objective.

8. Click **Create**.

   The Model Objective is created and displayed in the **Model Objective Summary** screen.

> **Note:**
>
> In the objective summary UI, a new button is introduced to show/hide empty objectives. By default "Show Empty Objectives" will be disabled (i.e. empty objectives wont be shown) User can enable this checkbox to view empty objectives. On successful addition of new objectives this checkbox is enabled automatically.

6.4.1.2 View a Model Objective

To view a Model Objective:

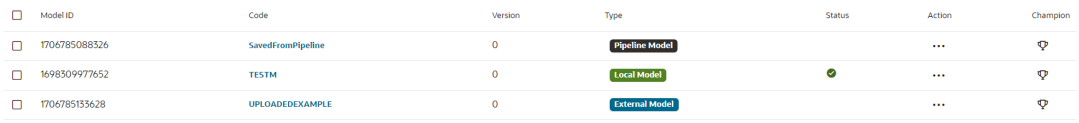
1. Navigate to **Model Objective Summary** page.
2. Click **Action** and select **View**.

   OR

   Click on the Model.

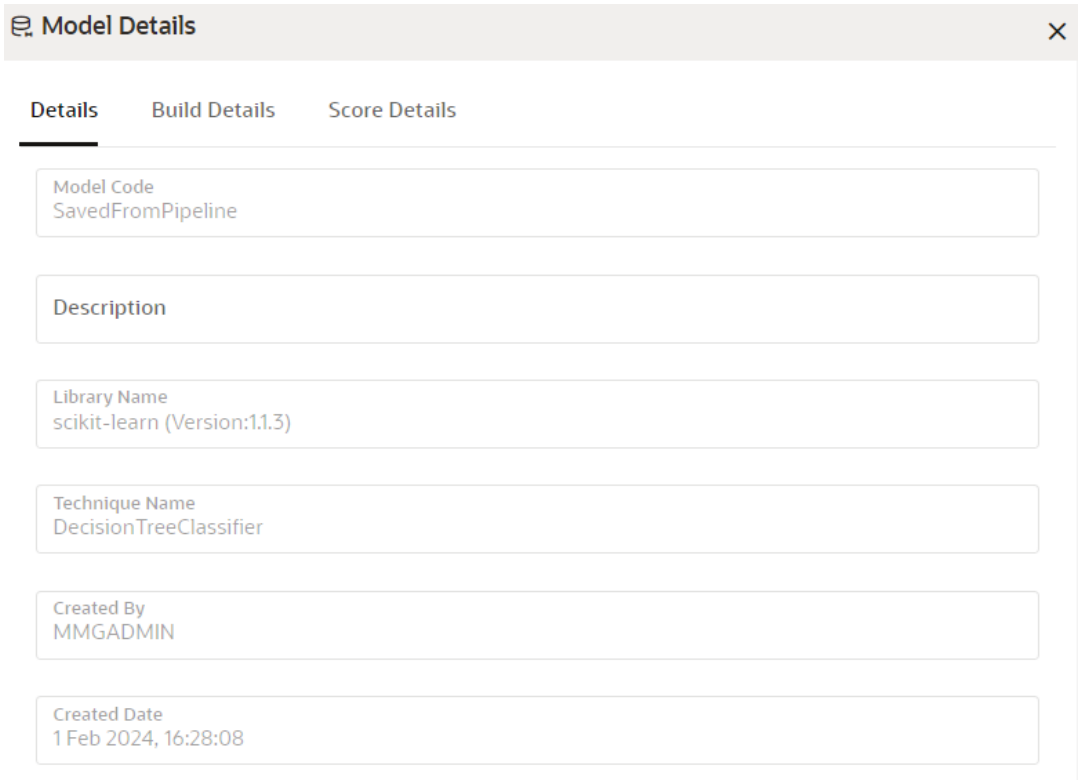
   The following page is displayed.

Figure 6-43 View page



Click on Type to view the model, build, and score details.

Figure 6-44 Model details



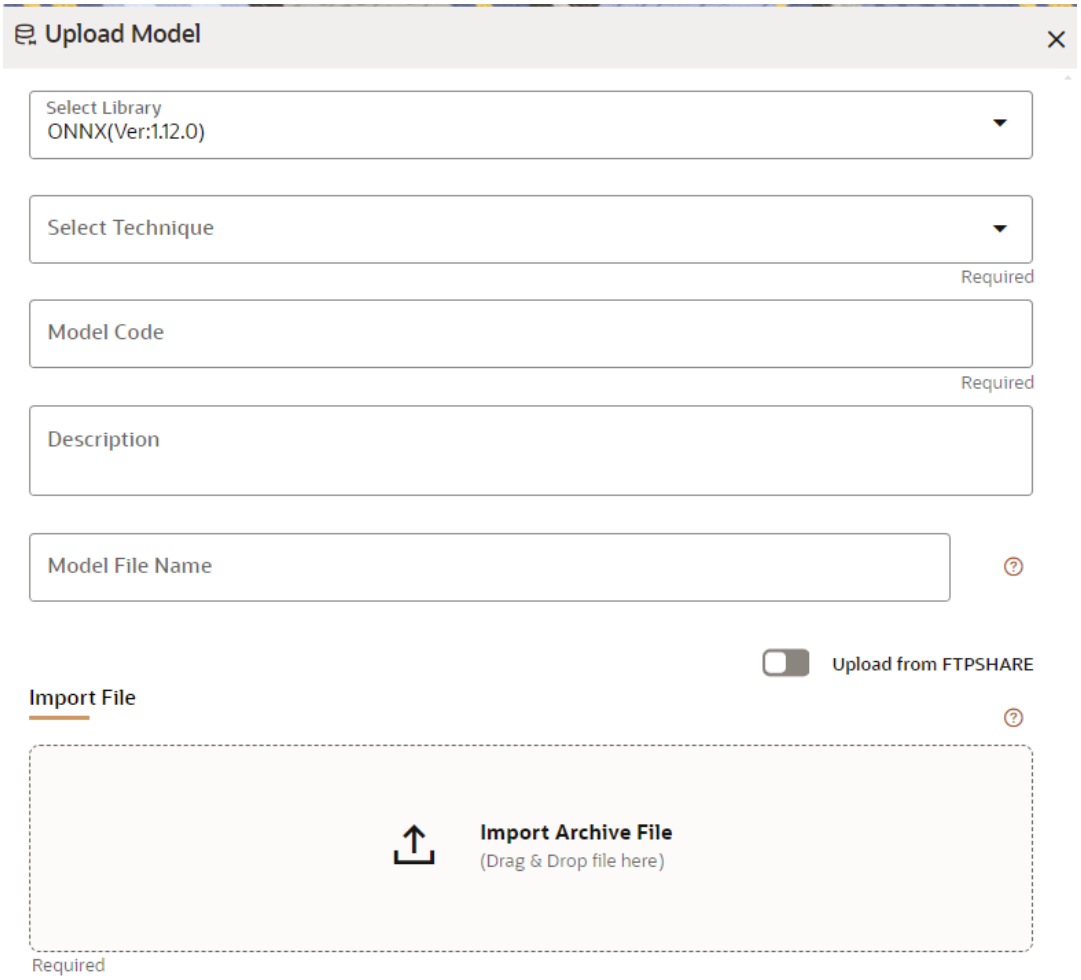
6.4.1.2.1 Upload a Model

To upload a model:

1. In the **View Model Objective** page, click **Add** and select **Upload Model**.

   The **Upload Model** page is displayed.

Figure 6-45 Upload Model page



2. Select the library from the drop-down.

   For more details on the model library, see the Model Libraries section.

3. Select the required technique from the drop-down.

   For more details on the model technique, see the Model Technique section.

4. Enter the model code and description for the Model Objective.

5. Enter the name of model file with extension for use during Save and Load.

6. Import the model using **Import Archive File**option. File format should in be zip format.

   OR

   Enable **Upload from FTPSHARE** option to import the model. File Path entered is the zip file path present on the Server relative to ftpshare path including file.
   Example: <installation path>/temp/models/model.zip

7. Click **Upload Model**.

   The model is uploaded successfully.

6.4.1.2.2 Build a Model

Building a model consists of the following steps:

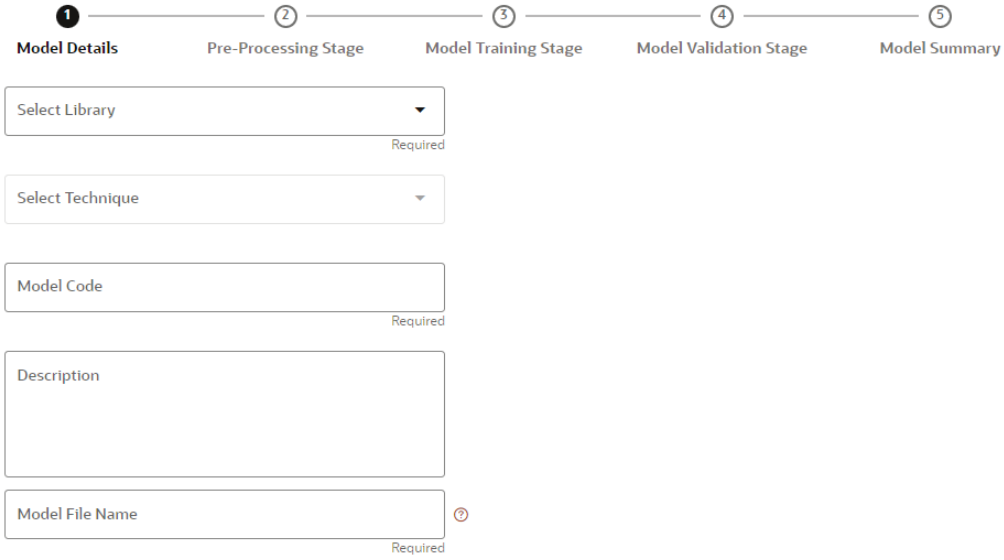
- Model Details
- Pre–Processing Stage
- Model Training Stage
- Model Validation Stage
- Model Summary

6.4.1.2.2.1 Model Details

1. In the **View Model Objective** page, click **Add** and select **Build Model**.

   The **Model Details** page is displayed.

Figure 6-46 Model Details page

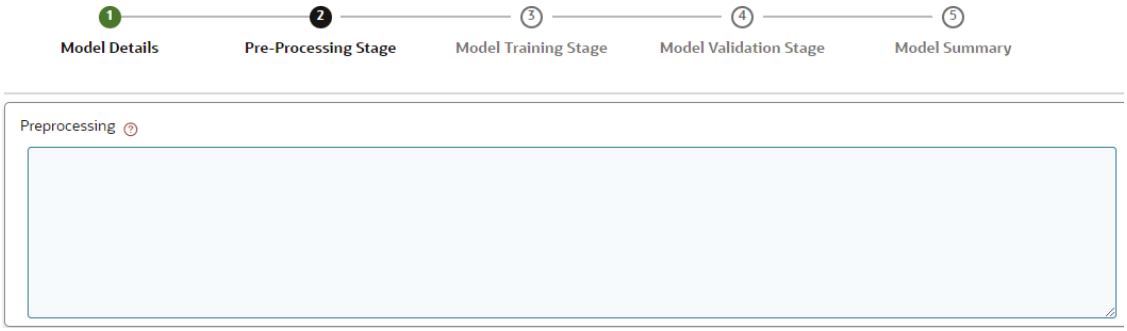


2. Select the library from the drop-down. Currently, the supported libraries are keras, ONNX, scikit-learn, and xgboost. For more details, see the Model Libraries section.
3. Select the required technique from the drop-down. For more details, see the Model Technique section.
4. Enter the model code and description for the model.
5. Enter the name of model file with extension for use during Save and Load.
6. Click **Next** to go to the next step.

6.4.1.2.2.2 Pre–Processing Stage

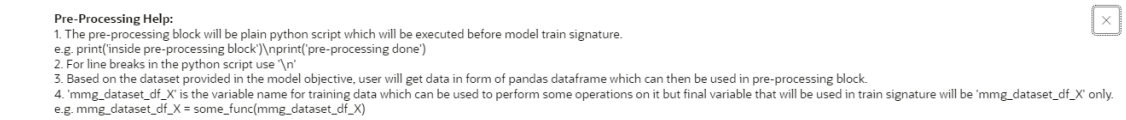
After adding the details in the Model Details page, the Pre-Processing Stage screen is displayed.

Figure 6-47 Pre-Processing Stage screen



1. Enter the data in the Pre-Processing text box as shown in an example is provided here.

Figure 6-48 Pre-Processing text box

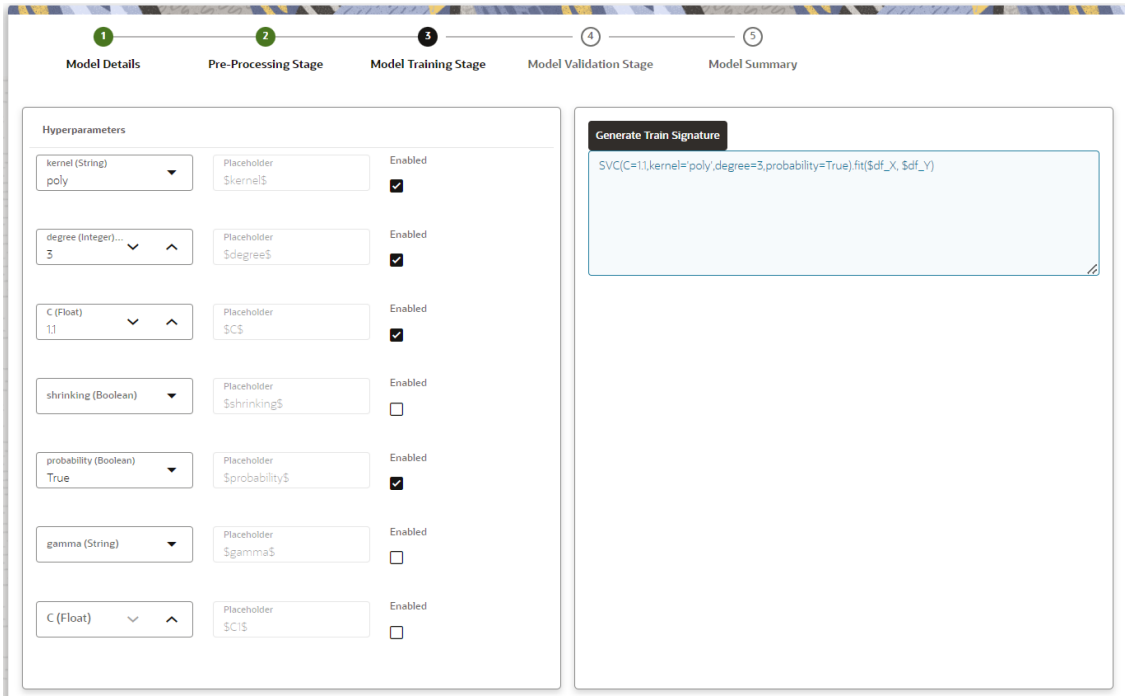


2. Click **Next** to go to the next step.

6.4.1.2.2.3 Model Training Stage

After adding the details in the Pre-Processing Stage page, the Model Training Stage screen is displayed.

Figure 6-49 Model Training Stage screen



1. Select the hyperparameters.

   When the parameters is selected, the **Enabled** check box is selected by default.

2. Click **Generate Train Signature**.

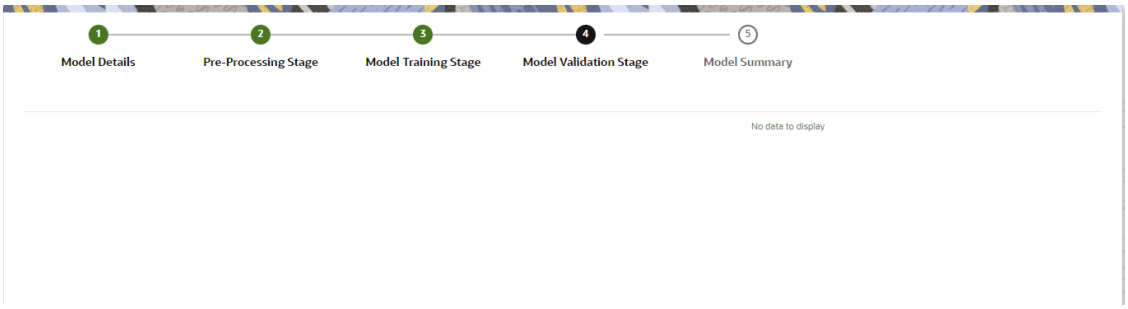
   The signature is displayed.

3. Click **Next** to go to the next step.

6.4.1.2.2.4 Model Validation Stage

After adding the details in the Model Training Stage page, the Model Validation Stage screen is displayed.

Figure 6-50 Model Validation Stage screen



The Validation is work in progress.

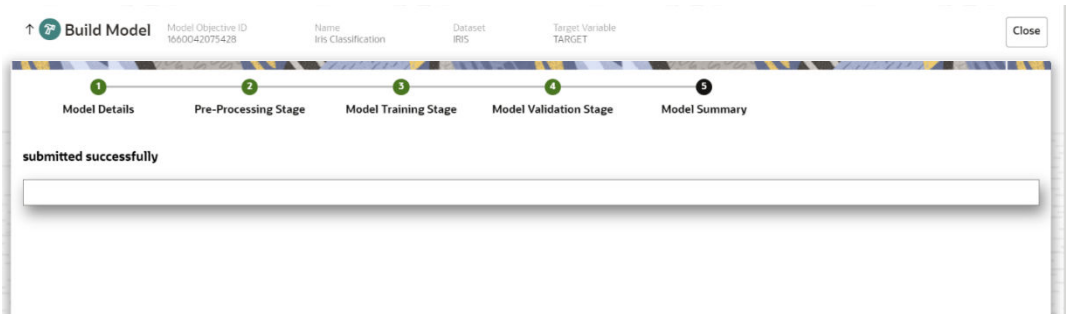
- Click **Finish** to start the training of the model.

6.4.1.2.2.5 Model Summary

Once the validation of the model is completed, the Model Summary screen is displayed. Follow the below steps to publish Models from Model Summary:

1. Navigate to Model Summary.
2. Click on the three dots.
3. Select publish Datastudio option.
4. Check the links.

Figure 6-51 Model Summary screen



- Click **Close**.

  The model is created and displayed in the Model Objective screen.

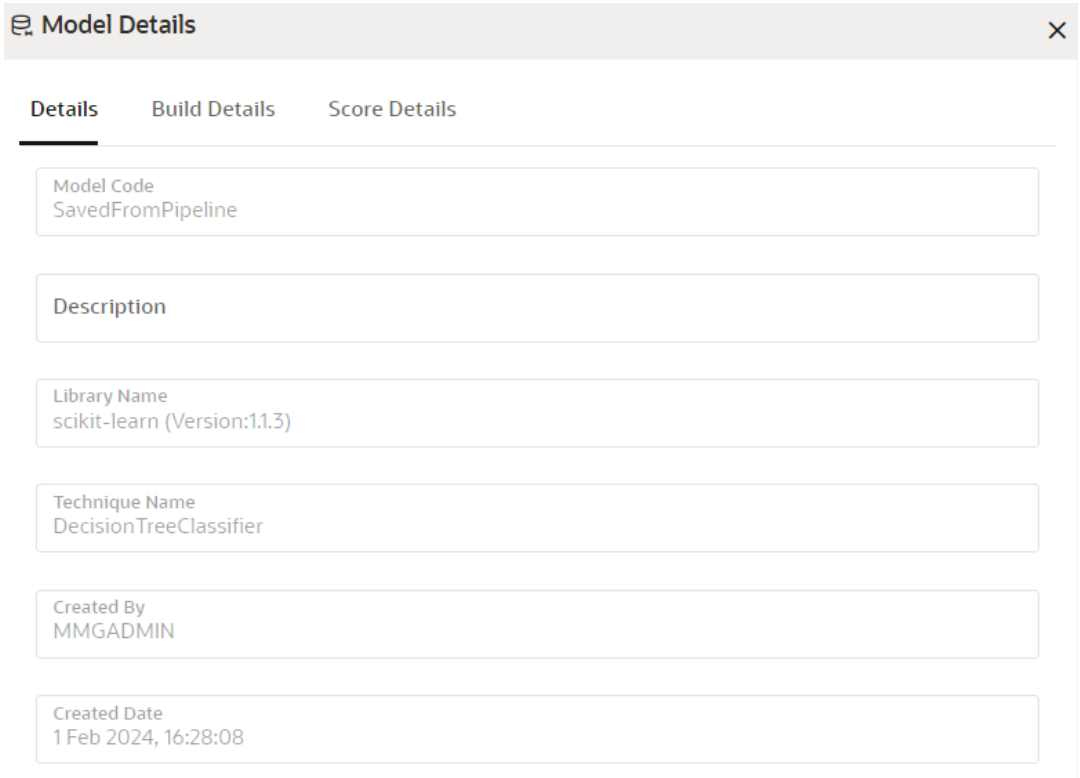
6.4.1.2.3 View Details of a Model

To view details of a model, follow these steps:

1. In the **View Model Objective** page, click **Action** next to corresponding model and select **View Details**.

   The model details are displayed.

Figure 6-52 Model details



2. Click on the **Build Details**tab to view the Preprocessing, and Hyperparameters details of the model.
3. Click the **Score Details** tab to view the Latest score, and Scoring details of the model.

6.4.1.2.4 Setting the Model as Champion

You can set the trained models as champion model.

To set a champion model:

1. In the **View Model Objective**page, hover over the model which you want to set as champion and click the **Set Champion**icon.

Figure 6-53 View Model Objective page

A confirmation message is displayed.

2. Click **Update**.

   The model is set as champion for the selected model objective.

6.4.1.2.5 Delete a Model

To delete a model:

1. In the **View Model Objective**page, select the models which you want to delete.
2. Click **Delete** and acknowledge the confirmation message.
3. Click **Delete**.

6.4.1.3 Delete a Model Objective

To delete a Model Objective:

1. Navigate to **Model Objective Summary** page.
2. Click **Delete** and acknowledge the confirmation message.
3. Click **Delete**. The selected model objective and all its associated models are deleted.

6.5 Pipelines

Model Pipeline enable you to create and publish models based on the workspaces created from datasources. The published models are then deployed in production to be consumed by other services and applications. Modelers create models by using model templates available as part of ML4AML or by developing them independently in Compliance Studio Notebooks. After building and evaluating, multiple models, the best model or champion model can be selected. The champion model can then be deployed for scoring. In this document the champion model can also be referred to as the scoring model.

6.5.1 Prerequisites

The prerequisites for model pipeline are as follows:

- To create a model, your user profile must be mapped to the Modeler Group.
- To create a model, a workspace must be deployed.
- To approve and deploy a model, your user profile must be mapped to the Modeling Administrator Group.

6.5.2 Access Model Pipelines

The Model Pipeline window allows you to create and publish models.

To access the Model Pipeline window:

1. Navigate to **Workspace Summary**page.

   The page displays workspace records in a table.

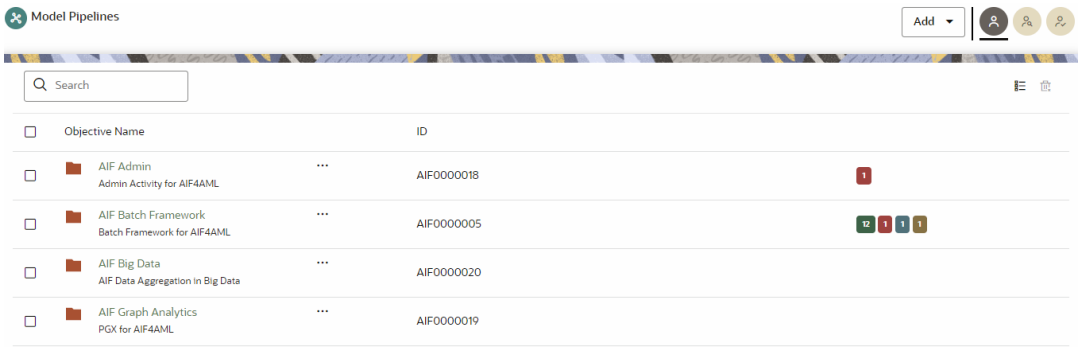
2. Click **Launch** next to corresponding Workspace to Launch Workspace.

   The **Dashboard** window is displayed with application configuration and model creation menu.

3. In the LHS menu, click **Model Pipelines** to display the **Model Pipelines** page.

   The window displays objectives that contain drafts and models. When you hover on the count that are next to the ID column, it displays the count of sub objectives, Drafts, Models, and Champion if available.

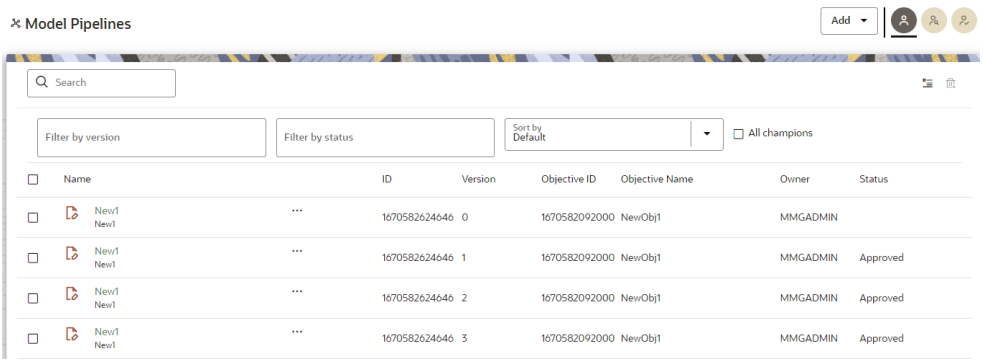
Figure 6-54 Model Pipelines page



You can switch the Model Summary page view from Flat to Hierarchical and vice-versa. Flat list is the default view in Production workspace while Hierarchical list is the default in Sandboxes.

- **Hierarchical option**: The Model Management window shows the hierarchical list of the Objectives using the **Hierarchical** option. In the Hierarchy view, you can see the following details of the Objectives such as Objective Name, and ID.
- **Flat option**: The Model Pipeline window shows the flat list of all the models (published and drafts) using the **Flat** option. Flat list is not objective-specific. It shows the models across all the objectives. You can search the models using the Filter by Version, Filter by Status, and All champions. You can also sort the drafts and models by Default or Latest first options.
  - In the Flat list of models, you can see the following details of the models such as Name, ID, Version, Objective ID, Objective Name, Owner, and Status.

Figure 6-55 Model Pipelines



The following table provides descriptions for the fields and icons on the Model Pipeline page.

Table 6-15 Fields and icons on the Model Pipeline page

Field or Icon	Description
Model Pipeline Page Header	The header that follows the global header in the application.
Breadcrumbs	Indicates the position of the current page in the Model Management hierarchy. Use breadcrumbs locator links to navigate back to higher levels in the hierarchy after you have drilled down through levels of functions. Click to navigate to Model Summary page. Click or to navigate back to Workspace Summary page.
Create Objectives and Models	Select Add to display a list with the following options: • Draft • Objective • Seeded Models To create Models, select Draft. To create Objectives, select Objective. See the Create Objective (Folders) and Create Draft Models sections for more information. However, if you want to know about the cycle of model creation, see the Create, Review, Approve, and Deploy a Model Section.
Requester	Displays that the logged-in user has the Requester privileges when the status is green. You can create a model. However, to approve and publish, the model must be reviewed by a user with reviewer privileges and approved by a user with approver privileges.
Reviewer	Displays that the logged-in user has the Reviewer privileges when the status is green. You can review models. However, to approve and publish, the model must be approved by a user with approver privileges.
Approver	Displays that the logged-in user has the Approver privileges when the status is green. You can approve models that are created and reviewed.
Model Pipeline Page Table	The table displays the objective and model records on the page.
Objective Name	Displays the name and description of the Objective
ID	Displays the ID of the objective.
Owner	Displays the owner who created the Drafts. This information does not display for an Objective.
Tags	Displays the tags associated with the Models or Drafts. This information does not display for an Objective.
Delete Objective	Select to delete the model from the Confirmation dialog box. Click Delete to process or click Cancel to cancel. This information does not display for an Objective.
Edit Objective	Select to edit the models in Data Studio. See the Edit Models section for more information. This information does not display for an Objective.

6.5.3 Manage Models

A Model has to go through an approval workflow before it can be deployed to production. The following types of users in the system have privileges that restrict the activities, they can do in the model creation and deployment workflow.

Table 6-16 Field or icons for the Model creation and deployment

Field or Icon	Description
Requester	You can create a model. However, to approve and publish, the model must be reviewed by a user with reviewer privileges and approved by a user with approver privileges. **NOTE**: User Groups must be mapped to the MDLDEPLOY role to access Requester functions.
Reviewer	You can review models. However, to approve and publish, the model must be approved by a user with approver privileges. **NOTE**: User Groups must be mapped to the MDLREVIEW role to access Reviewer functions.
Approver	You can approve models that are created and reviewed. You can then promote to production and make the model the champion in the production. **NOTE**: User Groups must be mapped to the MDLAUTH role to access Approver functions.

The following sections in this topic provide details for the cycle of creation of a model, review, approval, and deployment:

- Create Objective (Folders)
- Create Draft Models
- Publish Models (Scoring)
- View Model Details
- Compare Models
- Understand Model Governance
- Request Model Acceptance
- Review Models and Move to Approve or Reject
- Approve Models and Promote to Production
- Deploy Models in Production and Make it a Global Champion

6.5.3.1 Create Objective (Folders)

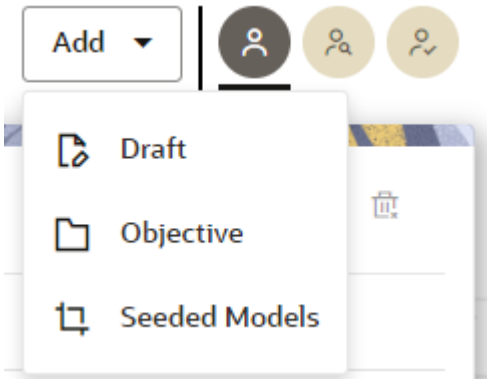
Create folders called Objectives within which you can create Models.

To create an Objective, follow these steps:

1. Click **Launch Workspace** next to the corresponding Workspace to Launch the **Dashboard** window with application configuration and model creation menu.
2. Click **Model Pipelines** to display the **Model Pipelines** window.

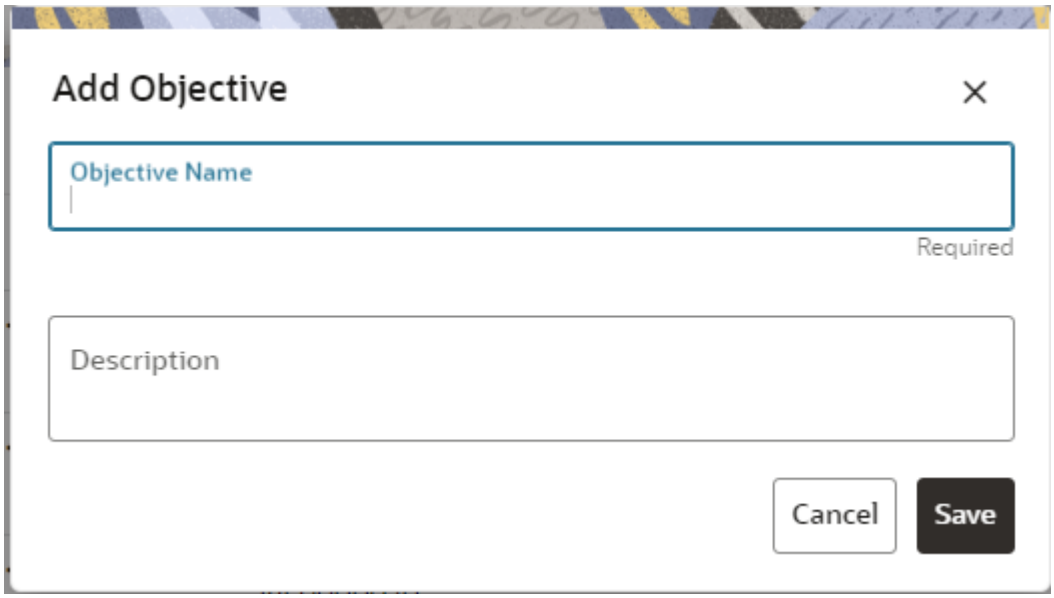
   The window displays folders that contain models and model records in a table.
3. Click **Add** and select **Objective** from the list to display the **Objective Details** dialog box.

Figure 6-56 Select Objective from Add



4. Enter details in Objective **Name** and **Description** fields in the **Add Objective** dialog box.

Figure 6-57 Objective Details Dialog box



5. Click **Save**.

6.5.3.2 Create Draft Models

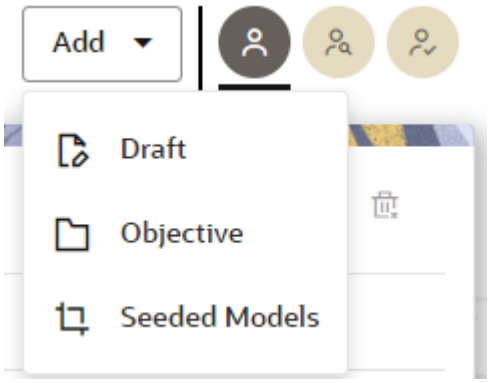
Once an objective has been created, a model draft or notebook can be created where model development can start.

To create a Draft Model:

1. Click **Launch Workspace** next to the corresponding Workspace to Launch the **Dashboard** window with application configuration and model creation menu.
2. Click **Model Pipelines** to display the **Model Pipelines** window.

   The window displays folders that contain models and model records in a table.
3. Click **Add** and select **Draft** from the list to display the **Add Draft** dialog box.

Figure 6-58 Create Model



4. **Create New Model** is the default setting in the **Model Details** dialog box. Drag the toggle button to select **Import Dump**. Use Create New Model to start from a blank Notebook in Compliance Studio. You can also create a draft under Objective (Folder) also. Click an Objective to open it.

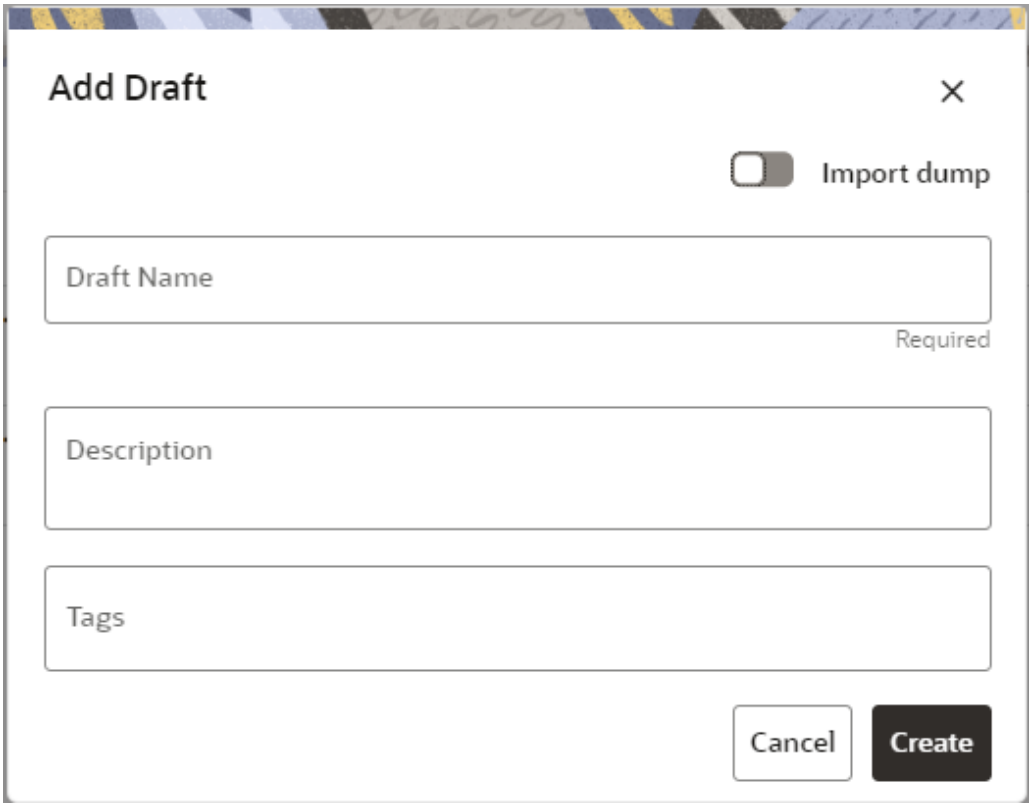
   OR

   Drag the toggle button to select **Import Dump**. Import Dump lets you drag and drop an existing file with model data and modify it. To import a model data dump from another model, see the Import a Workspace Model Data into a New Model section.

   To create a new model:

   a. Enter details for Draft **Name** and **Description**.

Figure 6-59 Model Details- Create New Model



b. Enter a tag in the **Tags** field.

c. Click **Create**.

6.5.3.3 Create Seeded Models

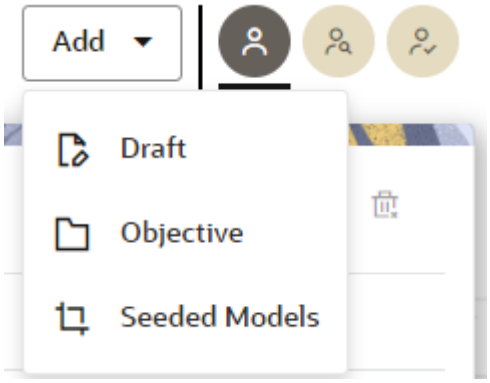
You can seed the models from the external sources which can be imported in the application.

To import the models:

1. Click **Launch Workspace** next to the corresponding Workspace to display the **Dashboard** window with application configuration and model creation menu.
2. Click **Model Pipelines** to display the **Model Pipelines** window.

   The window displays folders that contain models and model records in a table.
3. Click **Add** and select **Seeded Models** from the list to display the **Add Draft** dialog box.

Figure 6-60 Add Seeded Models

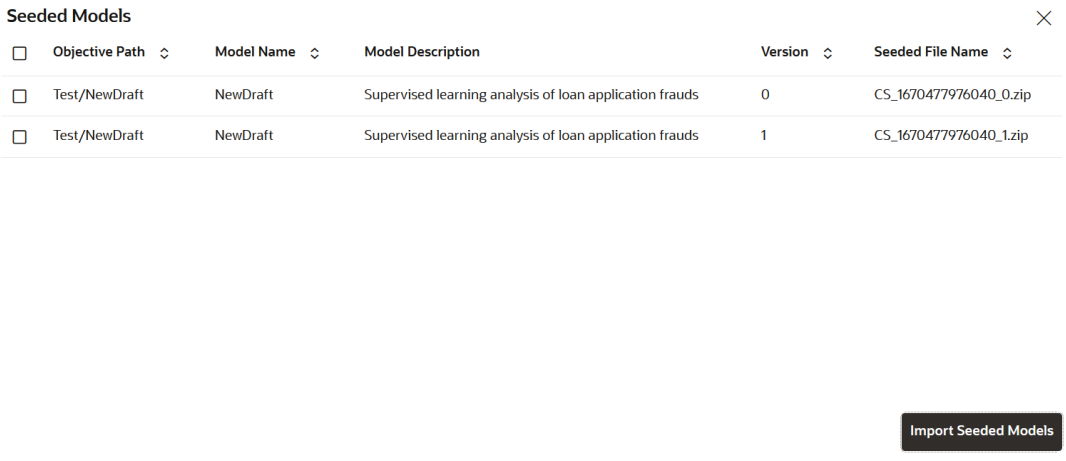


4. You must add the models in the following installed path location:

   `/scratch/ofsaaweb/ftpshare/mmg/seeded/models`

   The added models is displayed in the Seeded Models page.

Figure 6-61 Seeded Models



5. Select the models which you want to import and click **Import Seeded Models**.

The selected models are imported and displayed in the **Model Pipelines** page.

6.5.3.3.1 Model Report

This option allows you to view the Model report and download the same in PDF format.

To view and download Report, follow these steps:

1. Open a Model in Pipeline Designer.
2. Click **Generate Model Report**.

   The Model Report window is displayed.
3. Select the Parameters to generate the report.
4. Click **Preview Report**.

   The report is displayed based on selected parameters.
5. Click **Export as PDF** to save the report as PDF in local system.

6.5.3.3.2 Download a Model

This option allows you to download the Model. To download a model, follow these steps:

1. Open a Model in Pipeline Designer.
2. Click **Download**.

   A zip folder is downloaded. This folder contains .cfg and .dsnb files.

6.5.3.3.3 Delete

This option is used to delete the Model.

6.5.3.3.4 Cloning a Model

You can pick any published model and clone the contents to a new draft in the same objective or clone the content to the current parent draft. The cloned draft can be edited and used further. Audit Trail window also captures the clone information.

To clone the model details, follow these steps:

1. Open a Published Model in Pipeline Designer.
2. Select Clone to new Draft or Reimage parent draft with current.

6.5.4 Pipeline Designer

Clicking on the model navigates you to the Pipeline Designer page.

The following sections are available on the Pipeline Designer window:

- Pipeline Canvas
- Dashboard
- Notebook
- Simulations
- Execution History
- Compare

Note:

1. Models in Production workspace have the 'Dashboard' as the default tab.
2. Drafts or Models in the Sandbox workspace have the 'Pipeline Canvas' as the default tab.

6.5.4.1 Pipeline Canvas

You can perform following functions on Pipeline Canvas:

- Creating a Pipeline
- Creating Script Template
- Viewing a Pipeline
- Using Link Connector Nodes
- Execution of Pipeline
- Publishing a Pipeline

Note:

Users are now able to view the sub-stages from the canvas with the help of the new stage dropdown.

6.5.4.1.1 Create Paragraphs using Pipeline Designer

After creating the Models in the Workspace, create Paragraphs using the **Pipeline Designer** window.

Pipeline Designer enables you to design the paragraph using widgets (graphical representation) instead of using python codes. In addition, if you add new paragraphs in Data Studio, the added paragraphs are displayed in the widget format on Pipeline. Similarly, if you create a Notebook using Pipeline Designer, it can be opened for editing in Data Studio using Studio Notebook option.

Note:

When you open the notebooks from the UI, the attach workspace call will be made from mmg service and proper workspace will get attached. If the Studio is opened outside of mmg, then the attach_workspace command has to be used.

This helps the Financial Institutes and Banks in following ways:

- Visualization of the data (for example, based on data tasks)
- View the dependency
- Modify the flow of execution or execution order
- Easy for Auditing purpose

You can execute the flow based on requirement. For example, if you have created one flow and want to execute a flow of training paragraphs out of that and other flow as experimental way, then you can modify using the Training and Experimentation link types. One flow can be break into 2-3 flows for execution purpose.

When a draft is edited using the Pipeline Designer/Data Studio, and published, then a new version of published model is displayed in Model Summary page.

> **Note:**
>
> When you add a new paragraph from studio and opened the same in pipeline designer, it gets linked to multiple paragraphs. For example, you have paragraphs P1, P2 and P3 in the same order in a Notebook. It shows in Pipeline Designer canvas as P1 > P3 > P2.
> If you add two new paragraphs P1a and P1b in the Notebook after the P1 and open the canvas, this gets reflected as the following:
>
> 1. P1 -> P1a -> P1b -> P2
> 2. P1 -> P3 -> P2

After opening a Model in the Pipeline Designer, following options are displayed:

- Generate Model Report
- Download
- Delete
- Clone Model

After opening a Draft in the Pipeline Designer, following options are displayed:

- Generate Model Report
- Download
- Delete
- Publish
- Script Template list

6.5.4.1.1.1 Create a Pipeline

To create a paragraph using pipeline:

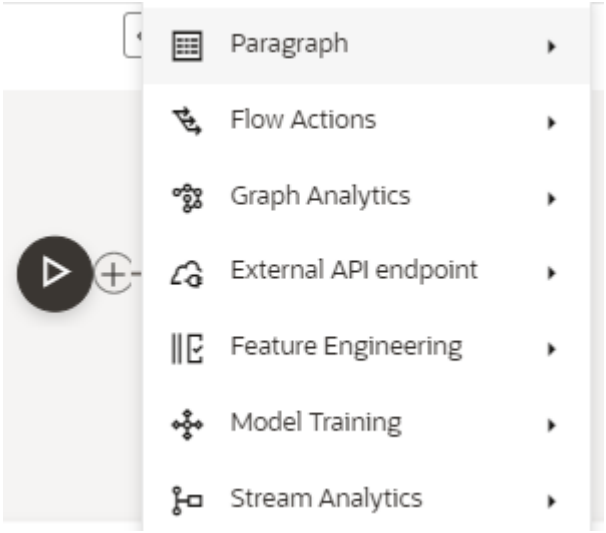
1. Navigate to the **Pipeline Designer** page.

   The Start widget is displayed by default in the canvas screen.

2. Hover over on Start widget and new nodes can be added using the add button in the Pipeline Canvas.

   This will open a menu which contains all the widgets sectioned into different categories. The spectra diagram builder is based on row and column positioning.

Figure 6-62 Create a Pipeline menu



> **Note:**
>
> Click on each menu to select the widgets of the required types.

You cannot edit or delete this widget. Whenever a new draft is created (not by importing dump files), the default paragraph created is converted into a start widget. The visibility of code/result/title in notebook of this node will be kept to invisible.

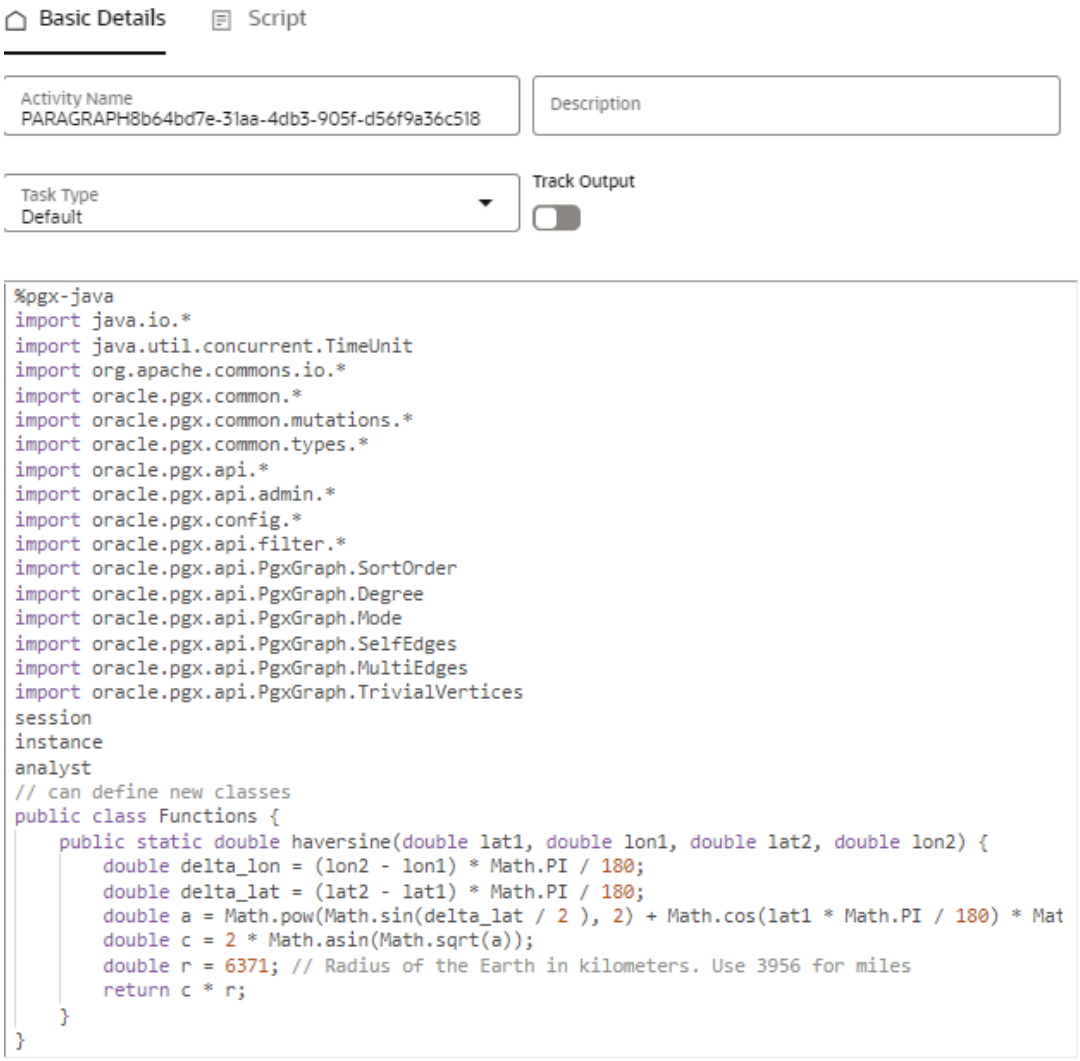
Whenever the notebook is opened, init script execution including workspace attach will happen in this start node.

Title of the node is called 'Start widget'. Publish /download and import /promote to production of model with start widget will keep this widget. Publish from canvas will keep the start widget in the published model. When you publish model from model summary, you can explicitly select the start widget paragraph from the list of paragraphs. After this only, start widget will appear in the published model.

3. Click on the node to add the basic details.

   The Basic Details page is displayed.

Figure 6-63 Basic Details for Paragraph



4. Provide details as described in the following table:

Table 6-17 Adding Details

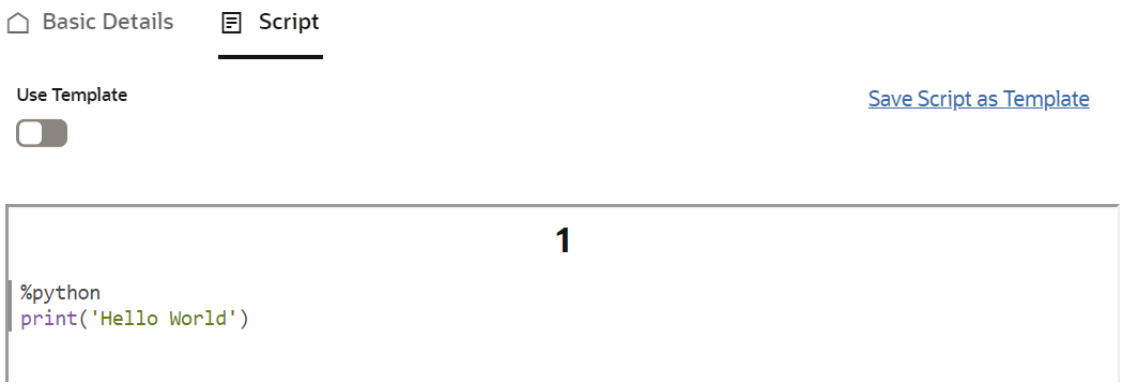
Field	Description
Activity Name	Enter the Activity Name
Description	Enter the description of Activity
Task Type	Select the task type. For example, Model Training, Data Analysis and so on. You can also search the task type.
Track Output	If this option is selected for any paragraph, then during the model comparison the output details are displayed. Keep the Track Output to ON in case you want to execute the paragraph and view the result from the Dashboard tab.
Script	Shows the script. You can edit the script in this screen or in the Script tab. This script can also be saved as the Script Template.

> **Note:**
>
> **If** is a conditional node, which behaves based on the execution of the user. It has a script which evaluates **"True"** or **"False"** and based on the value, it chooses either of them. Additionally, we can also add a new node that comes with a default link. The default path will execute irrespective of the result.

6.5.4.1.1.2 Create Script Template

Navigate to Script tab to manage scripts. These scripts can be called for paragraph. You can also create the Script template by clicking on the More options icon in the Pipeline Designer.

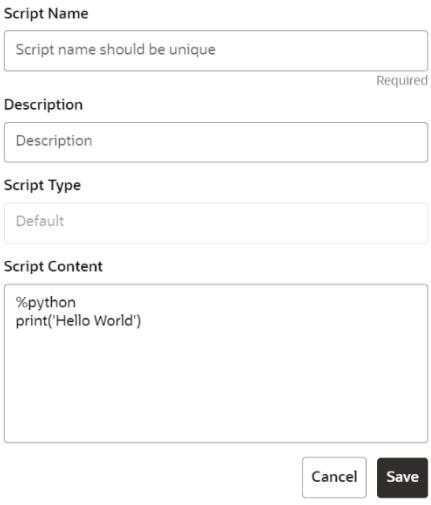
Figure 6-64 Script template



To add a script template:

1. In the **Script** tab, Click **Save Script as Template** link.

Figure 6-65 Script tab



2. Enter the following details:

Table 6-18 Creating Script Template

Field	Description
Script Name	Enter the Script Name
Description	Enter the description of Script
Task Type	Select the task type. For example, Model Training, Data Analysis and so on. You can also search the task type.
Script Content	Enter Python script

3. Click **Save**.

   A node is created.

   You can perform the following functions on Pipeline Designer window:

   - Parameter Sets: Allows you to view, edit, clone, and delete the Parameter Set. Use the Clone option to duplicate the parameter set with different values and name.
   - Publish: Allows publishing the pipeline. This option is displayed for the Drafts.
   - Deploy: Allows you to deploy the model. This option is displayed for the published models.
   - More Actions icon:
     - Generate Model Report: To generate model report
     - Download: Use Download to download the current working version in opened in canvas
     - Script Template list: Allows you to add, edit, view, and delete the script templates.
     - Delete: Allows to delete the current working version. If this is first draft of Model, it will delete all the dependent published version in the Sandbox. If the Model is not first version, then it will delete only the current working version.

     The following options are displayed only for the published models.

     - Clone to new draft: Allows to create a new draft with the same pipeline.
     - Overwrite existing draft: Allows to overwrite the existing draft with the current published model.

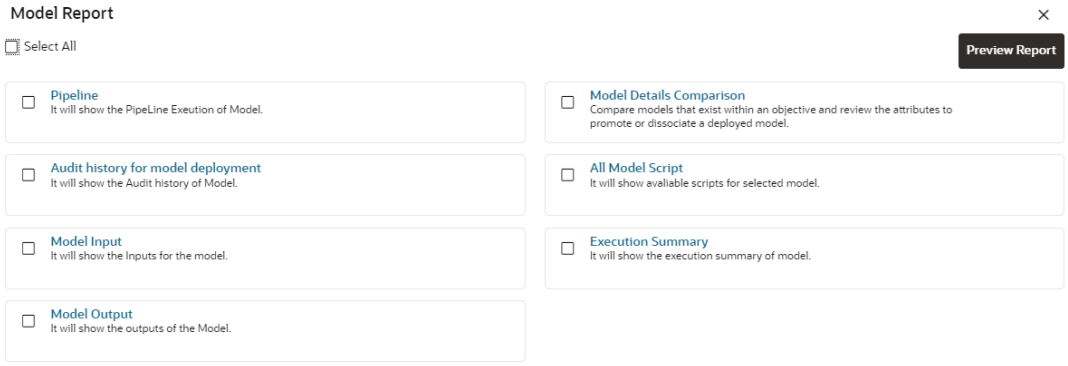
6.5.4.1.1.2.1 Model Report

This option allows you to view the Model Report and download the same in PDF format.

To view and download report:

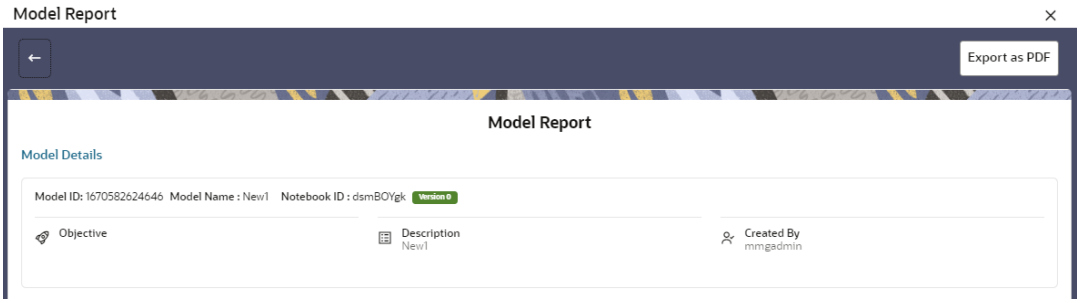
1. Open a Model in Pipeline Designer.
2. Click **Generate Model Report**.
   The Model Report window is displayed.

Figure 6-66 Model Report window



3. Select the Parameters to generate the report.
4. Click **Preview Report**.
   The report is displayed based on selected parameters.

Figure 6-67 Model Report



5. Click **Export as PDF** to save the report as PDF in local system.

6.5.4.1.1.2.2 Download a Model

This option allows you to download the Model.

To download a model:

1. Open a Model in Pipeline Designer.
2. Click **Download**.
   A zip folder is downloaded. This folder contains .cfg and .dsnb files.

6.5.4.1.1.2.3 Delete

This option is used to delete the Model.

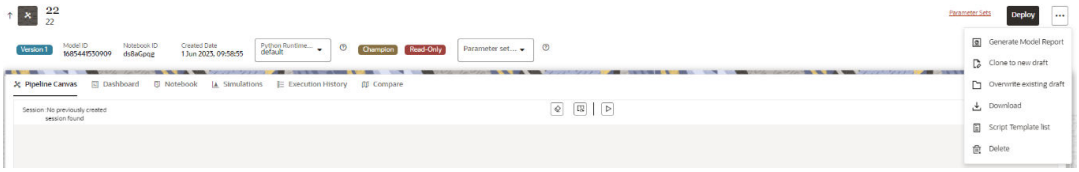
6.5.4.1.1.2.4 Cloning a Model

You can pick any published model and clone the contents to a new draft in the same objective or clone the content to the current parent draft. The cloned draft can be edited and used further. The Audit Trail window also captures the cloning information.

To clone the model:

1. Open a Published Model in Pipeline Designer.

Figure 6-68 Published Model in Pipeline Designer

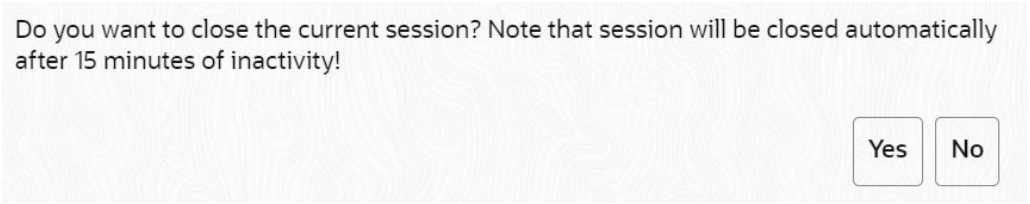


2. Click **More Actions** icon and select **Clone to new draft** or **Overwrite existing draft** option.
   You can perform the following functions on Pipeline Canvas window:

   - **Clear Execution Results:** Clears the execution details.
   - **Invalidate session:** Deletes the previous session details.
   - **Execute:** Allows you to execute the pipeline.
   - **Save Now:** Allows saving the pipeline.
   - **Open Notebook in session:** Allows you to open the notebook in canvas. This is displayed after the pipeline is executed.
   - **Execute Notebook in another session:** Allows you to provide different data for the same notebook execution. This option is only displayed when execution is in progress.
   - **Stop Execution:** Allows you to stop the execution which is in progress. This option is only displayed when execution is in progress.
     Whenever user executes a batch, a user session is created. This execution time can be less or more for any execution. Sometime, user doesn't want to wait for execution to complete and navigate away from the Pipeline Canvas page.

     For example, if user does not want to execute all the paragraphs and want to execute only Paragraph 1 and Paragraph 2. Paragraph 1 takes 15 minutes time for execution and Paragraph 2 takes 10 minutes for execution. Paragraph 2 execution also wants to use the execution of Paragraph 1. In this case, user can navigate away from the page. A confirmation message is displayed to close the current session. Here, this session time is configurable.

     Figure 6-69 Confirmation message

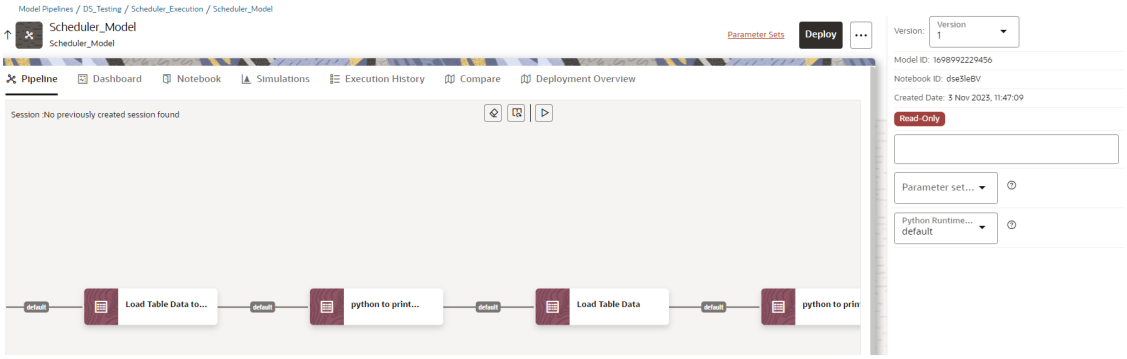
     - If user clicks **Yes**, then execution thread will be closed for this given session time.
     - If user clicks **No**, then execution thread will be valid for this given kill time and run in the background.

     You can configure these values in Application.xml file. For more information, see the *MMG Installation Guide*.

6.5.4.1.1.2.5 Add a Version, Parameter set, and Python Runtime parameters

The Pipeline canvas window allows you to add a Version, Parameter set, and Python Runtime parameters.

Figure 6-70 Pipeline canvas window



- **Version**: Select the version for the pipeline.
- **Parameter Set**: Select the required parameter set. The selected parameter set will be promoted along with the model pipeline at the time of deployment and you can update this dependency based on your requirements.
- **Python Runtime Parameters**: Select the required Python Runtime parameter. The selected Python runtime parameter will be used during all the executions and you can update this dependency based on your requirements.

6.5.4.1.1.2.6 View a Pipeline

The Pipeline canvas window allows you to view the Pipeline using the following options:

- **Auto-align**: Arrange all the widgets in vertical order. After saving, the reverting option will not work.
- **Revert-align**: Revert all the widgets if they are Auto-aligned.
- **Refresh**: Refresh the pipeline canvas.

6.5.4.1.1.2.7 Execute a Notebook

The Execute icon on Pipeline Canvas allows us to execute the notebook.
The following link types are available in the Pipeline Designer:

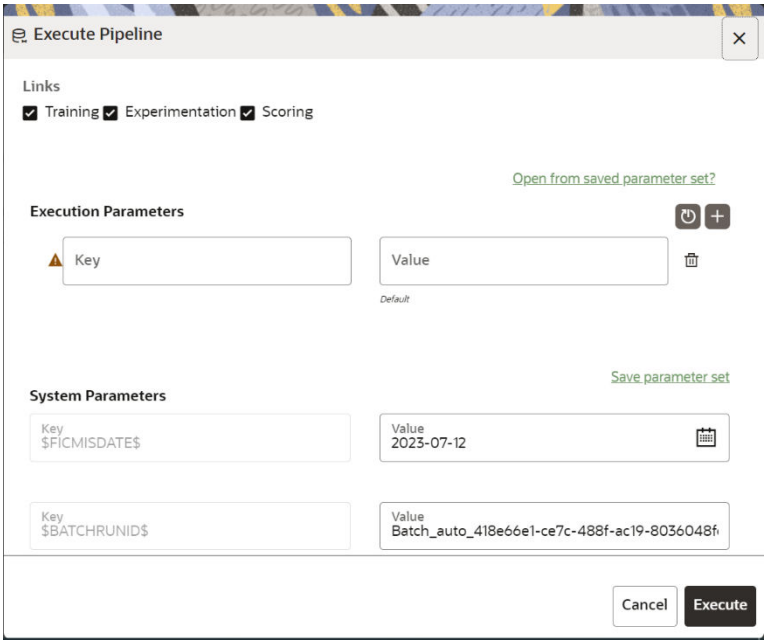
- Default
- Scoring
- Training
- Experimentation

> **Note:**
>
> When a model gets published from Model summary page, the Link types configured in Pipeline Designer are set to Default link type.

To execute the notebook:

1. Click **Execute** to view **Execute Pipeline** window.

Figure 6-71 Execute Pipeline



2. Execution parameters are the parameters defined in the notebook required for execution. Select the flow, which you want to execute Scoring, Training, and Experimentation. It displays all the keys defined for all the paragraphs in the notebook with a placeholder for providing the values.
3. Enter the execution Key and Value.

   You can also use Runtime parameters for execution. This runtime parameter must be defined in Notebook. If this is defined, you can enter execution value during the process execution.

   For more information, see Create Paragraphs in Model Studio Notebooks section.

   The System Parameters window also shows the execution ID, execution Date, and execution Batch. These are required for executing all the paragraphs along with other parameters. It also shows from where the parameter comes from as a subscript.

   - **Parameter Sets**: These are the set of parameters with a specific value required for an instance of execution. It consists Key and Value. You can save the parameters set that can be used for one execution instance and reuse it for the next execution. It consists of parameters with a specific value to each parameter. Parameters containing no value will not be taken. Each set is identified with a unique code for each objective. While saving the parameter, you have to provide a code for identifying the name and description which is not mandatory.
     You can save Key Value parameter set using the Save Parameter Set option. To Save Parameter Set, enter the Threshold Value and Description in the Parameter Set window.
   - **Selecting Parameter Set**: These saved Parameter Set can be selected during the execution. It will replace the values of the parameters from the chosen Parameter Set. To select the Parameter Set, follow these steps:
     - Click “Open from saved Parameter Set". The Threshold Code window is displayed.
     - Select the Parameter Set from the available list. You can select multiple Parameter Set in the same execution instance. In that case, if there are any common keys, value will be replaced with that from the latest Parameter Set selected.

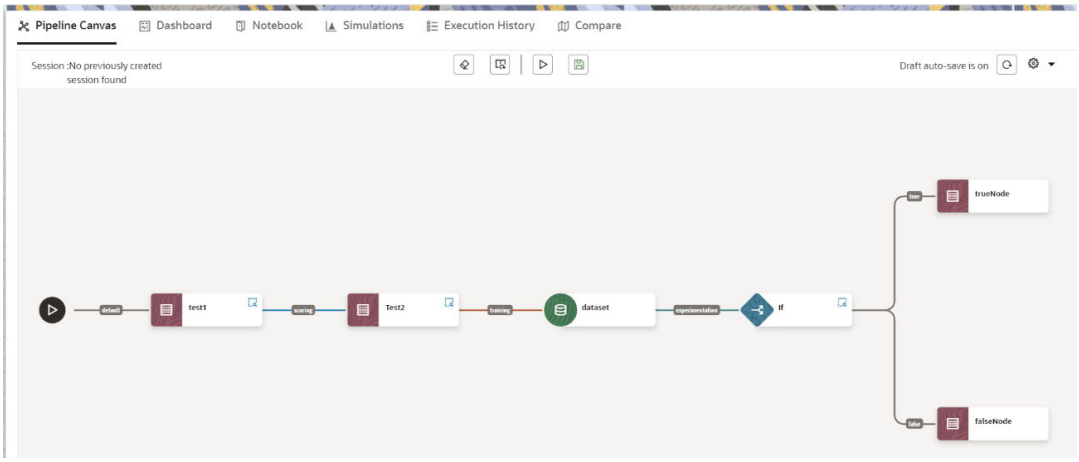
4. You can add new parameters using **Save parameter set** option.

> **Note:**
>
> If the parameter is not defined in the notebook, it will not be used for the execution. In case of multi select, if there are common parameters among the chosen scenarios, it will take the value based on the order of selection. that is first chosen scenario parameter will be taken.

5. But if open from saved Parameter Set again (not on single go), then already added will get replaced by the newly added (same as what existed).
6. Execution is performed based on selected link types. It filters out all the not required/ unused parameters. And all the unused parameters for the current execution are displayed with a warning. To view the only required parameters, click **Show only required**link.
7. Click **Reset**to reset the entered data.
8. Click **Delete** to delete the entered Key and Value.

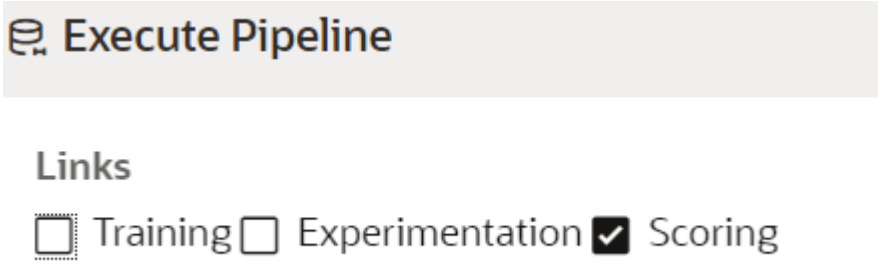
   For example, refer to below figure.

Figure 6-72 Example of Pipeline Canvas



- Here, if you want to execute this Notebook for scoring purpose, then the flow will be executed till Test 2. To perform this, Click Execute Notebook and select Links as Scoring.

Figure 6-73 Execute Pipeline



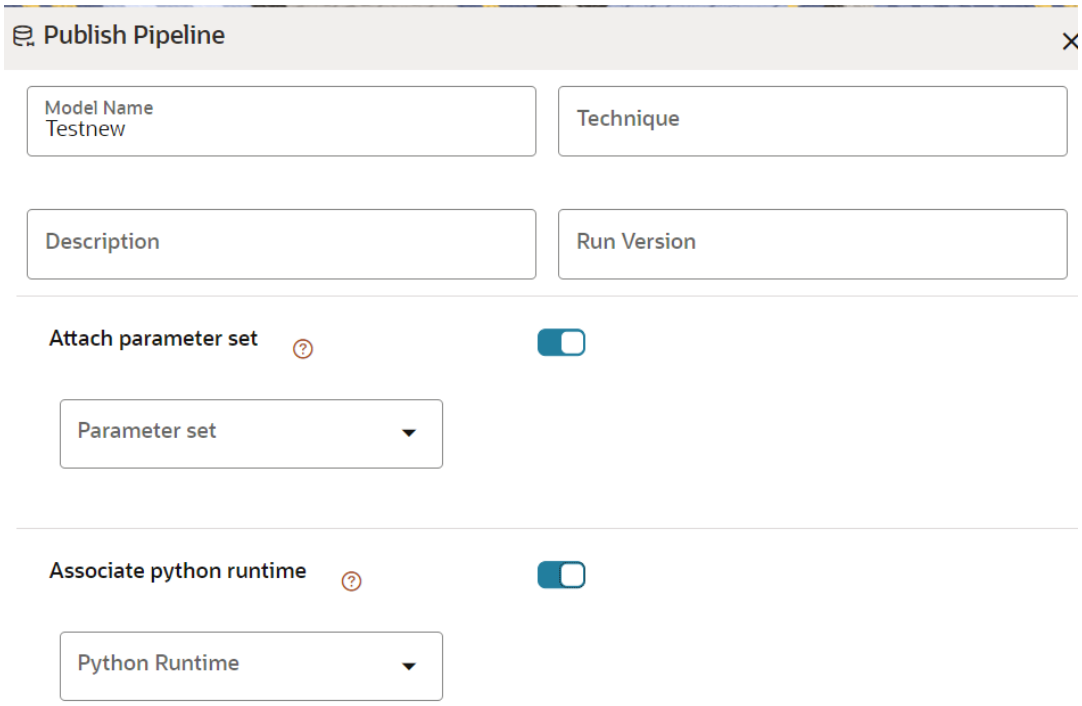
- Similarly, if you want to execute this Notebook for training purpose, then the flow will be executed till dataset with default paragraphs. To perform this, Click Execute Notebook and select Links as Training.

6.5.4.1.1.2.8 Publish a Pipeline

To publish the pipeline:

1. Click **Publish** to view **Publish Pipeline** window.

Figure 6-74 Publish Pipeline



2. Enter the details as shown in the following table.

Table 6-19 Field and its description for the Publish popup window

Field or Icon	Description
Model Name	The field displays the name of the Model. Modify the name if required.
Model Description	The field displays the description for the Model. Enter or modify the description if required.
Technique	Enter the registered technique to use.
Run Version	Select a run version.

Table 6-19 (Cont.) Field and its description for the Publish popup window

Field or Icon	Description
Attach Parameter set	If this option is enabled: • Selected parameter set will be associated with the published model. • Best performing parameter set or prime reference of the objective is recommended to be associated with the model. • This set is further used for promoting to production at the time of deployment. • Users can update this dependency from model details screen.
Associate python runtime	If this option is enabled: • Selected python runtime will be associated with the published model. • This python run time will be further used for promoting the model to production at the time of deployment. • Users can update this dependency from model details screen.

3. Click **Publish**.

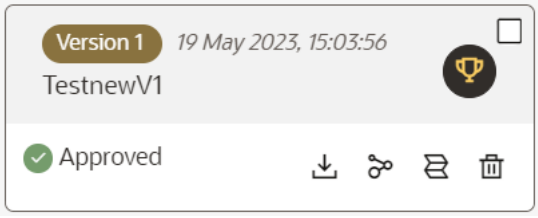
6.5.4.1.1.2.8.1 View Model Details

You can view model information for deployed models, models that require approval, and so on.

To view model details:

1. Click **Launch Workspace** next to the corresponding Workspace to display the **Dashboard** window with application configuration and model creation menu.
2. Click **Model Pipelines** to display the **Model Pipelines** window.
   The window displays folders that contain models and model records in a table.
   - The icon indicates that Model 1 is the champion.
3. Hover over the model records to view the various icons.

Figure 6-75 Mouse over the Model



The icon actions are listed in the following:

a. Download
b. Open in Pipeline Designer. See the Create Paragraphs using Pipeline Designer section for more details.
c. Scope Detail. See the Scope Detail section for more details.

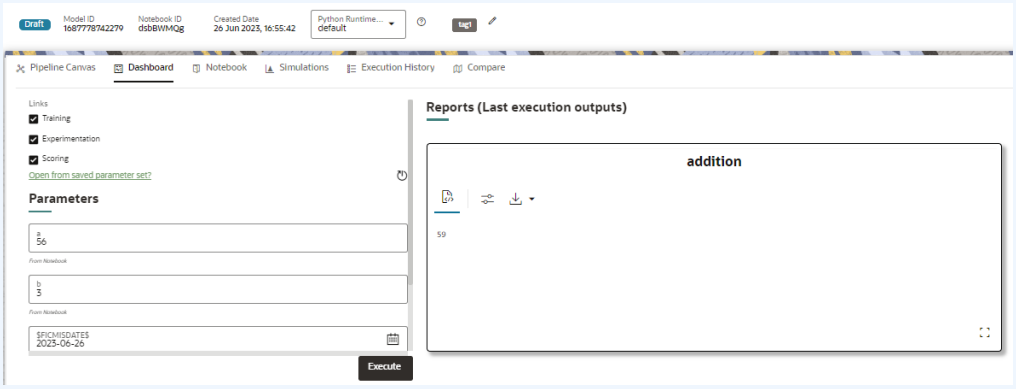
d. Delete Model.

6.5.4.1.1.3 Dashboard

This section of Pipeline Designer allows users to execute the Models and also shows the execution output of Model if the widgets are saved with Track output option enabled.

> **Note:**
>
> For Promoted models, the 'Dashboard' is the default tab.

Figure 6-76 Dashboard



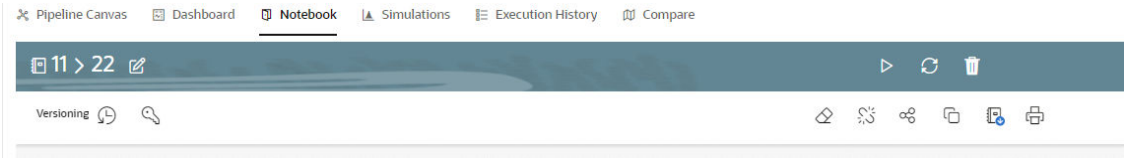
> **Note:**
>
> There is no Cancel button in Settings tab of reports in the Dashboard tab. You can press 'Escape' key to close the Settings.

6.5.4.1.1.4 Notebook

> **Note:**
>
> The Username is case sensitive. Ensure you use the correct case for the username when accessing and executing the notebook.

Navigate to **Notebook** page to view the paragraphs. You can run, invalidate session, edit, add, export the notebook, and so on.

Figure 6-77 Notebook page



> **Note:**
>
> The following features on Notebook are not supported in the current release.
>
> - Cloning the notebook
> - Sharing the notebook
> - Versioning the notebook
> - Modifying the notebook
> - Deleting the notebook
> - Attaching credentials to the notebook
> - Entering dependency modes to the notebook

6.5.4.1.1.4.1 Create Paragraphs in Model Studio Notebooks

After creating the Models in the Workspace, create Paragraphs in the Model Studio window. To create Paragraphs, you must have a working knowledge of scripting and Python.

Navigate to **Notebook** page to view and modify the paragraphs.

6.5.4.1.1.4.1.1 Create Paragraphs

The paragraph should be added after the start widget paragraph. Navigate to Notebook page to create the paragraphs.
The following types of paragraph creation are supported for model creation:

- Python Paragraph
- Shell in Python Paragraph
- Conda Paragraph
- PGQL Paragraph
- PGX - Java Paragraph
- Paragraph
- PGX - Python Paragraph
- PGX – Algorithm Paragraph
- PGX - PySpark Paragraph
- MySQL Paragraph
- Markdown Paragraph

To create Paragraphs in the Model Studio Notebooks:

1. Open the draft in Pipeline Designer and click **Notebook** tab.
2. Hover your mouse above or below a paragraph to display the interpreter toolbar.

Figure 6-78 Interpreter toolbar

3. Click an icon in the toolbar to select an interpreter and add a new paragraph that uses that interpreter. When you click some of the icons, the paragraph that's created includes example code that you can run.

> **Note:**
>
> You should not delete the start widget paragraph or rewrite the contents of this seeded paragraph.

6.5.4.1.1.4.1.1.1 Add Python Paragraph

- Click **Add Python Paragraph** in Model Studio to add a python paragraph in the Notebook and add the following scripting instructions.

  1. To fetch connection objects:
     `conn = mmg.getConnection(<workspace name>)`

     This creates a cx_oracle based connection object to the datadom of the workspace being passed:

  2. To fetch current Notebook and Model Objective details, use the following predefined parameters:

     - `currentNotebookId` - fetches the current notebook ID
     - `objectiveId` - fetches the current objective ID

     To fetch the runtime parameters supported in Model Studio runtime or to access optional parameters passed from scheduler services:

     - To access predefined batch runtime parameters like `taskid`, `batchrunid`, and `ficmisdate` passed during the model execution from scheduler, use the following variables : $BATCHRUNID$, $TASKID$, $FICMISDATE$.
       Example:

       ```
       %python
       print('BATCHRUNID value is : ${$BATCHRUNID$}')
       print('TASKID value is : ${$TASKID$}')
       print('FICMISDATE value is : ${$FICMISDATE$}')
       ```

     - To access any optional parameters passed during the model execution from scheduler, use the below sample script:

       ```
       %python
       print('threshold value is : ${threshold}')
       ```

> **Note:**
>
> Threshold is the optional parameter passed during the model execution.

> **Note:**
>
> For more details on the predefined functions available in MMG for python scripting, see Appendix -I.
> Click the icon to run the script.

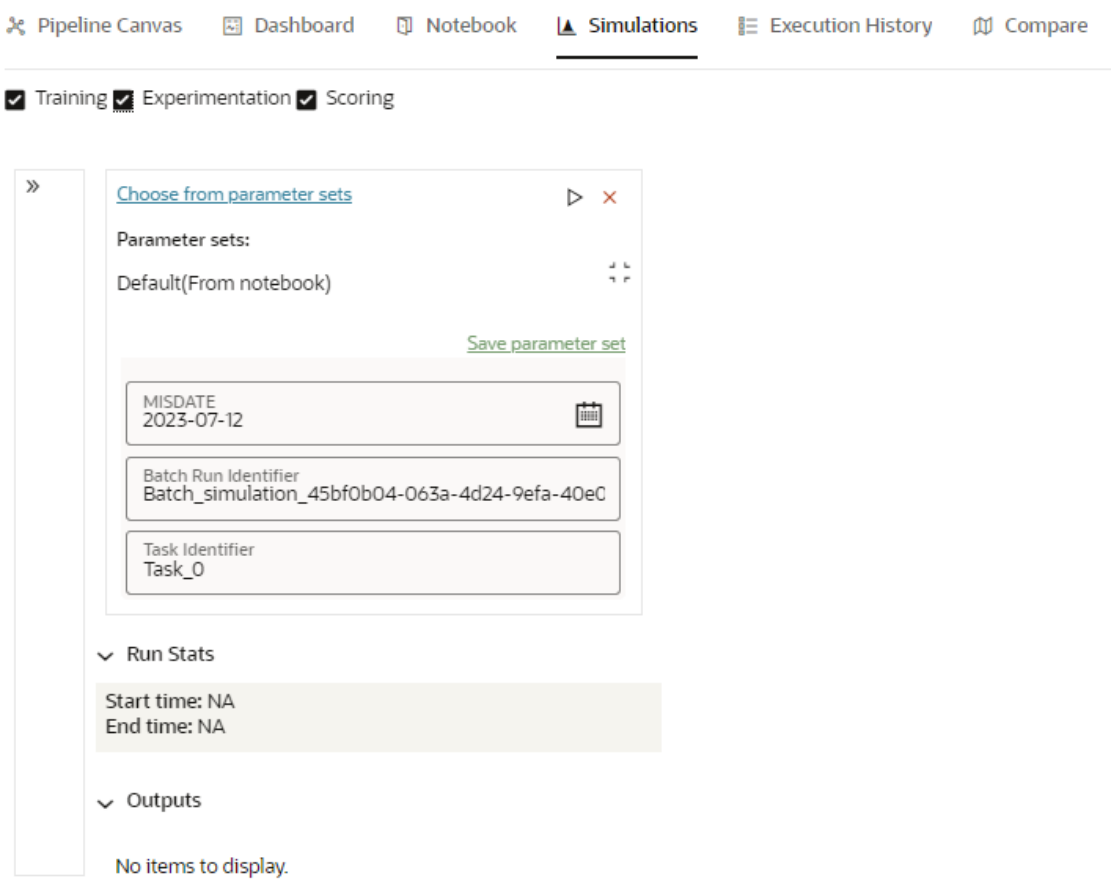
> **Note:**
>
> The execution result from the Pipeline Canvas screen is not displayed in the notebook tab.

6.5.4.1.1.5 Simulations

Simulation run is for executing a single notebook in parallel by giving different values for same parameters. The simulation flow allows for iterative execution along that path with input drivers (variables) that are passed through a parameter set.

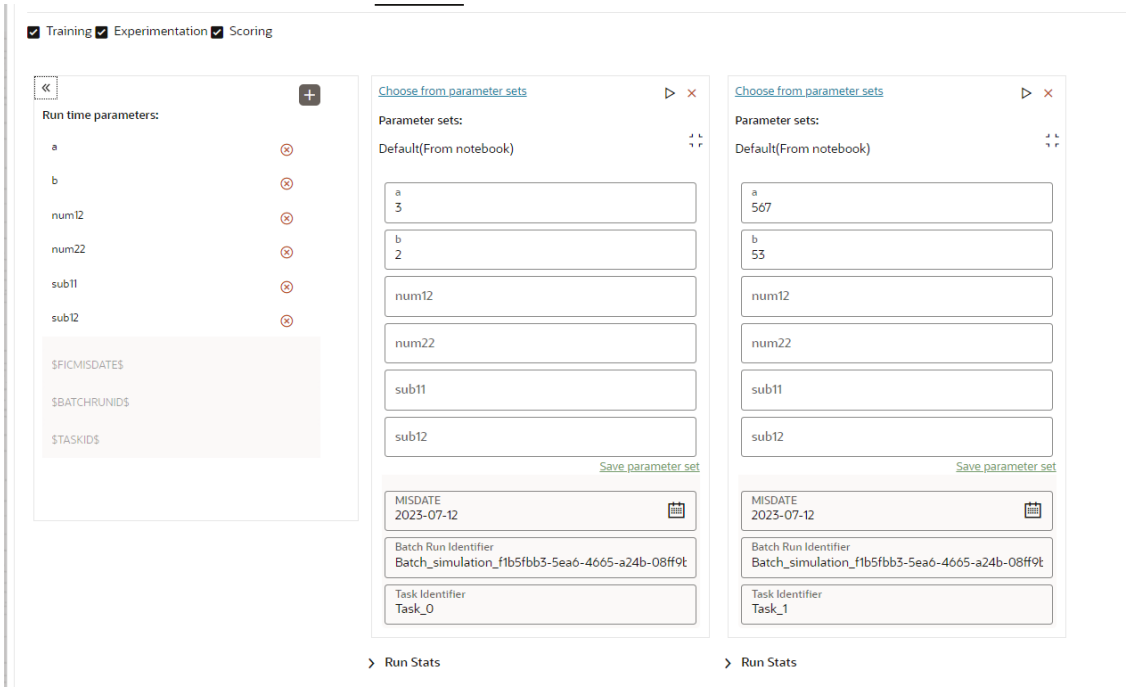
You can either create a new parameter set or use the existing parameter set and execute it from this tab. In addition, you can select or deselect the link types which you want to execute.

Figure 6-79 Simulation tab



If you want to add different values for the same parameter, click Add Run. You can add any number of Run based on your requirements and then click Execute all to execute at one time. The Batch Run Identifier will be same for all runs and the Task id will be incremented for newly created run.

Figure 6-80 Run Stats



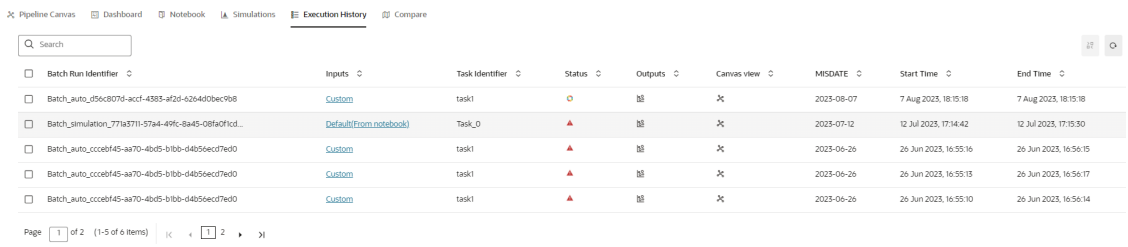
The Run Stats displays the execution start and end date. The Outputs of the tracked nodes will be shown under the inputs of each run panel.

6.5.4.1.1.6 Execution History

This section of Pipeline Designer shows the history of the executed pipelines.

> **Note:**
>
> Add file output button in execution history tab that opens file manager at a specific path exposed by the save dataframe widget in the pipeline window. When the save dataframe widget is not available in the pipeline, show that the path does not exist from file manager.

Figure 6-81 Execution History tab



You can compare and refresh the executions by clicking on **Compare**and **Refresh** icons.

To view the inputs, output, and canvas view, click on the corresponding options in the table.

Clicking the **Output** icon displays the Output Details page, where you can view "Executed-only" paragraph outputs. By default, this option is enabled, you can disable this option based on your requirements.

By default, **Tracked** option is selected that displays outputs tracked from the time of the execution. Select **All** option to view all the outputs.

6.5.4.1.1.7 Compare

This section allows you to compare the models with Champion model.

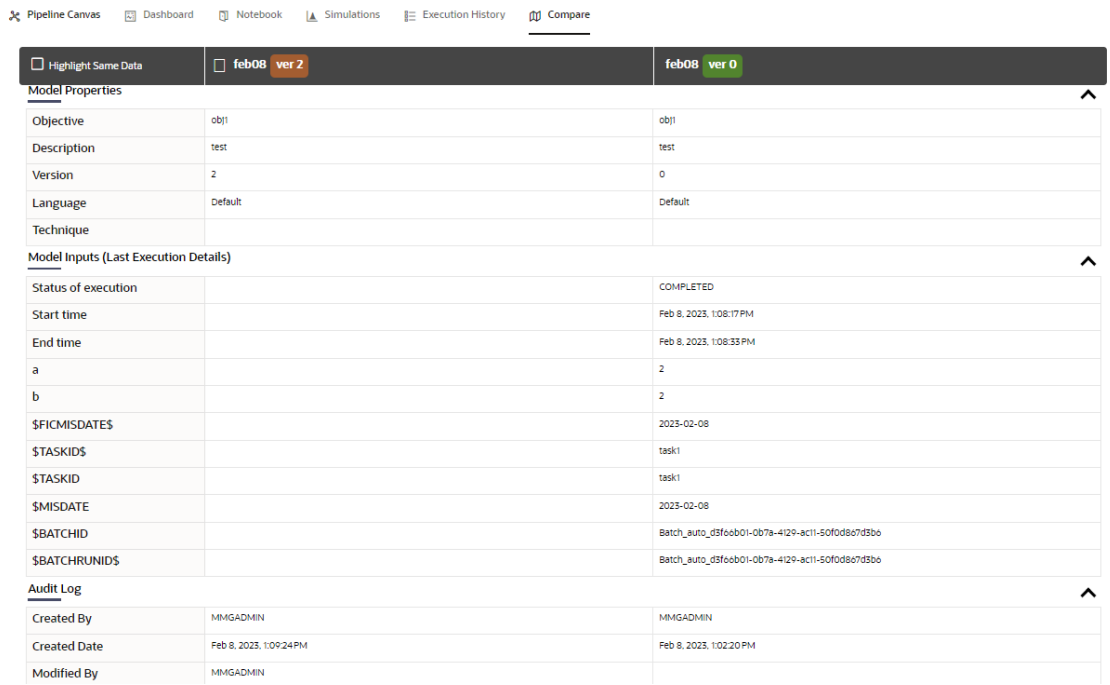
To compare:

Navigate to Pipeline Designer window and click **Compare** tab.

This shows the following comparison details:

- Model Properties
- Model Inputs (Last Execution Details)
- Audit Log
- Model Script
- Model Output

Figure 6-82 Compare tab



6.5.5 Publish Data Studio

After creating the Draft Models, publish the Notebooks, which have the Model script.

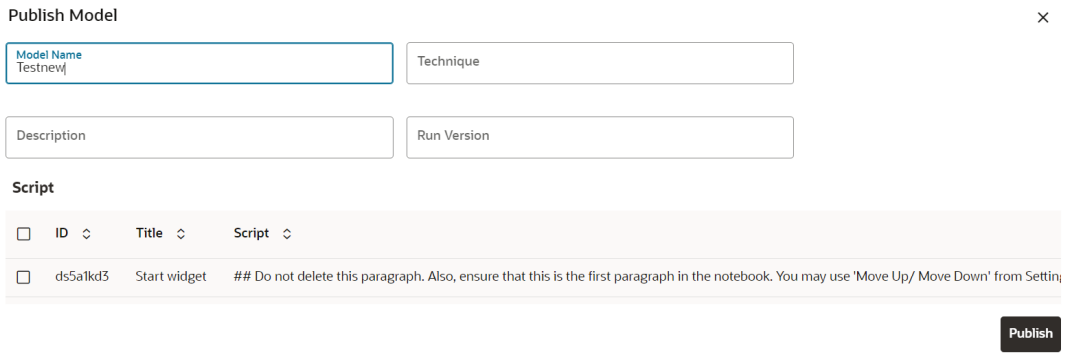
To create a Scoring Model:

1. Create a Draft Model. See the Create Draft Models section for more information.

2. Click next to corresponding Draft Model and select **Publish Data Studio** option.

   The Publish Model window is displayed.

Figure 6-83 Publish Model



3. Enter the details as shown in the following table.

Table 6-20 Field and its description for the Publish popup window

Field or Icon	Description
Model Name	The field displays the name of the Model. Modify the name if required.
Model Description	The field displays the description for the Model. Enter or modify the description if required.
Technique	Enter the registered technique to use.
Run Version	Select a run version.
Script	The table displays the Paragraphs created in the Training Model. Select the Paragraphs that you want to use to create the Scoring Model. Track Output - Select this to track the output of the paragraph.

4. Click **Publish**.

6.5.6 Scope Detail

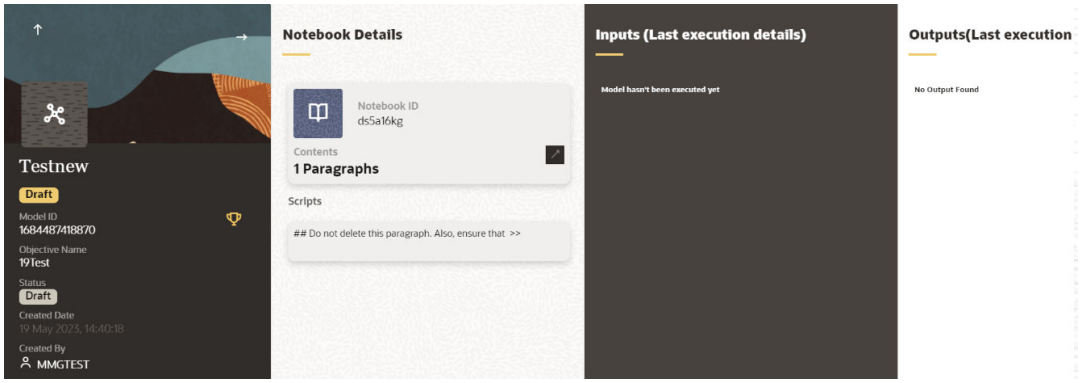
To view the scope details of the model:

1. Click **Launch Workspace** next to the corresponding Workspace to display the **Dashboard** window with application configuration and model creation menu.
2. Click **Model Pipelines** to display the **Model Pipelines** window.

   The window displays folders that contain models and model records in a table.
3. Click next to corresponding Model and select **Scope Detail** option. This is available for both Draft and Published models.

   The details such as Notebook, Inputs (Last execution details), Deployment Details, and Outputs (Last execution outputs) are displayed.

Figure 6-84 Scope Details



6.5.6.1 View Model Details

You can view model information for deployed models, models that require approval, and so on.

To view model details:

1. Click **Launch Workspace** next to the corresponding Workspace to display the **Dashboard** window with application configuration and model creation menu.
2. Click **Model Pipelines** to display the **Model Pipelines** window.

   The window displays folders that contain models and model records in a table.

   The icon indicates that Model 1 is the champion.
3. Hover over the model records to view the various icons.

   The icon actions are listed in the following:

   a. Download

   b. Open in Pipeline Designer. See the Create Paragraphs using Pipeline Designer section for more details.

   c. Scope Detail. See the Scope Detail section for more details.

   d. Delete Model.

6.5.7 Model Governance

After comparing models, you must understand the Model Governance system in the application. The Model Governance has an impact on how the application functions with the various user types and the requests they can place from the Model Details window. You require to understand Model Governance before you request model acceptance, review models, or approve models for production.

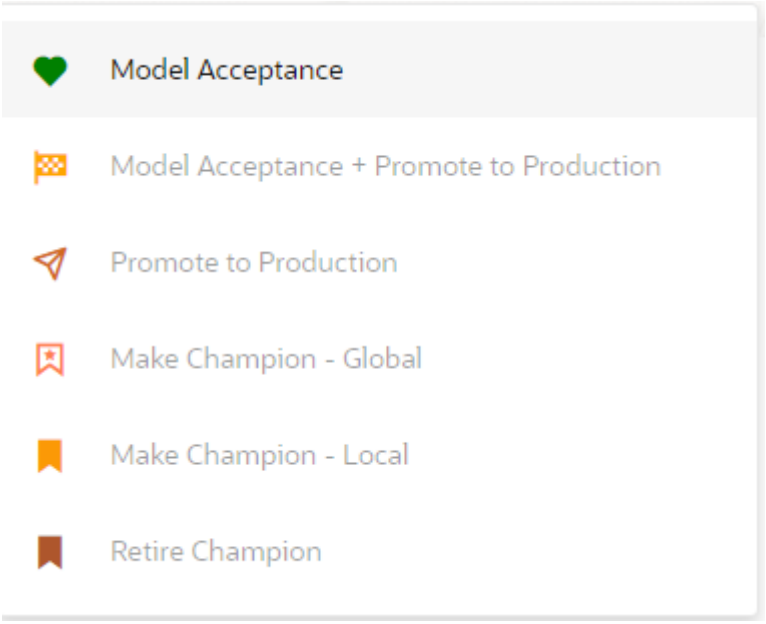
As discussed earlier, the users are of three types:

- Requester
- Reviewer
- Approver

The request consists of the following phases in the Request drop-down list (see the Request Model Acceptance section for how to access the drop-down list in the Model Deployment window):

- Model Acceptance
- Model Acceptance + Promote to Production
- Promote to Production
- Make - Champion - Global
- Make - Champion - Local
- Retire Champions

Figure 6-85 The Request Drop-down List



The values in the drop-down list are active based on the type of user (Requester, Reviewer, or Approver) and the phase that the model is in (accepted, promoted to production, global champion, and so on). Let us look at these with a few examples.

Example 1:

Assume that you are a user with Requester privileges, and you create a model. Now you can request for the model to be accepted on the Model Details window from the **Request** drop-down list. The values enabled for selection are **Model Acceptance** and **Model Acceptance + Promote to Production**. Let us proceed and assume that you select Model Acceptance, then a user with Reviewer privileges forwards your model to a user with Approver privileges. At this stage, the Approver can choose to reject or accept your model acceptance request. A rejection would bring the model back to the initial state with comments on the updates required before it can be requested for acceptance again. However, if the Approver accepts your model, then the **Make Champion- Local** selection is enabled when you log in. You can create many models and send them for acceptance. After acceptance, any model that is accepted can be made the champion on your local workspace at any time replacing the earlier local champion.

Example 2:

Assume that in the previous example, you selected **Model Acceptance + Promote to Production**, then a Reviewer forwards it to the Approver. The Approver, at this stage, chooses to promote the model to production by selecting **Promote to Production**. The model is now available in the production environment and the Approver can choose at any time to select a

model from these models in production and select **Make Champion- Global**. If there exists a Champion model in the production environment, then it will be replaced by the new global champion. However, the earlier champion will still be available in the production environment along with other models and the Approver can choose to make it the global champion again at any time or select any of the other models and make one of them the global champion.

6.5.7.1 Request Model Acceptance

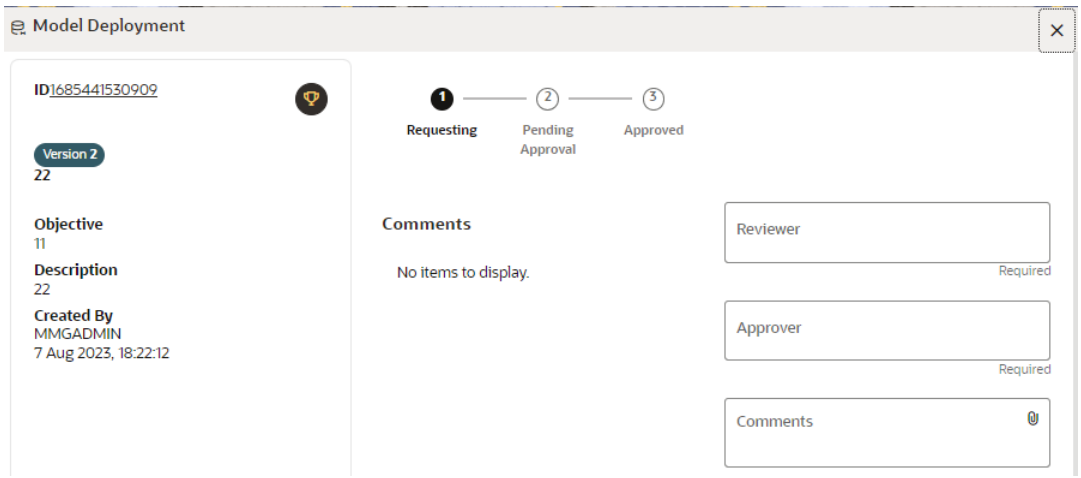
After comparing models, move the selected models to acceptance. Only a user with the Requester role can request for model acceptance. The model will be moved to review which will be available to Reviewer and Approver role, and then to acceptance is available to users with the Approver role, who can promote to production. See the Understand Model Governance section before you start here.

To request a model to promote to production, follow these steps:

1. Click **Model Pipeline** to display the **Model Pipeline** window.

   The window displays folders that contain models and model records in a table.
2. Click on the Deploy icon on the Model to open it.
3. If a model is a champion, the icon is displayed on it.
4. To make a model a champion from this window, place the mouse in the selected model columns. and click to display the **Model Details** window. If a model is a champion model, the icon is displayed.

Figure 6-86 Model Deployment



5. Select the Reviewer group from the **Reviewers** drop-down list.
6. Select the Approver group from the **Approvers** drop-down list.
7. Enter comments in Comments and click **Attach** to attach files supporting the comments.
8. Use the following features on the window to perform additional actions.
   - View the model status change in the progress indicator. The Progress Indicator displays the various states of progress that the model has been through. Accordingly, you must request, review, or approve models.

Figure 6-87 Model Approval Progress Indicator

- Click the type of **Request** from the drop-down list:
  - Model Acceptance: To review and accept the model creation.
  - Model Acceptance + Promote to Production: To review and promote the model to production.
  - Make Champion- Local: If the model is not the champion model, select to make it the local champion.
  - Promote to Production: To promote a model to production
  - Make Champion- Global: If the model is not the champion model, select to make it the Global champion.
  - Retire Champion: To retire a Champion model
- Comment History: A record of comments entered in the cycle of model creation and approval with the feature to download attachments.

The model sent for acceptance or for promotion to production is now displayed to a Reviewer to review it and then to Approver when signed in, who must either accept the request or reject it.

6.5.7.2 Review Models and Move to Approve or Reject

The Reviewer must provide comments describing the action (approve or reject). If comments are related to rejection and if the Approver rejects, then model goes back to the Requester to make changes or to delete it. If comments are related to approval, then model moves further in the workflow and is displayed to an Approver. See the Understand Model Governance section before you start here.

To review a model:

1. Click **Launch Workspace** next to the corresponding Workspace to display the **Dashboard** window with application configuration and model creation menu.
2. Click **Model Pipelines** to display the **Model Pipelines** window.

   The window displays folders that contain models and model records in a table.
3. Click to display the **Model Details** window.
4. Review the details and send it back to the Requester for modifications or send it to an Approver.

6.5.7.3 Approve Models and Promote to Production

The models reviewed and set to promote to production by either the Requester or Reviewer is displayed to the Approver when signed in. The Approver has to either reject the model and send it back to the requester with supporting comments or approve it for pushing to production. See the Understand Model Governance section before you start here.

> **Note:**
>
> When dataset has used the datasource which is of order (N) for example 5 , and the Production workspace does not contain the datasources at order 5, then promotion of models containing dataset from Sandbox to Production workspace fails.
> To remedify this issue, ensure that Sandbox and Production workspace contain the same number of Datasources before you perform promotion of models.

To approve or reject models:

1. Click **Launch Workspace** next to the corresponding Workspace to display the **Dashboard** window with application configuration and model creation menu.
2. Click **Model Pipelines** to display the **Model Pipelines** window.

   The window displays folders that contain models and model records in a table.
3. Select a model from the records in the objective.
4. Click to display the Model Details window.
5. Click **Approve** or **Reject** with appropriate comments.

6.5.7.4 Deploy Models in Production and Make it a Global Champion

After approving the models, deploy it to the production environment. You must have an Approver function role and privileges to do this activity.

> **Note:**
>
> Sandbox should have the production workspace attached in order to have this option enabled.

To deploy Models in production:

1. Click **Launch Workspace** next to the corresponding Workspace to display the **Dashboard** window with application configuration and model creation menu.
2. Click **Model Pipelines** to display the **Model Pipelines**window.

   The window displays folders that contain models and model records in a table.
3. Select the model to deploy and click **Model into Champion** to display the details of the model. If a model is a champion, the **Champion** icon is displayed.

6.5.8 Import a Model Data into a New Model

The model data from existing models in .dmp format and the existing model data in .dsnb can be imported during the creation of a new model.

> **Note:**
>
> The import should happen inside an Objective only.

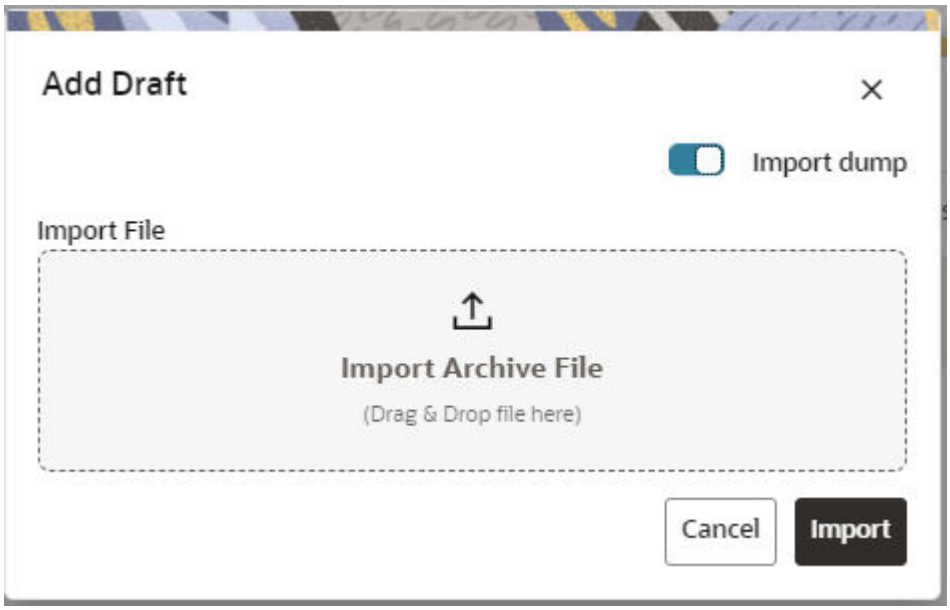
The import of model data lets you reuse and extend on model creation. This topic is part of the procedure of creating Draft Models and after creating a new model using this method, see the Create Draft Models section for instructions on how to proceed further.

To import model data:

1. Click **Launch Workspace** next to corresponding Workspace to Launch Workspace to display the **Dashboard** window with application configuration and model creation menu.
2. Click **Model Pipelines** to display the **Model Pipelines** window.

   The window displays folders that contain models and model records in a table.
3. Select an **Objective**.
4. Double click a model to display the model versions in the expanded display.
5. Hover over a model and click **Download** to download the model data dump.
6. Click **Add** and **Drafts** to display the **Model Details** dialog box for the creation of a new model.

Figure 6-88 Model Details - Import Dump



7. Drag the toggle switch to select **Import Dump**.
8. Drag and drop the file into the **Import Dump File** field or click in the box to open the file selector dialog and select a file.

> **Note:**
>
> If the upload file size is more than 8 MB, you can increase the file size upto 100 MB by modifying the below properties. Restart of services are required once the file size is updated.
> mmg-ui/conf/application.properties:
>
> **********
>
> spring.servlet.multipart.max-file-size=8MB
>
> spring.servlet.multipart.max-request-size=8MB
>
> mmg-service/conf/application.properties:
>
> **********
>
> spring.servlet.multipart.max-file-size=8MB
>
> spring.servlet.multipart.max-request-size=8MB

9. Click **Import**. A new model is created by importing the model data dump of another model.

6.5.9 Execute Models using Scheduler Service

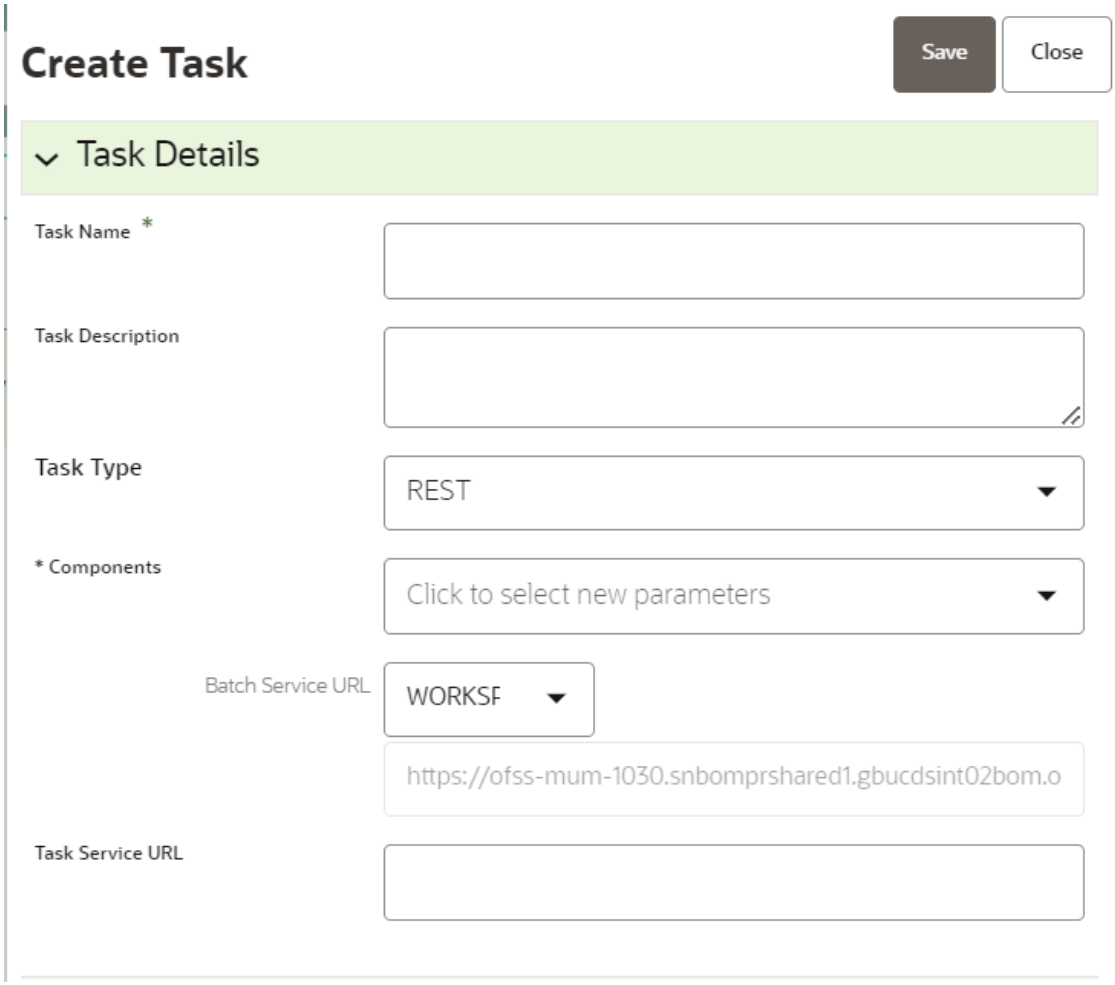
The models that you have created require that they are executed using Scheduler Service before they can be available to the users of OFSAA applications. For more information on this, see the Scheduler Service.

6.5.9.1 Define a Task

The following three Components are supported during the task creation:

- Model
- Populate Workspace
- Custom

Figure 6-89 Create Task



Enter the following details in Task Details section:

- **Task Name**: Enter the task name.

> **Note:**
>
> The Task Name must be alphanumeric and should not start with a number. The Task Name should not exceed 60 characters in length. The Task Name should not contain any special characters except underscore (_).

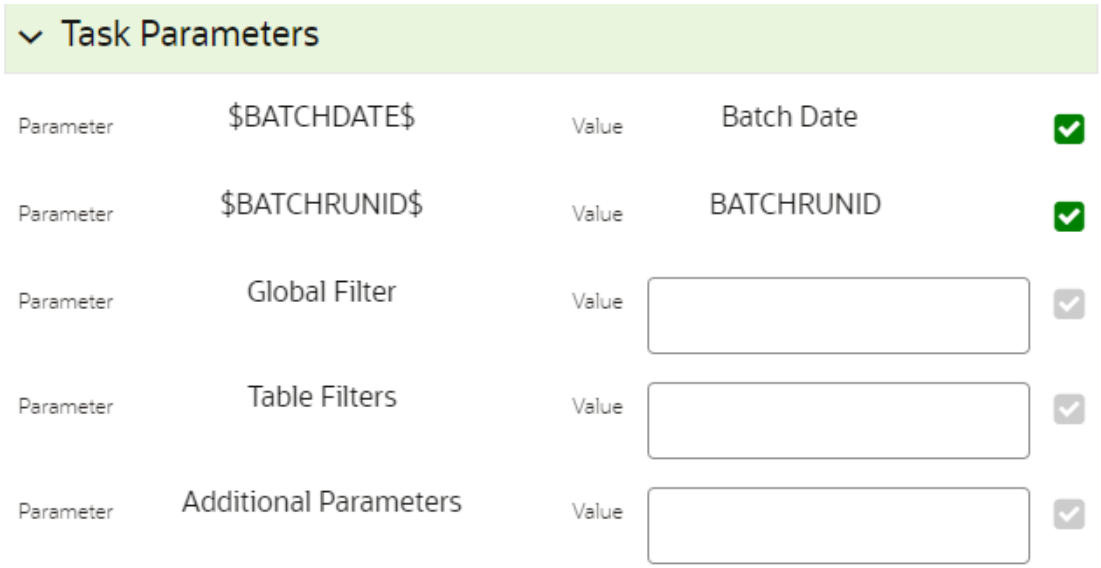
- **Task Description**: Enter the task description. No special characters are allowed in Task Description. Words like Select From or Delete From (identified as potential SQL injection vulnerable strings) should not be entered in the Description
- **Task Type**: Select the task type from the drop-down list. The options are REST and SCRIPT. You can enter Shell script for Model, Populate Workspace, Custom components. Status key in the curl command should be in uppercase as STATUS.
- **Batch Service URL**: Select the required Batch Service URL from the drop-down list. This can be blank, and you can provide the full URL in the Task Service URL field.
- **Task Service URL**: Enter task service URL if it is different from Batch Service URL.

> **Note:**
>
> Task Parameters will vary based on the selected Component.

6.5.9.2 When Component is Model

The following window shows the Task Parameters for Model Component.

Figure 6-90 Task Parameters



> **Note:**
>
> Fields marked with * are mandatory fields.

- **Batch Date**: Shows the batch execution date. You can enable or disable this parameter.
- **Batch Run ID**: Shows the batch execution run ID. You can enable or disable this parameter.
- **Objective**: Select the Object which you want to use for execution. For more information, see the Create Objective (Folders) section. The Sub Objective is displayed with path. For example, if Test1 is Objective and Test11 is Sub Objective, and you want to use Test11 Objective for execution, then select this field as Test1/Test11.
- **Model**: Select the Model of selected Objective. It can be any specific model of Objective or All models of Objective.
  - If the ALL_CHAMPION is selected here, then Objectives with no Champion Model is skipped, and the Objectives with Champion Models gets executed.
  - If CHAMPION is selected, and no Champion is present, then Model Execution gets fail.

- **Link Type**: Select the link type for execution. For example: Training, Scoring, or Training+Scoring. For more information, see the Links in the Pipeline Designer section.
- **Synchronous Execution**: You can set this parameter to Yes or No.
  - If Synchronous Execution is set to Yes, then execution will wait for the notebook execution status.
  - If Synchronous Execution is set to No, then execution will not wait for the notebook execution status, it will trigger the notebook and update task status as successful in batch monitor.
- Optional Parameters: This is used pass the parameters dynamically.
  For example:

  model_group_name=LOB1,benford_flag=Y,benford_digit=1,from_date=01-Jul-2020,to_date=31-Jul-2021

  Model_group is parameter defined in model and value can be passed here.

The Create Task window also shows the following Header Parameters, which are not editable:

- User: logged in user name
- Workspace: shows the launched workspace name
- Locale: shows the locale. For example: en_US

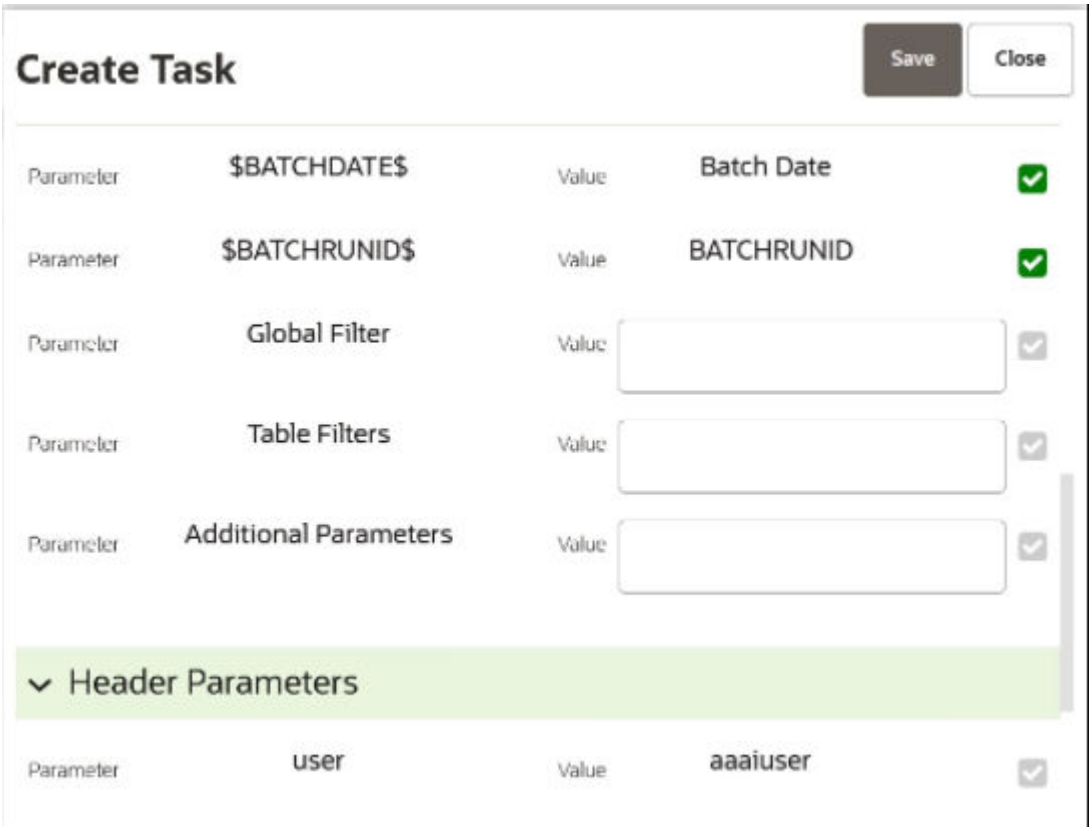
6.5.9.3 When Component is Populate Workspace

If you select Component as Populate Workspace, then to use the data population, enter the following parameters.

- Additional Parameters
- Data Filters - Global level
- Data Filters - Table level
- Data Filters - Hint

For more information, see the Populating Workspace section.

Figure 6-91 Populate Workspace screen



6.5.9.4 SS $RUNSK support for processing Delta Data by Data Pipeline and Match Rules

A new parameter to pass RUNSK is being implemented in the Scheduler UI.

Dynamic Parameters screen for new batches are created.

1. RUNSKEY will be present in Task Parameters while copying task for existing batches.
2. RUNSKEY should be visible in View Execution Parameters - Task Parameters in Monitor screen.
3. RUNSKEY will change just as BatchRunID in "Rerun" Case.
4. RUNSKEY will remain same just as BatchRunID in "Restart" Case.
5. RUNSKEY will be stored for batches in "AAICL_SS_BATCH_RUN" in column "N_RUNSKEY".

6.5.10 Export/Import Models via Utility

You can import and export the models between Sandbox to Production or vice versa using the Utility.

Prerequisite:

Before you import, ensure the Model artifacts are available in the <installed path>/ftpshare.

To import the models:

- Navigate to `<MMG_PACK>OFS_MMG/bin` and execute in the following format:

  ```
  ./model_export_import_utility.sh IMPORT WORKSPACE LOGIN_USER LOCALE FILE_NAME Y/N
  ```

  Example:

  ```
  ./model_export_import_utility.sh IMPORT CS SAUSER en_US CS_1638105398036_0.zip Y
  ```

Table 6-21 Fields and icons on the Sandbox Summary page

Field	Description
Import	Command to import the Model into Workspace
Workspace	The name of the Workspace to which you are importing the Model
Login_User	The logged in user name.
Locale	The application language preferences For example: en_US
File_name	The file name with the following format: Workspace name_Model ID_Model version.zip For example: `CS_1638105398036_0.zip`

> **Note:**
>
> If you enter input as **Y**, the utility imports the models in the approved status.
> If you enter as **N**, the models are imported but not in the approved status.

To export the models, perform the following steps:

Prerequisite:

Before you export, ensure the Models and Drafts are available in the UI / setup.

1. Navigate to <MMG_PACK>OFS_MMG/bin and execute in the following format:
   `./model_export_import_utility.sh EXPORT WORKSPACE LOGIN_USER LOCALE MODELID MODEL_VERSION`

   Example:

   ```
   ./model_export_import_utility.sh EXPORT CS SAUSER en_US 1638105398036 0
   ```

Table 6-22 Fields and icons on the Sandbox Summary page

Field	Description
Export	Command to export the Model from the Workspace
Workspace	The name of the Workspace from which the model is exported
Login_User	The logged in user name.
Locale	The application language preferences. For example: en_US
Model ID	The Model ID

Table 6-22 (Cont.) Fields and icons on the Sandbox Summary page

Field	Description
Model Version	The version of the Model

6.5.11 View Models

The View Models feature launches the OFS Data Studio window. You can view models on this window.

To use View Models:

1. Click **Launch Workspace** next to the corresponding Workspace to display the **Dashboard** window with application configuration and model creation menu.
2. Click **Model Pipelines** to display the **Model Pipelines** window.

   The window displays folders that contain models and model records in a table.
3. Click **Action** next to corresponding Model and select Open in Pipeline Designer.

   See the Create Paragraphs in Pipeline Designer section for details on how to use the OFS Data Studio.

6.5.12 Edit Models

The editing of models created versions that are different from the previously saved model and the cycle of Model Governance applies to any edited model. You can edit models from the OFS Data Studio window using Python scripting language.

You can edit the script of version 0 only even in Pipeline Designer and Studio. It is not possible in other versions.

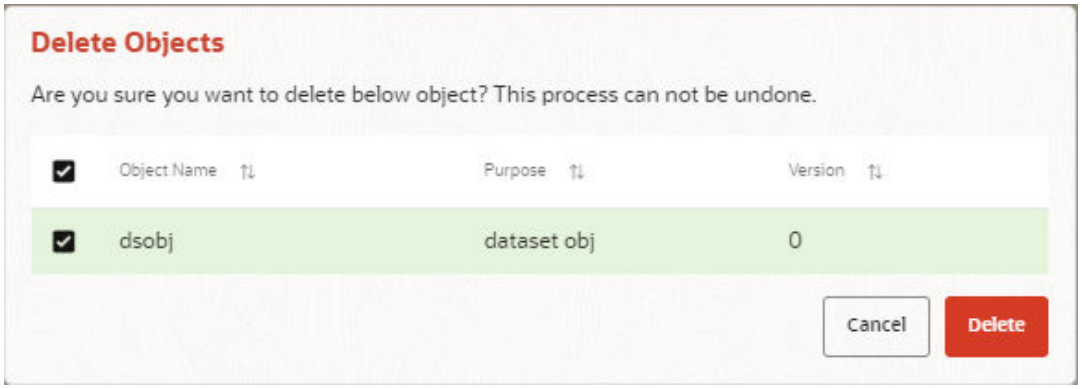
6.5.13 Delete Objectives and Draft Models

To delete Objectives and Models that exist in the Objectives, follow these steps:

1. Click **Launch Workspace** next to the corresponding Workspace to display the **Dashboard** window with application configuration and model creation menu.
2. Click **Model Pipelines** to display the **Model Pipelines** window.

   The window displays objectives that contain models and model records in a table.
3. Click **Action** next to corresponding record (objective/draft) and select **Delete**.
4. You can select two or more objectives or models from the records.
5. Click **Delete** to display the Delete dialog box.
6. Click **Delete** to delete or click **Cancel** to cancel.

Figure 6-92 Delete Objects screen



> **Note:**
>
> If the Model is promoted to Production, you cannot delete the model even it is retired.

6.5.14 Programmatic Dynamic Forms and Fixed Dynamic Forms

The Programmatic Dynamic Forms and Fixed Dynamic Forms scripts are supported.

A dynamic form is a user input field that is generated from the code of a paragraph. Dynamic forms allow users to bind free variables in a paragraph.

For Example:

Textbox

The Textbox dynamic form allows users to input any string of characters. Its format is as follows:

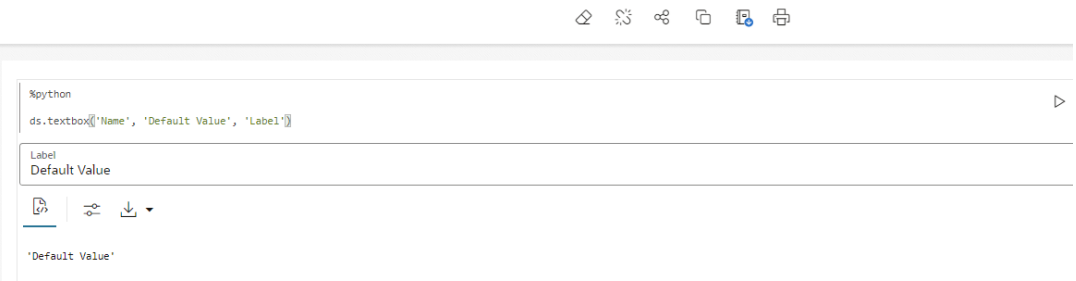
- Click **Add Python Paragraph** in Model Studio to add a Paragraph in the Notebook and fetch the runtime parameters supported in Model Studio runtime as shown in the following example. This is available for all the interpreters.

  For example, enter as follows:

  ```
  %python
  ds.textbox('Name', 'Default Value', 'Label')
  ```

  Click **Play** to run the script and display.

Figure 6-93 Run the script and display tab



> **Note:**
>
> In the current release, the MMG Studio supports Textbox, Select, Slider, Checkbox, Date Picker, Time Picker and DateTime Picker dynamic forms.

6.6 Graphs

The Graph Pipeline feature enables you to transform data in relational tables into a graph.

This feature uses the latest technology to harness the power of Graph Analytics to give Financial Institutions the ability to monitor the data financial institutions effectively. The data is organized as nodes, edges, and properties (property data is stored on the nodes or edges). The results of analytics algorithms are stored as transient properties of nodes and edges in the Graph.

Users can harness the power of Graph Analytics using our in-built in-memory Oracle Graph Analytics Engine (PGX).

The Graph Pipeline functionality allows users to define graphs easily, attach underlying data, and match pipelines to populate data in the graph.

The Graph Pipeline functionality allows users to quickly create and configure a graph for use in advanced analytics. It can also manage and schedule the tasks required to run populate the graph on a periodic basis.

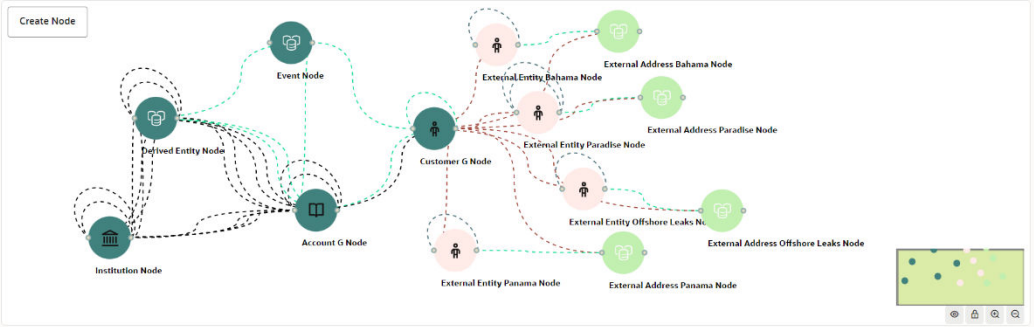
The Graph Pipeline functionality allows you to:

- Creating Graph Model from their existing relational data model
- Configuring the Graph Model through Pipeline UI
- Creating and Scheduling the data pipeline and matching (creation of similarity relationships) for created Graph Model
- Adding new sources and contextualizing the links quickly
- Standardizing the data before pushing it to Graph Model
- Using the following data source for Graph Model:
  - Oracle
  - File System (CSV)
- Using BD data source definitions for pre-seeded Workspace
- Using pre-configured Financial Crime Graph Model pipelines

- Using pre-seeded mappings of Graph Model properties with BD Data source properties
- Scheduling to create Graph by running pre-configured batches
- Blending data and getting insights from data quickly
- Load data into elastic search indexes so matching and entity resolution can occur in the graph
- Define matching rules for the generation of similarity edges in the graph.

An example of a preconfigured Financial Crime Graph Model with Nodes and Edges is provided here.

Figure 6-94 Example of Graph Model



The Graph Model defines the nodes and the relationship between them:

- **Node**: Represents a single entity with its attributes. The entity can be a single table or combination of tables merged on some common feature, it could be from a file, or any generalized component of multiple substructures merged for a specific reason/objective. For example, Customer.

  A customer could be a single entity.

  Customer, Salesperson, Manager, and so could be individual entities that can be structured together as a “Person” entity with a set of attributes persisted across individual entities.
- **Edge**: A connector between nodes or the same Node itself. Each edge can have a set of attributes associated with it. These edges will map to 'join' between two entities that could be direct or transitive.
- **Modeler Attributes**: Properties of the Node/Edge, and these are mandatory inputs. For example:

  A Customer Node will have properties like Customer ID, Name, Age, Gender, etc.

  A Transaction edge will have properties like Transaction ID, Date, Amount, From/To account, etc.

See the Adding a Graph Pipeline section for creating Graph, Nodes, Edges, and Scheduling.

6.6.1 Accessing the Graph Summary Page

The Graph Summary page gives access to the various graph pipeline functions such as create, view and delete.

To access the Graph Summary page, follow these steps:

1. Click **Launch Workspace** next to the corresponding Workspace to display the **Dashboard** window with application configuration and model creation menu.
2. Click the **Graphs** icon.

   This displays the graph pipeline in a tabular format.

   The following table provides descriptions for the fields and icons on the Graph Summary page.

Table 6-23 Graph Summary Details

Field or Icon	Description
Search	The field to search for a graph pipeline. Enter specific terms in the field for which you want to search, and press Enter on the keyboard to display the results.
Name	The name of the graph pipeline.
Number of Nodes	The number of nodes present in the graph pipeline.
Number of Edges	The number of edges present in the graph pipeline.
Owner	The owner of the graph pipeline.
Created Date	The date on which the graph pipeline was created.
Delete	Click to delete multiple graph pipelines.
Add	Click to create a new graph pipeline. See Adding a Graph Pipeline section.
Action	Click **Action** icon to View/Edit/Delete/Download functions on the selected graph pipeline.

> **Note:**
>
> The user is now able to execute/schedule a batch with SYSDATE-X days functionality.

6.6.1.1 Adding a Graph Pipeline

To add a graph pipeline:

1. Navigate to Graph Summary page.

   This page displays the graph summary records in a table.
2. Click **Add**.

   The **Add Graph**pop-up window is displayed. Figure: Add Graph window.

Figure 6-95 Add Graph window

3. In the **Graph Name** field, enter a name for the graph. The name must be unique to a particular workspace.
4. In the **Graph Description**field, enter the description for the graph and click Add.

   OR

   You can import an existing Graph. Use the toggle button to select **Import Dump**.

   The following page is displayed.

   Browse the file and click **Import**. Once the Import is successful, the Graph Model page is displayed.

   The Maximum Age of Old session for a graph is 7 days by default. The Maximum age of old session of 7 days specifies that graph would be retained for a period of 7 days. You can modify the description and Max Age of Old session by clicking on the **Setting** icon. If you want the batch to be created in read-only mode for scheduler screen, enable the option.
5. Click **Save**.

   The user session of the Graph Pipeline will get refreshed after the set timeline.
6. You can perform the following:
   a. Create a Node
   b. Create an Edge
   c. Toggle drawing which will enable or disable the user from dragging components
   d. Zoom In and Zoom Out

6.6.1.2 Create a Node

To create a Node:

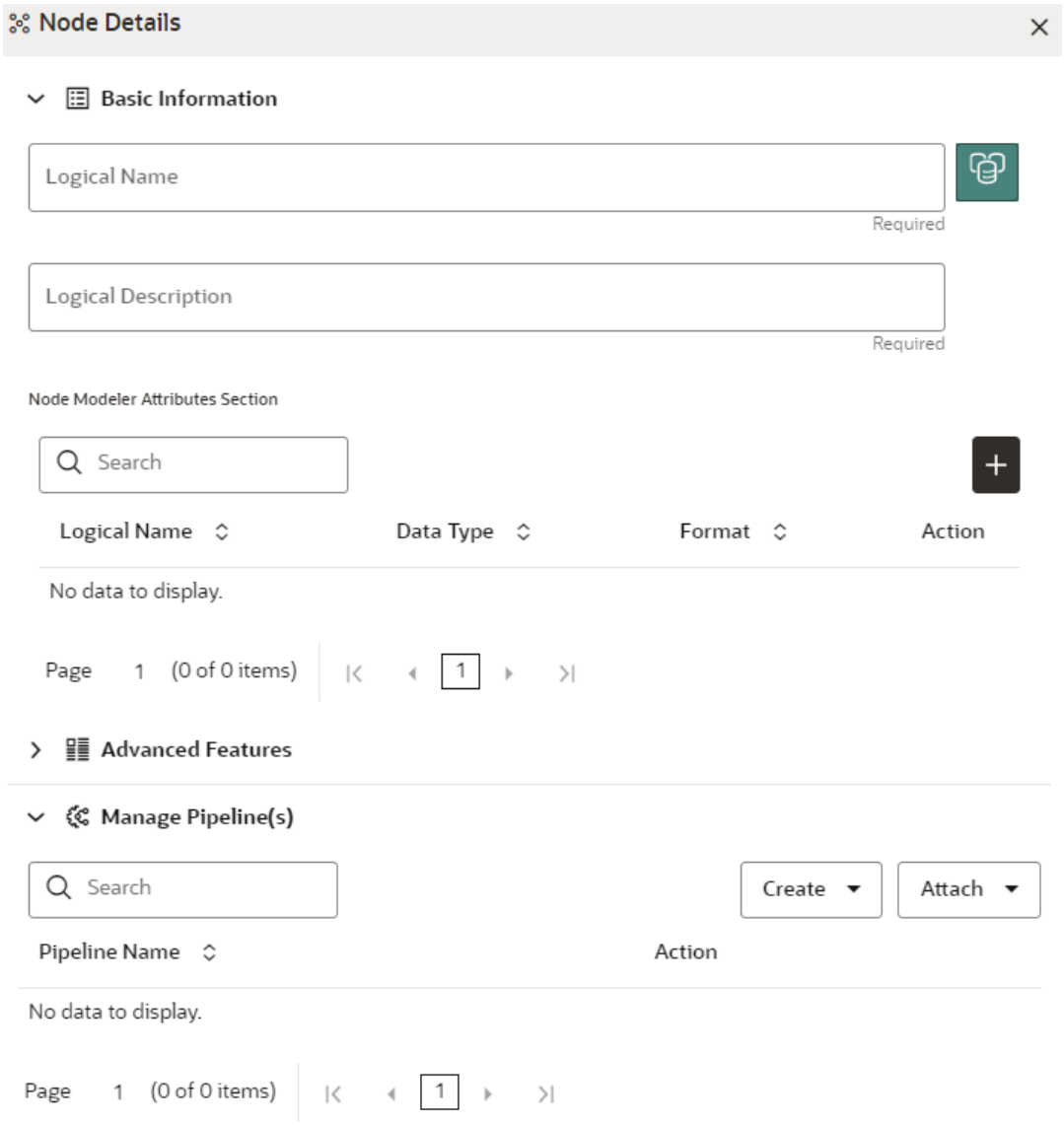
1. In the Graph Model page, click Create Node.

   OR

Select the graph for which you want to create a Node, click on the Action button, and select the **Edit** option. In the Graph Model page, click **Create Node**.

The **Node Details** screen is displayed.

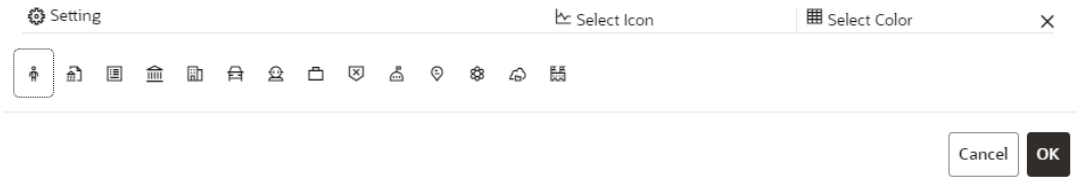
Figure 6-96 Node Details



2. In the **Logical Name** field, enter a name for the Node or click **Setting** .

The Setting pop-up window is displayed.

Figure 6-97 Setting



3. Select the required representation from the above Node icon.

   For example: Person, Institution, Account etc.

   Hover over the icon to view the definition of the node.

4. Based on your selection, the related Nodes are displayed. Select the Node, and you can choose the required color for the Node using the **Select Color** option.

5. The selected Node is displayed on the graph. To add attributes to the Node, click **Add** .

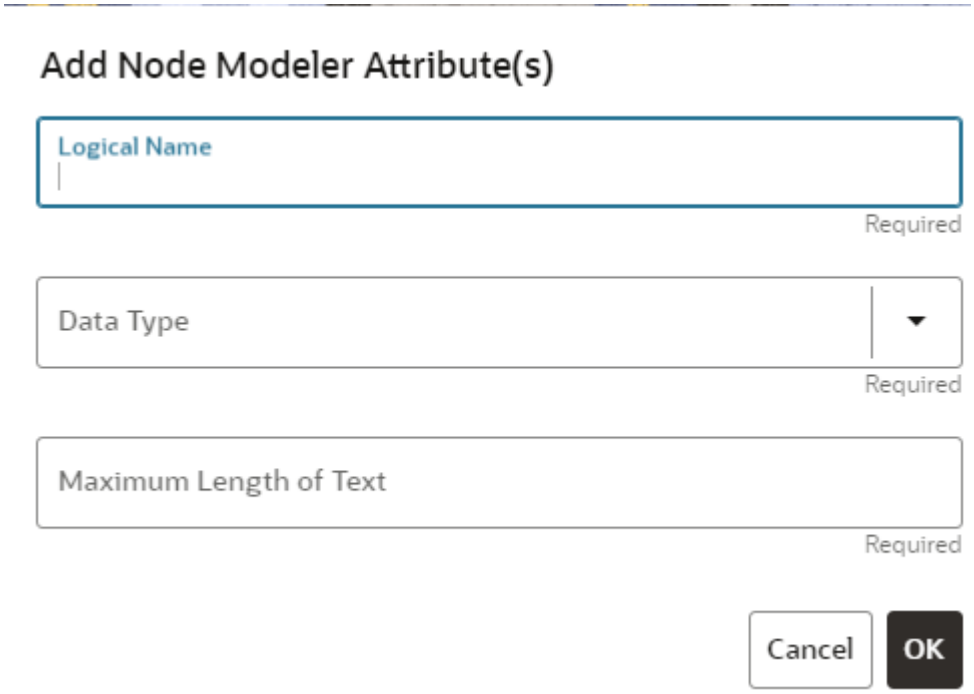
   Modeler attributes are the properties of Node/Edge, and these are mandatory inputs.

   For Example:

   Customer Node will have properties like Customer ID, Name, Age, Gender, etc.

   The **Add Node Modeler Attribute(s)**screen is displayed.

Figure 6-98 Add Node Modeler Attribute(s) screen



6. In the **Logical Name** field, enter the logical name for the modeler attribute.

   This field is mandatory.

7. Select the data type based on your requirements in the **Data Type** field drop-down.

   The available data types are Text, Number, and Date. This field is mandatory.

   Based on the data type selection, the following input field is displayed.

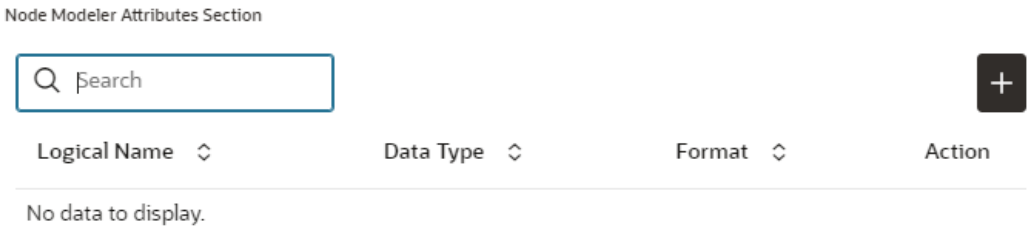
   For example, the **Maximum Length** of Text field is displayed if the data type is selected as Text.

> **Note:**
>
> Ensure the Maximum Length attribute is set to at least the field's length from which the data is to be populated.

8. Enter the data and click **OK**.

   The Node attributes are displayed on the Node Details page.

Figure 6-99 Search



You can add any number of modeler attributes based on your requirements by clicking on the Add icon.

9. Under the **Advanced Features** group, select the modeler attribute from the Column Identifier drop-down. All the created modeler attributes are displayed in this list.

Figure 6-100 Advanced Features

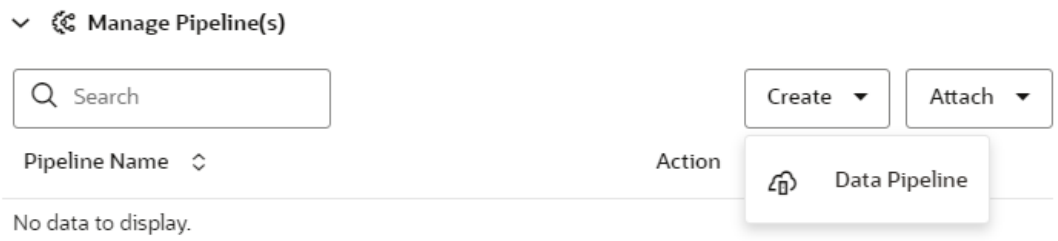
Under the **Manage Pipeline(s)** group, you can either create a new data pipeline or attach the existing pipeline.

The application uses data pipelines to prepare filtered data which can be used to create and run scenarios. Data pipelines prepare data by selecting and joining data sources to create virtual data tables, adding derived attributes to data, running derivations on the data to determine the risk associated with the entity, and so on.

> **Note:**
>
> The newly created data pipeline is not automatically added to the Node. You must attach the pipeline using the **Attach** drop-down option.

The selected data pipeline is displayed under the Manage Pipeline(s).

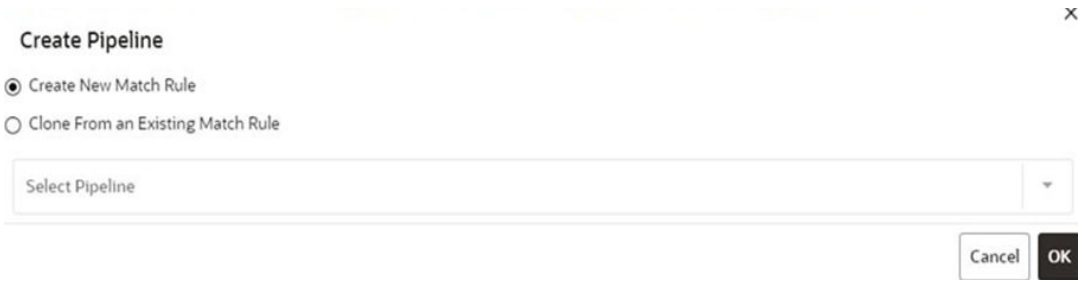
Figure 6-101 Manage Pipeline(s)



10. Click **Create** drop-down and select **Data Pipeline**.

    The **Create Pipeline** window is displayed.

Figure 6-102 Create Pipeline



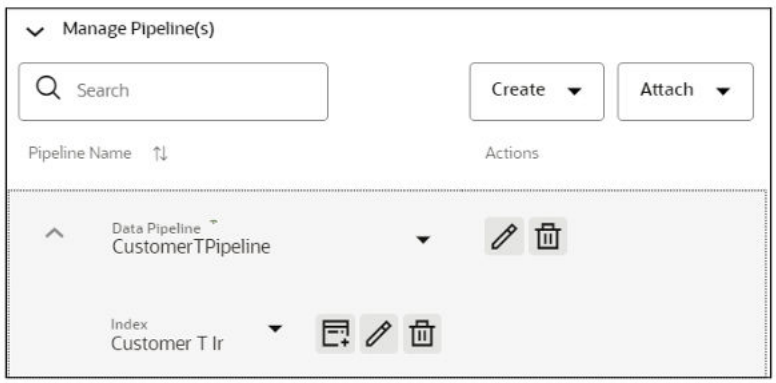
11. Select the **Create New Data Pipeline** option and then click **Ok**.

    The Pipeline Designer page is displayed.

    - To create a data pipeline, see the Managing Data Pipeline section.
      OR
    - To clone from an existing data pipeline, see the Cloning from an Existing Data Pipeline section.

> **Note:**
>
> The Persist of the Data pipeline of the corresponding node should be defined with the prescript. See the Prescript Condition section for more details.

Figure 6-103 Create an Index

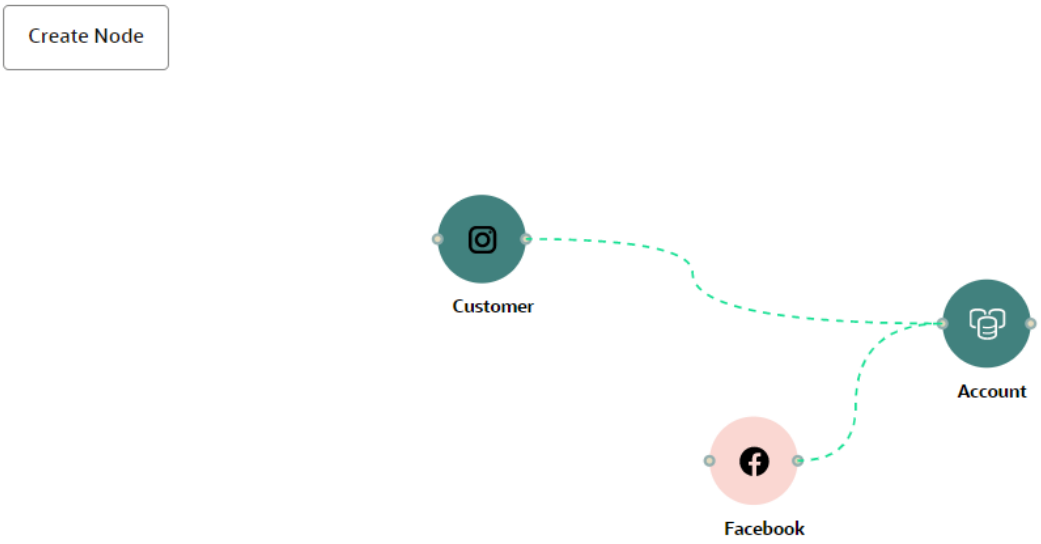


- To define an index for the data pipeline, see the Defining an Index section.

12. By clicking on the **Attach** drop-down, you can add any number of data pipelines to the Node-based on your requirements.
13. Click **OK**.

    The Node is created with pipeline Facebook and an account attached to it as shown in the below example.

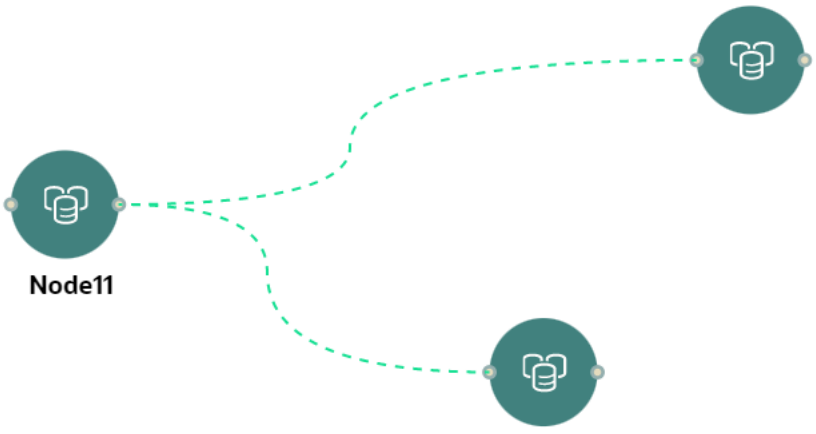
Figure 6-104 Graph with Node and Edge



6.6.1.3 Create an Edge

An Edge can be used to connect two different nodes or connect to itself. When compared to nodes, edges also capture the basic information like name, description, attributes to an edge, and possible pipelines that can be attached to an edge.

Figure 6-105 Edge

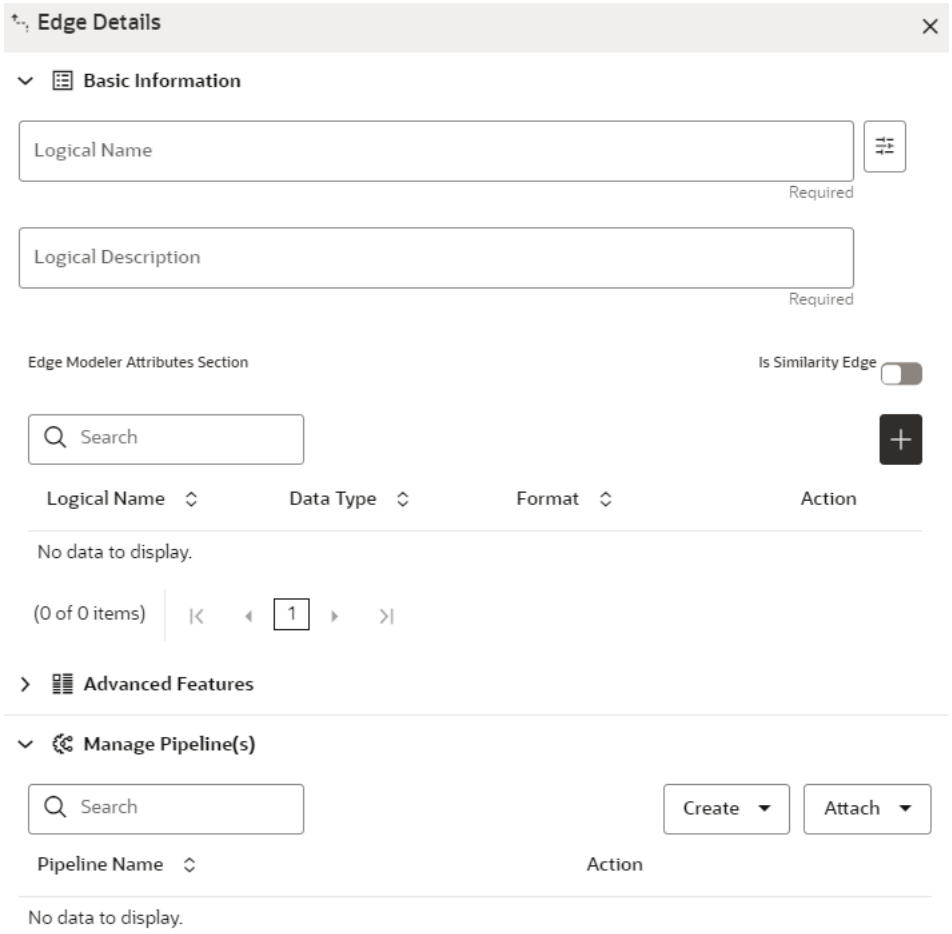


To create an edge, perform these steps:

1. Hover on the Node to create a relationship between them.

   The **Edge Details**screen is displayed.

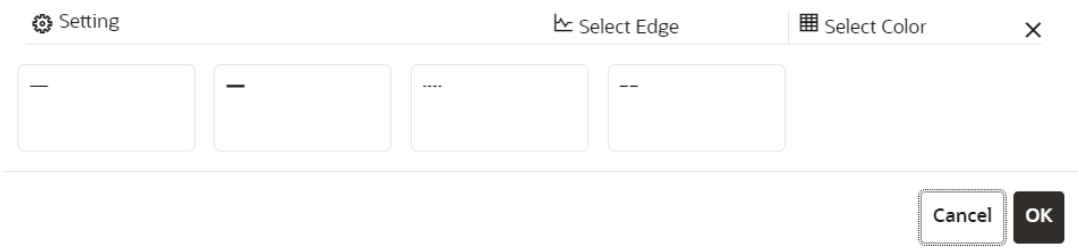
Figure 6-106 Edge Details



2. In the **Logical Name** field, enter the Node's name or click the **Setting** icon.

   The **Setting** pop-up window is displayed.

Figure 6-107 Setting



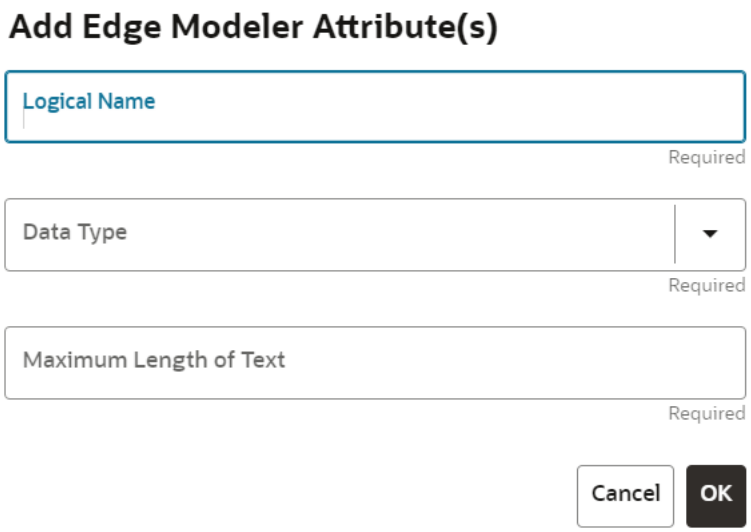
3. Select the required edge from the options displayed and choose the required color for the edge using the **Select Color** option.

4. Click **OK**.

   The selected edge is displayed on the graph pipeline.

5. To add attributes to the edge, click the **Add** icon.

Figure 6-108 Add Edge Modeler Attribute(s) screen



   The **Add Edge Modeler Attribute(s)** screen is displayed.

6. In the **Logical Name** field, enter the logical name for the edge attribute.

   This field is mandatory.

7. In the **Data Type** drop-down, select the data type based on your requirements.

   The available data types are Text, Number, and Date. This field is mandatory.

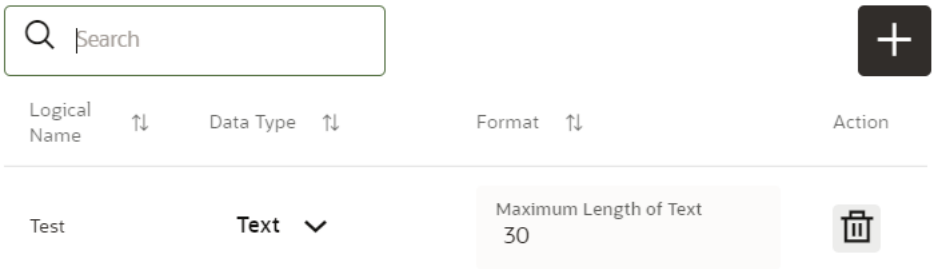
   Based on the data type selection, the following input field is displayed.

   For example, the **Maximum Length of Text** field is displayed if the data type is selected as **Text**.

8. Enter the data and click **Ok**.

   The edge attribute is displayed on the **Edge Details** page.

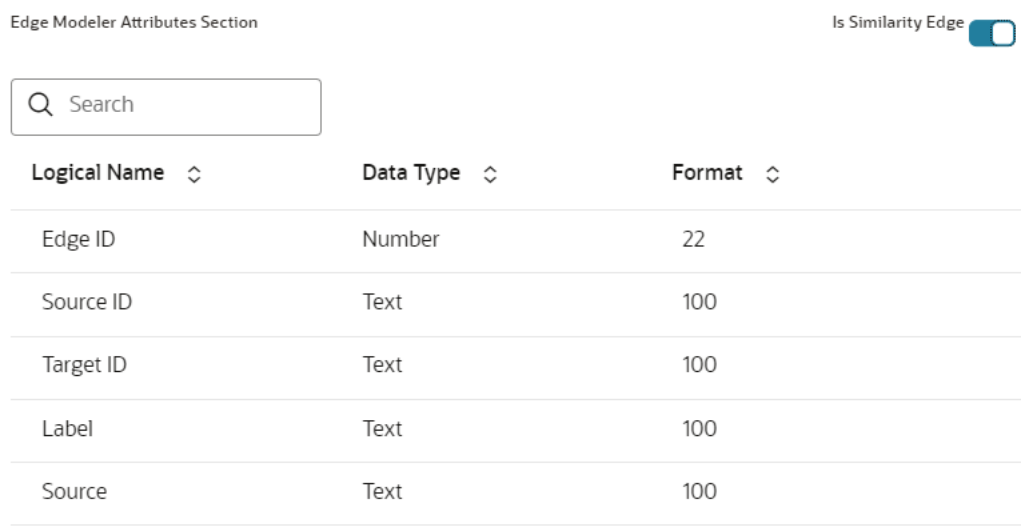
Figure 6-109 Search



By clicking on the **Add** icon, you can add any number of edge attributes based on your requirements.

OR

Figure 6-110 Edge Details



Click the **Is Similarity Edge** toggle button the **Edge Modeler Attributes** section to add the pre-defined set of attributes available in the edge.

Similarity Edgesin the graph are created by Match Rules. If the edge is defined using Match Rule (under Manage Pipeline, by selecting Match Rule from drop down), the toggle button should be enabled to make it a Similarity Edge.

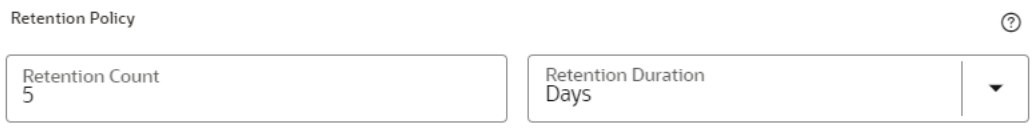
9. Under the **Advanced Features** group, you can specify the Edge Retention Policy.

   For example, An Edge Retention Policy of 7 Days specifies that the current edge would be retained in the physicalized Graph for a period of 7 Days.

10. Enter the retention count and select the retention duration from the drop-down.

    The available retention durations are Days, Weeks, and Months.

Figure 6-111 Retention Durations tab



11. Enable the **Pluggable Attribute**, then the corresponding drop-down will contain the list of attributes of the Edge.

    Pluggable edges are those edges whose data can be directly plugged in to the graph. In Financial Crime and Compliance domain, Transaction is a pluggable edge, which is generally large and is difficult to find out difference/ delta between two subsequent loads. Hence, it's data is provided for a range of dates which can be directly plugged in to the PGX memory while refreshing the graph.

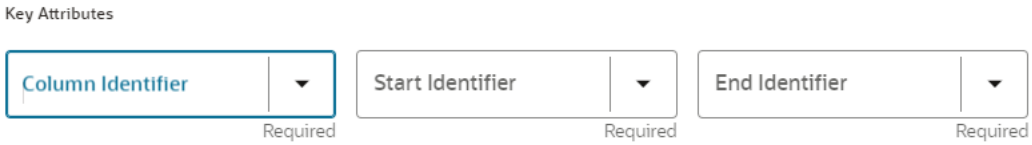
Any pluggable edge should have an identifier attribute, which is of “Date” data type. Figure: Pluggable Attribute.

Figure 6-112 Pluggable Attribute

12. Under the **Key Attributes** field, select the Column, Start, and End identifier from the respective Identifier drop-down. Without these attributes, an edge cannot be created.

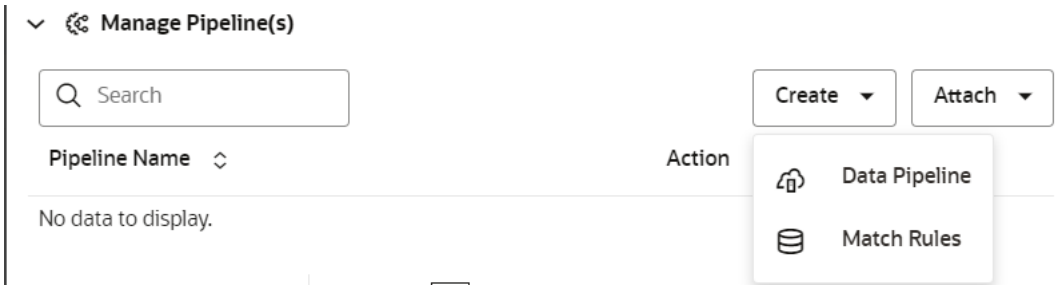
    All the created edge attributes are displayed under this drop-down.

Figure 6-113 Key Attributes



13. Under the **Manage Pipeline(s)** group, you can either create a new data pipeline if the edge is coming from source data and Match Rules, if a similarity edge is to be created between nodes or attach the existing pipeline and Match Rules.

Figure 6-114 Manage Pipeline(s) for Edge



14. Click **Create** drop-down and then select the Data Pipeline.

    The **Create Pipeline** window is displayed.

Figure 6-115 Create Pipeline

15. Select the **Create New Data Pipeline** option and then click **OK**.

    The Pipeline Designer page is displayed.

    - To create a data pipeline, see the Creating a New Data Pipeline section.
      OR
    - To clone from an existing data pipeline, see the Cloning from an Existing Data Pipeline section.

> **Note:**
>
> Define persist of the Data pipeline of the corresponding edge with the prescript. See the Prescript Condition section for more details.

16. Click **Create** drop-down and then select the **Match Rules** from the Manage Pipeline(s) group.

    The **Create Pipeline** window is displayed.

Figure 6-116 Create Pipeline

17. Select the **Create New Match Rule** option and then click **Ok**.

    The Create Match Ruleset window is displayed

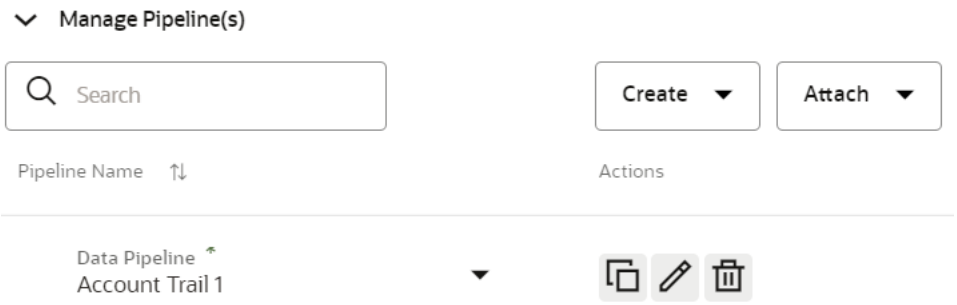
    - To create a Match Rule, see the Creating Match Ruleset section.
      OR
    - To clone from an existing match rule, see the Cloning Ruleset (Match) section.

> **Note:**
>
> The newly created data pipeline and match rule are not automatically added to the edge. You must attach using the Attach drop-down option.

By clicking on the **Attach** drop-down, you can add any number of data pipelines/match rules to the Node-based on your requirements.

The selected data pipeline/match rule is displayed.

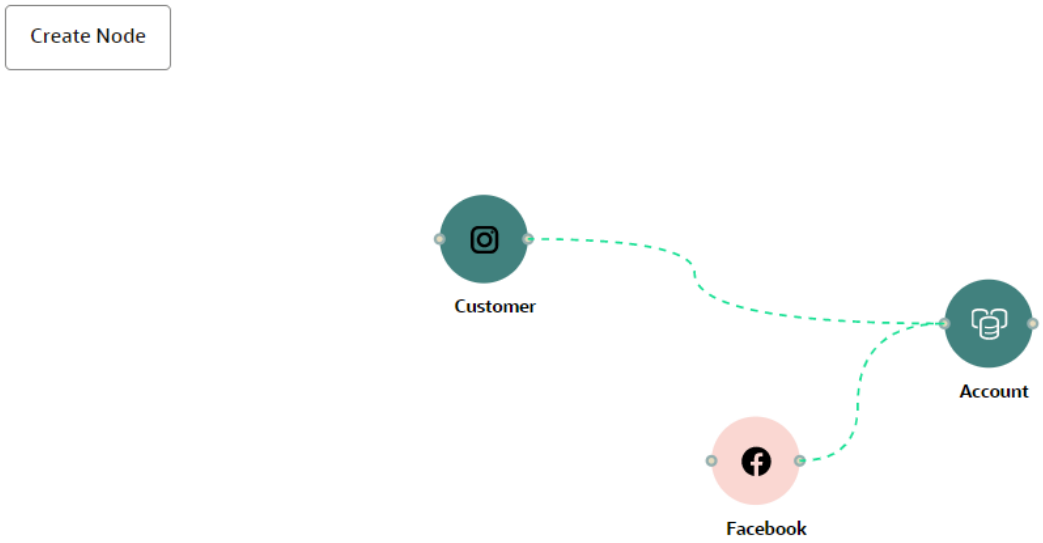
Figure 6-117 Create Pipeline



18. Click **OK**.

    An Edge is created between the Customer, Facebook, and Account as shown in the below example.

Figure 6-118 Graph page



6.6.1.4 Zoom In and Zoom Out

You can zoom in or zoom out of the graph pipeline and view the data pipelines using the diagram present in the bottom right corner of the page.

6.6.1.5 Edit a Graph Pipeline

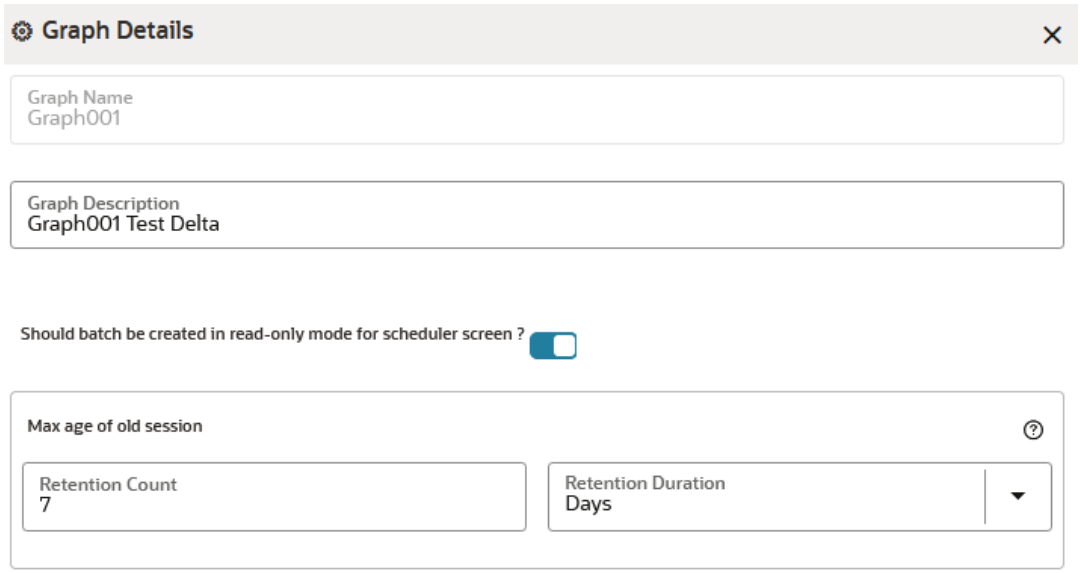
To edit a graph pipeline, perform these steps.

1. In the **Graph Summary**page, select the graph pipeline you want to edit, click the Action icon, and select **Edit**.
2. In the **Graph Name**field, modify the graph pipeline name.

   The name must be unique to a particular workspace.
3. Click the **Setting** icon.
4. In the **Graph Description**field, enter the description for the graph pipeline and click **Save**.

The following page is displayed.

Figure 6-119 Graph page



The Maximum Age of Old session for a graph is 7 days by default. The Maximum age of old session of 7 days specifies that graph would be retained for a period of 7 days. You can modify the description and Max Age of Old session by clicking on the **Setting** icon. If you want the batch to be created in read-only mode for scheduler screen, enable the option.

5. Click **Save**.

   The user session of the Graph Pipeline will get refreshed after the set timeline.

6.6.1.6 Delete a Graph Pipeline

To delete a graph pipeline:

- On the **Graph Summary**page, select the graph pipeline you want to delete, click the **Action**icon, and select **Delete**.

  To delete multiple graph pipelines, select the graph pipelines which you want to delete and click **Delete** icon in the Header.

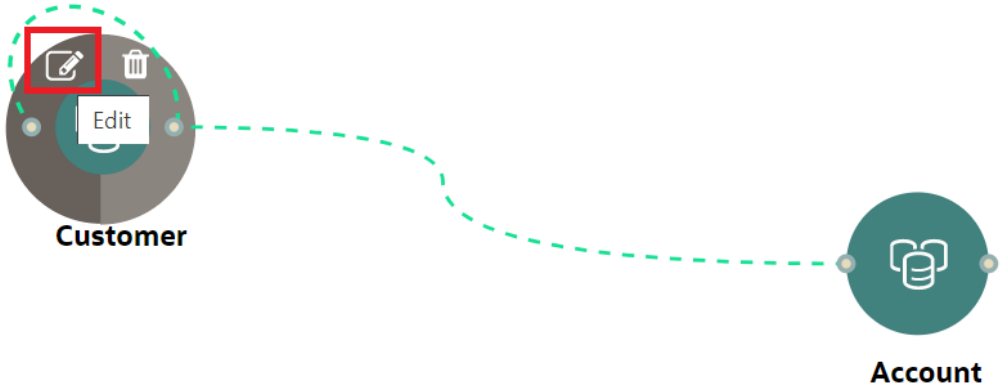
6.6.1.7 Edit a Node

To edit a node, perform these steps.

1. On the **Graph Summary** page, select the graph pipeline you want to edit, click the **Action** icon, and select **Edit**.
2. Hover over the graph pipeline and click **Edit** icon.

   The **Node Details**page is displayed.

Figure 6-120 Node Details page



3. Perform the steps from 3 to 12 in the Creating a Node section.

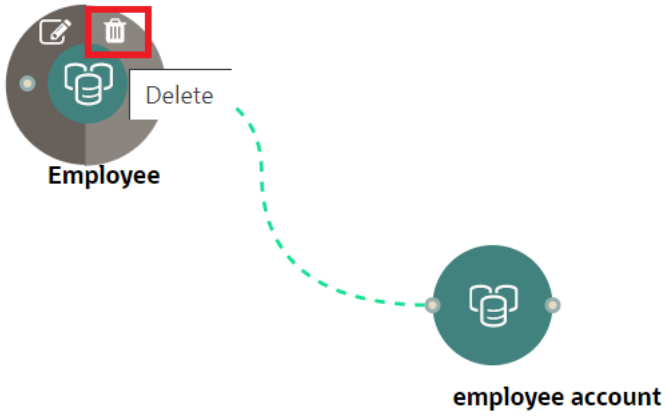
> **Note:**
>
> If you delete any attribute in the Node Modeler Attribute section of the Node Details page, you must delete the same attribute in the match rule details.
> For example:
>
> If the edge has two match rules and one is attached, and another match rule is not attached to the edge details. You need to remove attributes in both math rules for this case. To remove the deleted attribute from the Node Details page in the unattached match rule, perform the following:
>
> a. Navigate to **Edge Details > Manage Pipeline(s) > Attach**. Attach the new match rule, which is not attached to the node details.
> b. Click **Edit** and delete the attribute in the Ruleset Details, which is removed from the Node Details page.
> c. Click **Save**. The attribute is removed.
> d. Click **Delete** in Manage Pipeline(s) of the Edge Details page to remove the attached match rule.
> e. After saving in the Edge Details page, navigate to the Node Details page and delete the attribute again.
> f. Click **OK**. The attribute is removed from the node.

6.6.1.8 Delete a Node

To delete a node:

1. On the **Graph Summary** page, select the graph pipeline you want to edit, click the **Action**icon, and select **Edit**.
2. Hover over the graph pipeline and click **Delete**icon.

Figure 6-121 Delete Node



A confirmation dialog is displayed.

3. Click **Delete**.

   The Node is deleted.

6.6.2 Schedule Details

> **Note:**
>
> Currently, when you execute a Graph pipeline for the first time and navigates to the "Monitor Batch" screen, two "run ids" are generated for that batch.

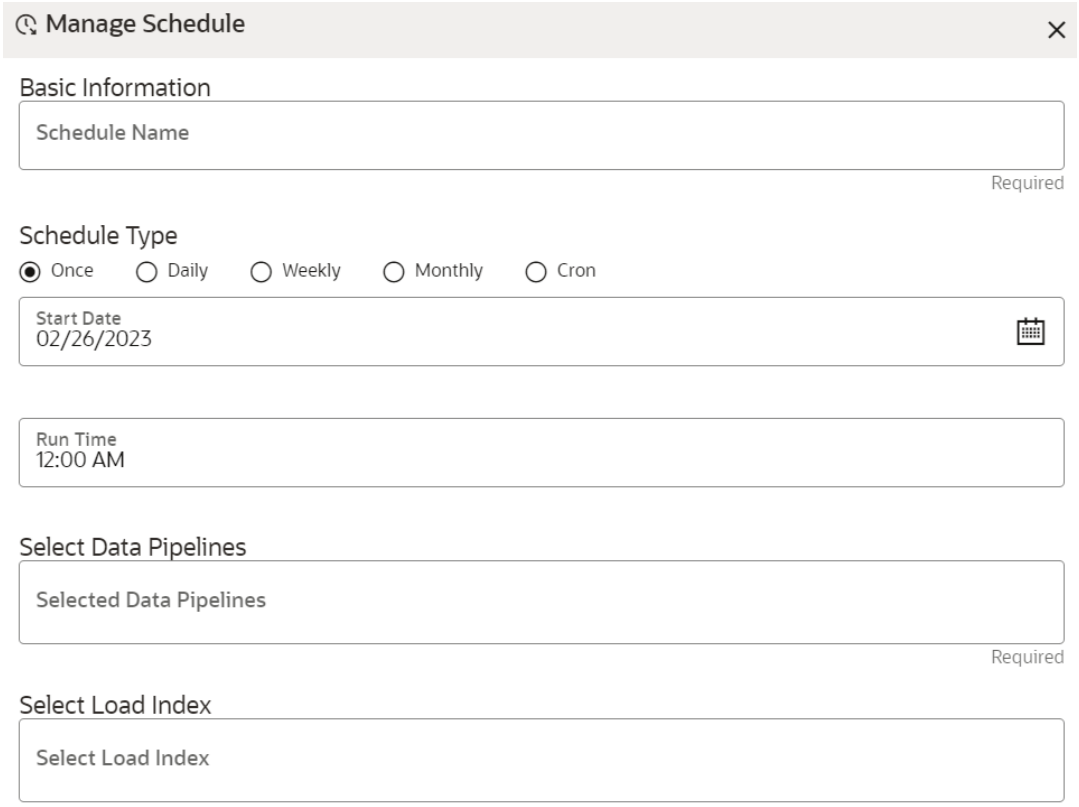
To create a Graph Refresh Schedule:

1. Click **Launch Workspace** next to the corresponding Workspace to display the **Dashboard** window with application configuration and model creation menu.
2. Click the **Graphs** icon.

   This displays the graph pipeline in a table.
3. Click on the graph for which you want to create refresh schedule.
   - The Graph Model page is displayed.
4. Navigate to **Schedule Details** page and click **Add Schedule**.

   The **Manage Schedule** screen is displayed.

Figure 6-122 Manage Schedule screen



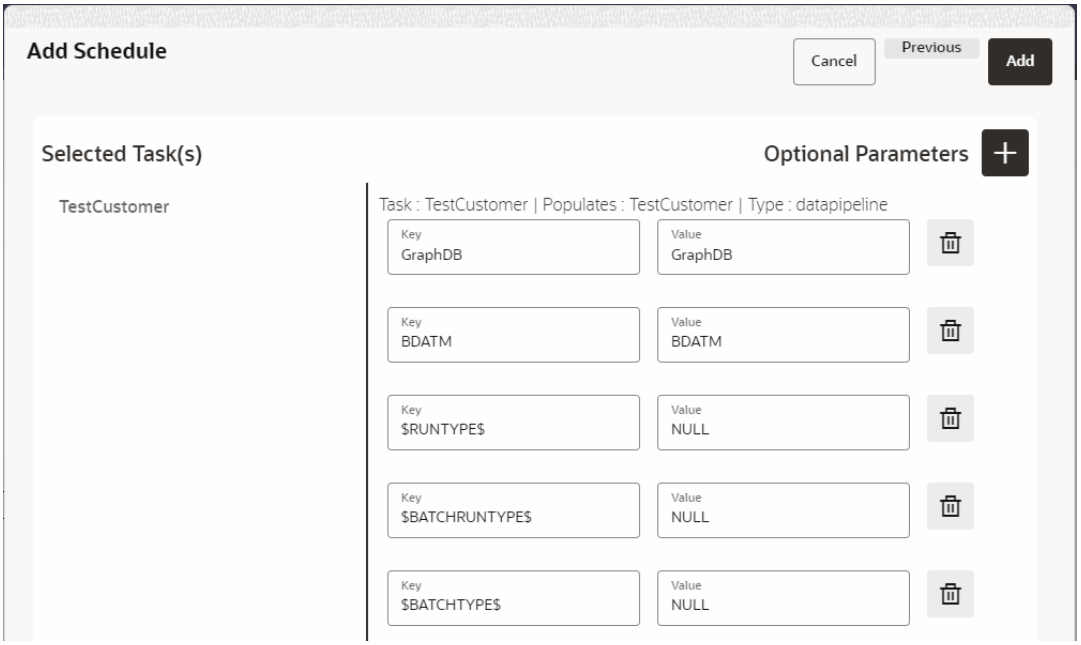
5. In the **Schedule Name**, enter the desired name.
6. Select the **Schedule Type** from the options.

   Based on the schedule type, the following input field is displayed.

> **Note:**
>
> For more details and expressions of schedule type, click the **help** icon.

7. Click inside the **Select Task** field and select the tasks. You can select multiple tasks.
8. Click **Next** and add the values for the selected task, and the values are mandatory fields.

Figure 6-123 Add Optional Parameters



9. Click the **+** icon to add optional parameters for the selected task.

   For the Pre-configured Data pipeline, respective Key fields will be displayed. You can configure the respective key-value in the Value field as required and specified in the following table:

Table 6-24 Optional Parameters for Data Pipeline

Name	Description	Default Value
$GRAPH_SCHEMA	Oracle wallet alias of graph schema. Navigate to <OFS_COMPLIANCE_STUDIO >/wallet/ tnsnames.ora and check the Graph Schema alias name. For example, graph_schema_alias	
$BD_SCHEMA	Oracle wallet alias of BD schema. Navigate to <OFS_COMPLIANCE_STUDIO >/wallet/ tnsnames.ora and check the BD Schema alias name. For example, bd_schema_alias	
$RUNTYPE$	Indicates Run type. It can be either a production batch or a test batch. If it is a production batch, set it as PROD and TEST for the test batch.	PROD

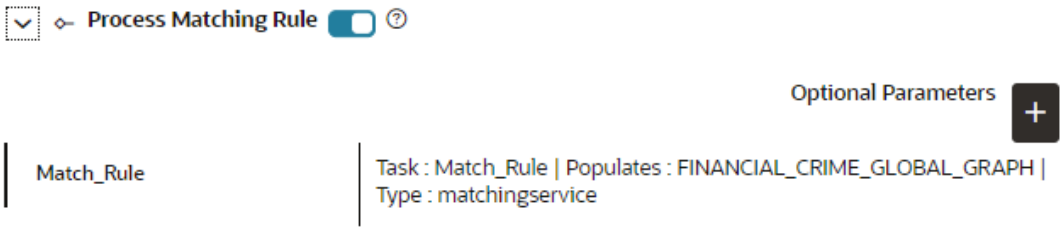
Table 6-24 (Cont.) Optional Parameters for Data Pipeline

Name	Description	Default Value
$batchRunType$	Indicates Batch Run type. It can be either run or re-run. If it needs to execute once, set it as RUN and RE-RUN for re-executing after failure.	RUN
$BATCHTYPE$	Indicates the batch type. It should be DATA for the data pipeline task.	DATA
$JOBNAME$	Name of the data pipeline to be executed.	

10. To add Match Rules, enable the Process Matching Rule option and click + icon to add optional parameters.

    For more information on Match Rule, see the Creating Match Ruleset section.

Figure 6-124 Process Matching Rule



For the Pre-configured Match Rules, respective Key fields will be displayed. You can configure the respective key-value in the Value field as required and specified in the following table:

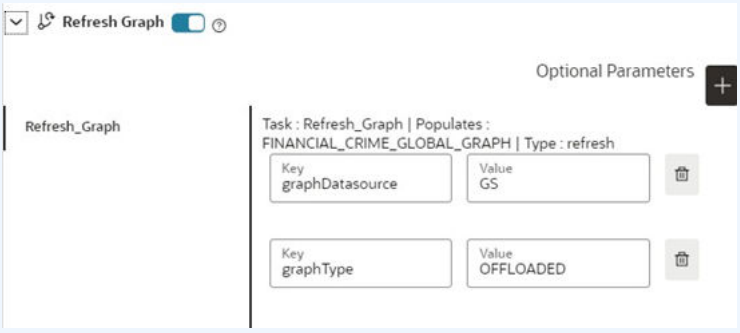
Table 6-25 Optional Parameter for Match Rules

Name	Description	Default Value
datasource	Oracle wallet alias of graph schema. Navigate to <OFS_COMPLIANCE_STUDIO >/wallet/ tnsnames.ora and check the Graph Schema alias name. For example, graph_schema_alias	-
loadType	Execution Load Type. Available types are full and delta. In this release, only the full mode is supported.	full
processingGroupName	Logical processing group name for the batch. You should set it as MAN.	MAN

Table 6-25 (Cont.) Optional Parameter for Match Rules

Name	Description	Default Value
basedOnPipelineId	Match rules can be executed based on pipeline Ids or ruleset Ids. The value should be true while running from the Graph.	true
runSkey	A numeric value indicates the runskey of the execution.	100
runType	Indicates matching job that can be executed as part of graph or Entity Resolution. It should be set as Graph.	graph

11. Enable the **Refresh Graph** option and click **Add** to refresh the Graph to add optional parameters.

> **Note:**
>
> The Key and Value parameters are not required.
>
> **Figure 6-125 Refresh Graph**
>
> 
>

> **Note:**
>
> If you do not enable the Refresh Graph option, the pipeline will add nodes and edges to the Graph Schema but does not load into PGX.

> **Note:**
>
> The user is now able to execute/schedule a batch with SYSDATE-X days functionality.

6.6.3 Audit History

At any time, you can audit the Graphs from the Audit History page. This page provides the events on the Graph. This shows the information such as, when the Graph was created, who created the Graph and its status, and so on.

The sequence of actions performed on the Graph is listed in the table view and the timeline view (graphical representation).

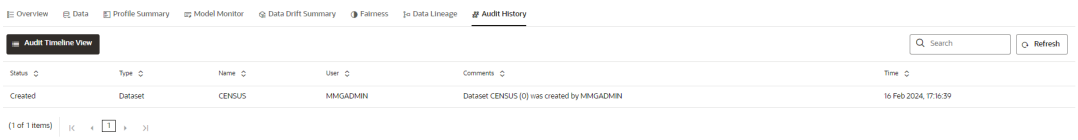
To view the audit history of the Graph, follow these steps:

1. Click **Launch Workspace** next to corresponding Workspace to Launch Workspace to display the **MMG Dashboard** window with application configuration and model creation menu.
2. Click the **Graphs**icon.

   This displays the graph pipeline in a table.
3. Click on the graph for which you want to view the audit history.

   The **Graph Model**page is displayed.
4. Navigate to **Audit History** page to view the details.

Figure 6-126 Audit History



Click for **Audit Timeline View** for Timeline View and **Audit Table View** to switch to the regular view. Click to **Refresh**to refresh the Audit History.

Figure 6-127 Audit History window Timeline View with Horizontal Time Axis

6.6.4 Execution History

To view the Graph Execution History:

1. Click **Launch Workspace** next to the corresponding Workspace to display the **Dashboard** window with application configuration and model creation menu.
2. Click the **Graphs** icon.

   This displays the graph pipeline in a table.

3. Click on the graph for which you want to view the execution history.

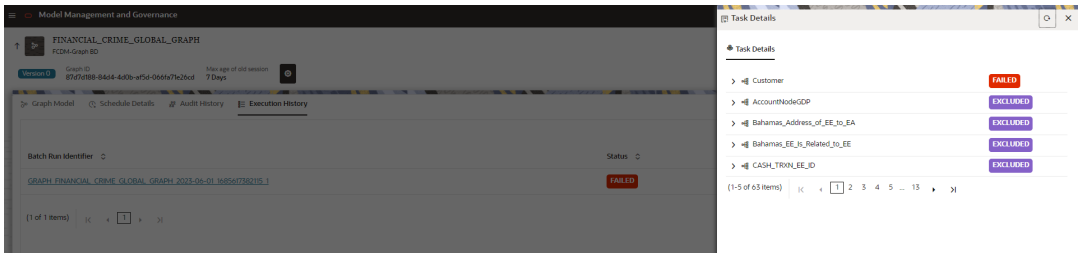
   The **Graph Model** page is displayed.

4. Navigate to **Execution History** page to view the details.

   This window displays the details such as Batch Run Identifier, Status, Start and end time of the execution. Click on the Batch Run Identifier to view the individual task details.

> **Note:**
>
> A plus icon has been added under Execution History of Details screen to support populate Sandbox from within the screen.

Figure 6-128 Execution History page



6.6.5 Analyses

You can analyse, search, add, and delete a notebook from this page.

> **Note:**
>
> In the current release, features such as Edit, Clone, and Credential version are not supported.

To analyse the Graph:

1. Click **Launch Workspace** next to the corresponding Workspace to display the **Dashboard** window with application configuration and model creation menu.
2. In the Mega menu, click **Modeling > Graphs**.

   This displays the graph pipeline in a table.

3. Click on the graph for which you want to view the execution history.

   The **Graph Model** page is displayed.

4. Navigate to **Analyses** page to view the details.

7
Orchestration

Orchestration covers the following:

- **Scheduler Dashboard**: Get an overview of the scheduled tasks and processes.
- **Define Batch**: Manage and configure batch definitions.
- **Define Tasks**: Create tasks, configure parameters, and set execution dependencies within a batch process.
- **Schedule Batch**: Set execution schedules for your batch processes.
- **Monitor Batch**: Track and monitor batch process executions.

7.1 Scheduler Service

The Scheduler Service automates the running of models in the Compliance Studio application.

The Scheduler Service contains a graphical user interface and a single control point for defining and monitoring background executions.

The Scheduler Service is a service in the Infrastructure system that automates behind-the-scenes work that is necessary to sustain various enterprise applications and functionalities. This automation helps the applications to control unattended background jobs.

The Scheduler Service contains a graphical user interface and a single point of control for the definition and monitoring of background executions.

Following are the concepts or terminologies in the Scheduler Service:

- **Batch**: Date and time-based execution of the background tasks based on a defined period during which the resources were available for batch processing.
- **Job**: A batch job is a piece of a program meant to meet specific and business-critical functions. The program is a RESTful API used in a batch.
- **Job Dependency**: When the batch job is submitted, it is moved to the job queue until the system is ready to process. The system process the job based on chronological order or priority in case if more jobs are required to be executed in the job queue.
- **Schedule**: Batch jobs are used to automate the tasks that require to be performed on a regular basis but don’t necessarily need to occur during the day or have an employee interacted with the system are batch schedule. Jobs that happen on a regular basis are incorporated into batch schedules.

7.1.1 API Endpoints for Export/Import Scheduler Objects

Header Parameters for each API:

- ofs_tenant_id - Tenant ID of the Application
- ofs_service_ID - Service ID of the Application
- ofs_workspace_ID - Workspace ID of the Application

- ofs_remote_user - Used ID of the User. This parameter should be mapped to proper function.
- Content-Type: application/JSON

1. Batch Export API:
   HTTP Method - GET

   URL - /v1/rest/batch/:code
2. Batch Import API:
   HTTP Method - POST

   URL - /v1/rest/batch

   Request Body:

   ```
   {
   "objectcode": "<<BATCH_NAME>>",
   "objecttype": "BATCH",
   "objectsubtype": "",
   "overwrite": "Y",
   "objectdefinition":""
   }
   ```
3. Batch Group Export API:
   HTTP Method - GET

   URL - v1/rest/batchgroup/:code
4. Batch Group Import API
   HTTP Method:

   POST URL- /v1/rest/BatchGroup

   Request Body:

   ```
   {
   "objectcode": "<<BATCHGROUP_NAME>>",
   "objecttype": "BATCHGROUP",
   "objectsubtype": "",
   "overwrite": "Y",
   "objectdefinition":""
   }
   ```
5. Schedule Export API:
   HTTP Method- GET

   URL- v1/rest/schedule/:code
6. Schedule Import API HTTP Method - POST URL - /v1/rest/schedule
   Request Body:

   ```
   {
   "objectcode": "<<SCHEDULE_ID>>",
   "objecttype": "SCHEDULE",
   "objectsubtype": "",
   ```

```
"overwrite": "Y",
"objectdefinition":""
}
```

> **Note:**
>
> 1. OverWrite flag can be **Y** or **N**
> 2. Objectdefinition will be obtained in export API same need to be paste during import in request body
> 3. In Schedule Export API code it is Schedule ID.
> 4. During import of schedule dependent batch or if BatchGroup is not migrated, the user have to use batch and BatchGroup import API to migrate updated definition in the Target environment.

7.1.2 Scheduler Service Dashboard

1. Click **Launch Workspace** next to corresponding Workspace to display the **Dashboard** window with application configuration and model creation menu.
2. In the LHS menu, click **Scheduler Service**.

   This displays the executed runs in a table.

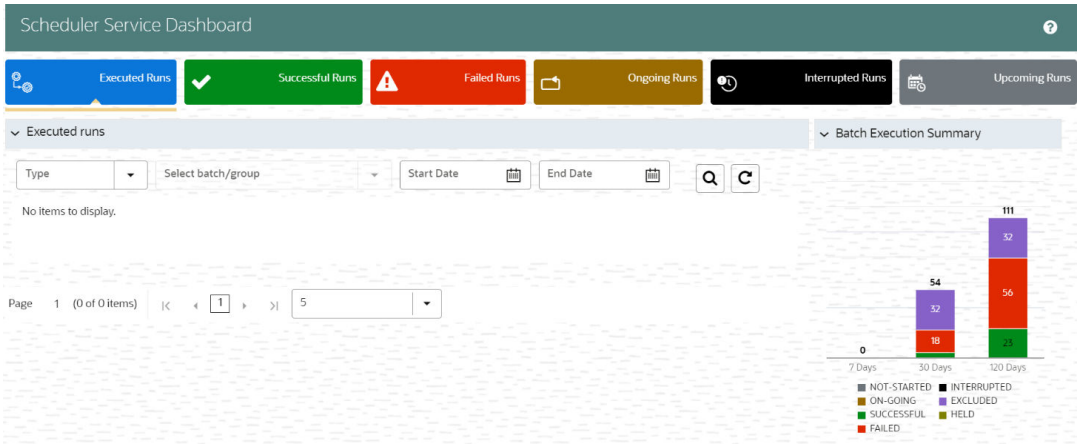
   To view the demonstration of the Dashboard window, see the Scheduler Service Introduction video.

   In the Scheduler Service window, you can view the following details:

   - The Executed Runs, Successful Runs, Failed Runs, Ongoing Runs, Interrupted Runs, and Upcoming Runs tabs. You can click the tabs to view the details of the Batches based on their status. For example, click **Ongoing Runs** to view the details of the batches that are currently running.
   - The Batches that were executed within the last 7 or 30 days contain details such as Batch Name, Batch Run ID, and Run Time. Click **30 days** to view the batches that were executed within the last 30 days.
   - The **Batch Execution Summary** pane displays the count of total batches executed that were executed within the last 7 days, 30 days, and 120 days. Additionally, you can see the separate count of successful batches, failed batches, interrupted batches, on-going batches, and the batches which are yet to start, by hovering your mouse over the batches.

   You can filter the executed runs based on Batch type, Batch/ group, start date and end date by selecting the corresponding options.

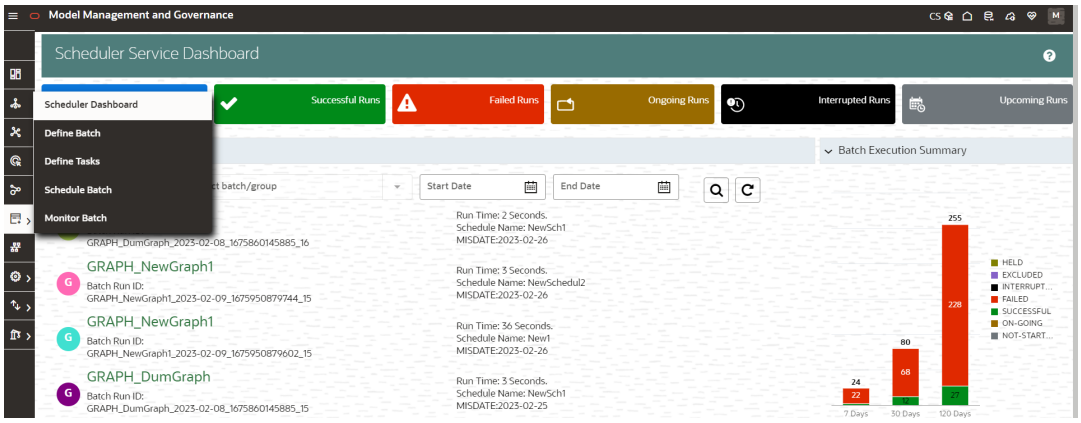
Figure 7-1 Scheduler Service Dashboard



On the **Scheduler Service** window, you can see:

- Menu Navigation: Shows the menu options available in Scheduler Service.
- Quick Actions: Quick actions can be performed by using the buttons in this section.
- User Details: View the logged in user's details, help, and set your profile preferences.

Figure 7-2 Header Details



On the left side of the window, click the **Scheduler Service** icon to see available options in the Scheduler Service.

7.1.3 Configure Batch and Batch Groups

This section describes about Configure Batch and Batch Groups.

7.1.3.1 Create a Batch

The **Define Batch** window displays the details of an existing batches such as Batch ID, Batch Name, Batch Description, Last Modified By and Last Modified Date. You can create, edit, copy, and delete batches. You can also create a Batch Group, a list of batches that you have selected and grouped, and execute it.

To create a new batch and schedule and monitor the batch that you create:

1. Click **Define Batch** from the Header in the Dashboard window. After selecting the batch, click corresponding to the batch to proceed to create or edit tasks.
2. In the Define Batch window, click **Add**.

   The **Create a New Batch** window is displayed.
3. Specify the details as described in the following table.

Table 7-1 Fields in the Create a New Batch window and their Description

Field	Description
Batch Name	The Batch Name is generated based on the values provided by you. **NOTE**: • The Batch Name should be unique across the Information Domain. • The Batch Name must be alphanumeric and should not start with a number. • The Batch Name should not exceed 60 characters in length. • The Batch Name should not contain any special characters except “_”.
Batch Description	Enter a description for the Batch based on the Batch Name. **NOTE**: The Batch description should be alphanumeric. The allowed special characters are: and <blank space>, along with spaces and alpha-numeric. It should not exceed 200 characters in length.
Service URL Name/ Service URL Notify on Mail	• Select the Service URL name from the drop-down list, if it is available. The Service URL is displayed in the Service URL field. • To add a new service URL, enter a name to identify it in the Service URL Name field and enter the proper URL in the Service URL field. You can give partial URL here and the remaining URL in the Task Service URL. Choose this option to notify by an e-mail the batch execution status depending on the property selected. E-mail for the email notification is fetched from the IDCS configuration for the logged-in user. **NOTE**: E-mail is a mandatory field in IDCS. The BATCH_NOTIFY_FUNCT function has to be enabled to access this feature. If it is not available for the logged-in user, this field is grayed out. – Every Time: An e-mail is triggered irrespective of the batch status. – Never: No e-mail will be triggered. – On Error only: An e-mail is triggered only when the batch execution has failed/Error. – On Interrupt only: An e-mail is triggered if the batch is successfully interrupted. **NOTE**: **Never** is the default option.

4. Click **Save**.

   The new Batch is created and displayed in the **Define Batch**window.

7.1.3.2 Manage Batch

You can manage a batch by editing, copying, or delete them as required.

- **Edit**: Use the **Edit Batch** icon to edit the Batch Description, Service URL Name, Service URL. You can also add a new batch parameter. You cannot edit seeded batches.
- **Copy**: Use the **Copy Batch** option to copy a Batch that you want to clone or create instances in the system from the **Define Batch** window.
- **Delete**: Use the **Delete** icon to delete a Batch that is no longer required in the system. You cannot delete seeded batches.

7.1.3.3 Create a Batch Group

You can create a new batch group in the Define Batch window and schedule and monitor the batch group that you created.

To create a new Batch Group:

1. In the **Define Batch** window, click **Add**.

   The **Create a New Batch** window is displayed.
2. Select **Batch Group** option.
3. Specify the following fields:
   - Name
   - Description
   - Add Batches: This is a multi-select field. You can select the batches that you want to add to the group using this field.
   - Notify on mail

> **Note:**
>
> The Add Batches is a multi-select field, you can select the batches that you want to add to the group using this field.

Choose the desired option to notify by an e-mail the batch execution status depending on the property selected. E-mail for the email notification is fetched from the IDCS configuration for the logged in user.

> **Note:**
>
> E-mail is a mandatory field in IDCS.

The BATCH_NOTIFY_FUNCT function must be enabled to access this feature.

If it is not available for the logged in user, this field is grayed out.

- Every Time: An e-mail is triggered irrespective of the batch status.
- Never: No e-mail will be triggered.

- On Error only: An e-mail is triggered only when the batch execution has failed/Error.
- On Interrupt only: An e-mail is triggered if the batch is successfully interrupted.

> **Note:**
>
> Never is the default option.

4. Click **Save**.

   The new Batch Group is created and displayed in the **Define Batch** window.

7.1.3.4 Manage Batch Group

You can manage a batch group by editing, copying, or delete them as required.

- **Edit**: Use the edit icon to edit the Batch Group name, Added Batches, and Batch Group Description.
- **Copy**: Use the copy option to copy a batch group that you want to clone or create instances in the system.
- **Delete**: Use the delete icon to delete a batch group that is no longer required in the system.

7.1.3.5 Define Batch

The Define Batch window displays the details of all existing Batch like Batch ID, Batch Name, Batch Description, Last Modified By and Last Modified Date. This window allows you to create a new, edit, copy, and delete the batches. You can also create a Batch Group and execute the Batch Group which has the list of batches that you have selected and grouped.

To navigate to the Define Batch window, click Define Batch option from the Header in the Dashboard window. After selecting the batch, you can select the **Monitor** icon corresponding to the batch to proceed to create or edit tasks.

7.1.3.5.1 Batch

Batch is a process of execution of the Date and time-based background tasks based on a defined period during which the resources were available for batch processing.

7.1.3.5.1.1 Edit a Batch

The **Edit Batch** option allows you to edit the Batch details such as Batch Description, Service URL Name and Service URL and also add a new Batch Parameter.

> **Note:**
>
> Seeded batches cannot be edited.

To modify a Batch, perform the following steps:

1. In the Define Batch window, click **Edit** corresponding to the Batch you want to modify.

   The **Edit Batch** window is displayed.

2. Modify the required Batch details.

For more information, see Create a Batch section.

3. Click **Save**.

   The edited batch is saved and displayed in the Define Batch window.

7.1.3.5.1.2 Copy a Batch

The **Copy Batch** option allows you to copy a Batch that you want to clone or create instances in the system from the **Define Batch** window.

To copy a Batch, perform the following steps:

1. In the **Define Batch** window, click **Copy Batch** corresponding to the Batch that you want to copy.

   The **Copy Batch** window is displayed.

2. Specify the Batch details as you want to clone and copy the existing batch.

   For more information, see Create a Batch section.

3. Click **Save**.

   The copied batch is saved and displayed in the **Define Batch** window.

7.1.3.5.1.3 Delete a Batch

The Delete Batch option allows you to delete a Batch that are no longer required in the system from the Define Batch window.

> **Note:**
>
> Seeded batches cannot be deleted.

To delete a Batch, perform the following steps:

1. From the **Define Batch** window, click **Delete** corresponding to the Batch you want to delete.
2. Click **OK** in the confirmation dialog to confirm deletion.

   If the batch has any active schedules a warning is displayed.

   Upon confirmation, all schedules of the batch are also deleted.

7.1.3.5.2 Batch Group

Batch Group is a process of grouping the batches that are required to be execute together for execution of the Date and time-based background tasks based on a defined period during which the resources were available for batch processing.

7.1.3.5.2.1 Edit a Batch Group

The **Edit Batch Group** option allows you to edit the Batch Group details such as Batch Group name, Added Batches, and Batch Group Description.

To modify a Batch Group, perform the following steps:

1. In the **Define Batch** window, click **Batch Group** option to list the batch groups.
2. Click **Edit** corresponding to the Batch Group you want to modify.

The **Edit Batch** window is displayed.

3. Modify the required Batch Group details.

   For more information, see Create a Batch Group section.

4. Click **Save**.

   The edited batch group is saved and displayed in the **Define Batch** window.

7.1.3.5.2.2 Copy a Batch Group

The **Copy Batch** option allows you to copy a Batch that you want to clone or create instances in the system from the Define Batch window.

To copy a Batch Group, perform the following steps:

1. In the **Define Batch** window, click Batch Group option to list the batch groups.
2. Click **Copy Batch** corresponding to the Batch Group that you want to copy.

   The Copy Batch window is displayed.

3. Specify the Batch Details as you want to clone and copy the existing batch.

   For more information, see Create a Batch Group section.

4. Click **Save**.

   The copied batch group is saved and displayed in the **Define Batch** window.

7.1.3.5.2.3 Delete a Batch Group

The **Delete Batch** option allows you to delete a Batch that are no longer required in the system from the **Define Batch** window.

To delete a Batch Group, perform the following steps:

1. From the **Define Batch** window, click **Batch Group** option to list the batch groups.
2. Click **Delete** corresponding to the Batch Group you want to delete.
3. Click **OK** in the confirmation dialog to confirm deletion.

7.1.3.6 Define Tasks

The **Define Tasks** window displays the list of tasks associated with a specific Batch definition. You can create new tasks, edit the existing tasks or delete unwanted tasks. Additionally, you can specify task precedence for each task in Task Precedence window and click the Schedule icon to Schedule the batch.

The **Define Tasks** window allows you to perform the following task operations for your batch and batch group:

- Batch
- Batch Group

7.1.3.6.1 Batch

Batch is a process of execution of the Date and time-based background tasks based on a defined period during which the resources were available for batch processing. You can perform the following operation for the batch based on the task.

- Add a Task

- Modify a Task
- Define Task Precedence
- Delete a Task

7.1.3.6.1.1 Modify a Task

Modifying a task option allows you to modify the details of existing tasks of a Batch definition such as Task Description, Task Type, Batch Service URL and Task Service URL. You can also add a new task parameter and enable or disable already existing task parameters.

To modify a Task, perform the following steps:

1. From the **Define Task** window, select the Batch whose task details you want to modify, from the **Select** drop-down list.
2. Click **Edit** corresponding to the Task whose details you want to modify.

   The **Edit Task** window is displayed.
3. Modify the required Task Details.

   For more information, see Add a Task section.
4. Click **Save**.

7.1.3.6.1.2 Delete a Task

You can remove a task from a Batch definition which are no longer required in the system by deleting it from the **Define Task** window.

To delete a Task, perform the following steps:

1. From the **Define Task** window, select the Batch whose task details you want to delete from the **Select** drop-down list.
2. Click **Delete** corresponding to the Task you want to delete.
3. Click **OK** in the confirmation dialog to confirm deletion.

7.1.3.6.2 Batch Group

Batch Group is a process of grouping the batches that are required to be execute together for execution of the Date and time-based background tasks based on a defined period during which the resources were available for batch processing. You can perform the following operation for the batch based on the task.

- Define Task Precedence

7.1.3.6.2.1 Define Task Precedence

Task Precedence indicates the execution-flow of a Batch. Task Precedence value facilitates you to determine the order in which the specific Tasks of a Batch are executed.

For example, consider a Batch consisting of 4 Task. The first 3 Task does not have a precedence defined and hence will be executed simultaneously during the Batch execution. But, Task 4 has a precedence value as task 1 which indicates that, Task 4 is executed only after Task 1 has been successfully executed.

You can set Task precedence between Tasks or define to run a Task after a set of other tasks. However, multiple tasks can be executed simultaneously, and cyclical execution of tasks is not permitted. If the precedence for a Task is not set, the Task is executed immediately on Batch execution.

To define the task precedence in the Define Task window, perform the following steps:

1. Click the **Menu** button corresponding to the task for which you want to add precedence task.

   The **Task Precedence Mapping** window is displayed.

   The Task Precedence option is disabled if a batch has only one task associated.

   a. Select the batch that you want to execute before the current task, from the Available Tasks pane and click **Move**. You can press Ctrl key for multiple selections.

   b. To select all the listed batches, click **Move All**.

   c. To remove a batch, select the task from the Selected Tasks pane and click **Remove**.

   d. To remove all the selected batches, click **Remove All**.

2. Click **Save** to update Task Precedence in the batches.

3. Click **Preview** to view the Precedence information.

7.1.4 Configure Tasks

The **Define Tasks** window displays the list of tasks associated with a specific Batch definition. You can create new tasks, edit existing tasks, or delete unwanted tasks. Additionally, you can specify task precedence for each task in the Task Precedence window and schedule the batch.

Use the **Define Tasks** window to perform task operations for your batch and batch group.

7.1.4.1 Add Task to a Batch or Batch Group

You can add new tasks to a selected Batch or Batch Group definition.

To add new task:

1. Click **Define Tasks** from the Header panel.

   The **Define Task** window is displayed.

2. Select the Batch or Batch Group for which you want to add new task from the Select drop-down list.

3. Click **Add**.

   The **Create a New Task** window is displayed.

   When creating a task, there are multiple types of components to choose from:

   Model

   - **Model**: Select the Model of selected Objective. It can be any specific model of Objective or All models of Objective. If the ALL_CHAMPION is selected here, then Objectives with no Champion Model is skipped, and the Objectives with Champion Models gets executed. If CHAMPION is selected, and no Champion is present, then Model Execution gets fail.
   - **Objective**: Select the Object which you want to use for execution. For more information, see the Create Objective (Folders) section. The Sub Objective is displayed with path. For example, if Test1 is Objective and Test11 is Sub Objective, and you want to use Test11 Objective for execution, then select this field as Test1/Test11.
   - **Link Types**: You can set this parameter to **Yes** or **No**. If Synchronous Execution is set to **Yes**, then execution will wait for the notebook execution status. If Synchronous

Execution is set to **No**, then execution will not wait for the notebook execution status, it will trigger the notebook and update task status as successful in batch monitor.

- **Synchronous Executions**: You can set this parameter to **Yes** or **No**. If Synchronous Execution is set to **Yes**, then execution will wait for the notebook execution status. If Synchronous Execution is set to **No**, then execution will not wait for the notebook execution status, it will trigger the notebook and update task status as successful in batch monitor.
- **Optional Parameters**: This is used pass the parameters dynamically.

Populate Workspace

- **Global Filter**: Provided input will be applied as a data filter on all the tables selected for data sourcing.
- **Table Filter**: Users can provide data filters individually on the tables here. Please note that you may provide multiple table names for the same SQL filter. Global Filters will not be applicable for those tables on which filters have been applied individually. In case the same table name is provided in more than one rows here, the filter condition will be generated as a conjunction of all the provided filters.
- **Additional Parameters**: Additional parameters include source/target pre-scripts, fetch size, and commit size to optimize data processing and control execution behavior.

4. Enter the details as given here:

Table 7-2 Fields in the Create a New Task window and their Description

Field	Description
Task Name	Enter the task name. **NOTE:** • The Task Name must be alphanumeric and should not start with a number. • The Task Name should not exceed 60 characters in length. • The Task Name should not contain any special characters except underscore (_).
Task Description	• Enter the task description along with spaces and alpha-numeric. No special characters are allowed in Task Description. • Words like Select From or Delete From should not be entered in the Description.
Task Type	Select the task type from the drop-down list: **Rest**: Used to invoke RESTful APIs. Only POST method is supported commonly used to trigger downstream services (For example: models, graphs and so on) or workflows. **Executable**: Used to run shell scripts (.sh) or other system-level executables.

Table 7-2 (Cont.) Fields in the Create a New Task window and their Description

Field	Description
Batch Service URL	Select the required Batch Service URL from the drop-down list. This can be blank, and you can provide the full URL in the Task Service URL field. **WORKSPACE_URL**: Seeded batch URL base endpoint of the batch service responsible for managing and orchestrating batch or batch group executions related to workspace APIs. **MMG_SERVICE_URL**: Seeded batch URL base endpoint of the batch service responsible for managing and orchestrating batch or batch group executions related to MMG APIs (For example: Model Execution). **CS_SERVICE_URL**: Seeded batch URL base endpoint of the batch service responsible for managing and orchestrating Batch or Batch Group executions related to CS APIs.
Component	These are preceded metadata types, which can be executed through Scheduler. The user is allowed to select the required components with regards to the select Bacth URL, or the user can create their own CUSTOM URLs by selecting the CUSTOM component.
Task Service URL	Enter task service URL if it is different from Batch or Batch Group Service URL. **Batch URL**: This is the base endpoint of the batch service responsible for managing and orchestrating batch or batch group executions. Users can append the task end points in the Define Task. There is nothing as Batch group service URL. **Task Service URL**: This is the endpoint for executing individual tasks (For example, REST calls). Users can either provide an end point path to be appended to the Batch URL provided in the batch definition or specify a full REST URL for independent execution.

5. From the Task Parameters pane, click **Add** to add a new Task Parameter. By default, all Batch level parameters are added and enabled as task parameters. To disable, deselect the check box corresponding to the task parameter ( )
   a. Enter the Parameter name in the **Param Name** field.
   b. Enter the Parameter value in the **Param Value** field.
   You can delete a parameter by clicking **Delete** corresponding to the parameter.
6. Click **Save**.

> **Note:**
>
> **Header Parameter**: This parameters are passed in the request header of the execution API to provide metadata like workspace, service ID and so on.

7.1.4.2 Manage Tasks

You can manage tasks by modifying or deleting tasks and also define task precedence.

- **Modify**: Use the edit icon to modify details of existing tasks of a Batch definition such as Task Description, Task Type, Batch Service URL, and Task Service URL. You can also add a new task parameter and enable or disable existing task parameters.

- **Delete**: Use the delete icon to remove a task from a Batch definition from the Define Task window.
- **Define Task Precedence**: Task Precedence specifies the execution-flow of a Batch. Task Precedence value facilitates you to determine the order in which the tasks of a Batch are executed.

For example, consider a Batch consisting of 4 tasks. The first 3 tasks do not have a precedence defined and hence will be executed simultaneously during the Batch execution. But, task 4 has a precedence value as task 1 which indicates that, task 4 is executed only after task 1 has been successfully executed.

You can set Task precedence between Tasks or define to run a Task after a set of other tasks. However, multiple tasks can be executed simultaneously; cyclical execution of tasks is not permitted. If the precedence for a task is not set, the task is executed immediately on execution of the Batch.

To define the task precedence in the Define Task window:

1. Click **Menu** button corresponding to the task for which you want to add a precedence task.

   The **Task Precedence Mapping** window is displayed.

   > **Note:**
   >
   > The Task Precedence option is disabled if a batch has only one task associated.
   >
   > a. Select one or multiple tasks you want to execute before the current task, from the **Available Tasks** pane and click >. Press Ctrl key to select multiple items.

2. Click **Save**.
3. Click **Preview** to view the precedence information.

7.1.4.3 Define Task Precedence

Task Precedence specifies the execution-flow of a Batch. Task Precedence value facilitates you to determine the order in which the specific Tasks of a Batch are executed.

For example, consider a Batch consisting of 4 Task. The first 3 Task does not have a precedence defined and hence will be executed simultaneously during the Batch execution. But, Task 4 has a precedence value as task 1 which indicates that, Task 4 is executed only after Task 1 has been successfully executed.

You can set Task precedence between Tasks or define to run a Task after a set of other tasks. However, multiple tasks can be executed simultaneously, and cyclical execution of tasks is not permitted. If the precedence for a Task is not set, the Task is executed immediately on Batch execution.

To define the task precedence in the Define Task window:

1. Click the **Menu** button corresponding to the task for which you want to add precedence task.

   The **Task Precedence Mapping** window is displayed.

> **Note:**
>
> The Task Precedence option is disabled if a batch has only one task associated.

a. Select the Task you want to execute before the current task, from the **Available Tasks** pane and click **Move**. You can press Ctrl key for multiple selections.

b. To select all the listed Tasks, click **Move All**.

c. To remove a Task, select the task from the Selected Tasks pane and click **Remove**.

d. To remove all the selected Tasks, click **Remove All**.

2. Click **Save** to update Task Precedence.
3. Click **Preview** to view the Precedence information.

7.1.4.4 Exclude or Include Tasks

You can exclude tasks or include the excluded tasks during Batch Group Execution. The excluded task components are therefore executed in the normal process assuming that the excluded task have completed execution.

To exclude/include tasks, perform the following steps:

1. In the **Schedule Batch** window, click **Exclude Tasks**.

   The **Select Tasks** window is displayed.
2. To exclude tasks:

   a. Select the required task from the **Included Tasks** list and click **Move**. You can press Ctrl key for multiple selections.

   b. To exclude all tasks, click **Move All**.
3. To include the excluded tasks:

   a. Select the required task from the **Excluded Tasks** list and click **Remove**. You can press Ctrl key for multiple selections.

   b. To include all excluded tasks, click **Remove All**.
4. Click **Save**.

7.1.5 Schedule Batch or Batch Group

Use the *Schedule Batch / Batch Group* feature to run, schedule, re-start, re-run batches in the Scheduler Service. After you upload the data in the required format into Object Storage, you must load the data into the system using the Scheduler Service. You can schedule them to run in a required pattern and view the run time status of the scheduled services using the Monitor Batch feature.

You can perform operations on batches and batch groups. Information in the following sections is applicable to both batches and batch groups. Where there are unique instructions for batches or batch groups, they are called out clearly.

7.1.5.1 Execute a Batch/Batch Group

Use the Execute batch option to run a batch/batch group instantaneously.

To execute a batch:

1. Click **Schedule Batch**from the Header panel.

   The **Schedule**window is displayed.
2. Select the Batch Name from the **Select Name** drop down menu.

   For example, AMLDataLoad.

   You can preview the schedule of a Batch/Batch Group by clicking on Preview button.
3. Click **Execute**.

   The Execution Status Dialog Box is displayed with the Batch executed successfully message.

   This indicates the unique identification reference number for the batch and date of the batch execution.
4. In the **Execution Status** Dialog Box, click **Monitor** to monitor the batch.
5. If you want to exclude/include some tasks, click **Exclude Tasks**.

   For more information, see Exclude/Include Tasks section.
6. If you want to hold/release some tasks, click **Hold Tasks**.

   For more information, see Hold/Release Tasks section.
7. If you want to edit the dynamic parameters of the batch, click **Edit Dynamic Parameters** .

   For more information, see Edit Dynamic Parameters section.

7.1.5.2 Schedule a Batch/Batch Group

You can schedule a Batch to run just for Once, Daily, Weekly, Monthly or a Cron Expression for scheduling the batches. You can also have a user defined schedule to schedule and run a batch/batch group.

7.1.5.2.1 Schedule Once

To schedule a Batch to run once:

1. Click **Schedule Batch** from the Header panel.

   The **Schedule Batch** window is displayed.
2. In the Schedule Batch window, click **Once**.
3. Select the Batch or Batch Name you want to schedule for once from the **Select** list.
4. Enter a **Schedule Name**.
5. Click the date and time picker icons and select the date and time when you want to run the batch.
6. Click **Schedule**.

7.1.5.2.2 Schedule Daily

To schedule a Batch to run daily:

1. In the **Schedule Batch** window, click **Daily**.
2. Select the Batch or Batch Name you want to schedule daily from the **Select** list.
3. Enter a **Schedule Name**.

4. Click the date and time picker icons and select the date and time when you want to run the batch.
5. Click **Schedule**.

7.1.5.2.3 Schedule Weekly

To schedule a Batch to run weekly:

1. In the **Schedule Batch** window, click **Weekly**.
2. Select the batch or batch name you want to schedule weekly from the **Select** list.
3. Enter a **Schedule** Name.
4. Click the date and time picker icons to select the start date, end date, and time when you want to run the batch.
5. Select the days of the week when you want to run the batch from the **Select Days of the Week** list.
6. Click **Schedule**.

7.1.5.2.4 Schedule Monthly

To schedule a Batch to run monthly:

1. In the **Schedule Batch**window, click **Monthly**.
2. Select the Batch or Batch Name you want to schedule monthly from the Select list.
3. Enter a **Schedule** Name.
4. Click the date and time picker icons to select the start date, end date, and time when you want to run the batch.
5. Select the days of the week when you want to run the batch from the **Select Days of the Week** list.
6. Click **Schedule**.

7.1.5.2.5 Schedule Cron Expression

To run a Batch in a user-defined schedule, you can have custom schedule with the help of Cron Expression. A Cron Expression is a string comprised of 6 or 7 fields separated by white space. Fields can contain any of the allowed values, along with various combinations of the allowed special characters for that field. For more information, click the information icon.

To create a customized schedule for a batch using a Cron Expression:

1. In the **Schedule Batch**window, click **Cron Expression**.
2. Select the batch or batch name you want to schedule from the Select list.
3. Enter a **Schedule Name**.
4. Enter the Cron Expression for your schedule. For more information about the Cron Expression, click the **Information**icon.
5. Click **Schedule**.

7.1.5.2.6 Pause a Batch

You can pause a batch that has been executed. User can click on pause button and can pause schedule for between scheduled dates. Time is not supported for pause activity, if any

schedule is paused for current date, it will be paused immediately. If pause start date is passed, user cannot change pause start date, rest fields can be manipulated and will be valid one minute before end of day.

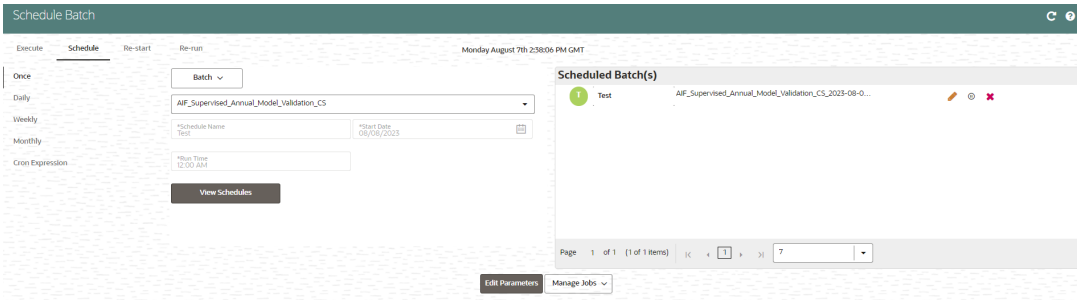
If pause is expired, user cannot edit and delete but will be shown in UI.

To pause a batch:

1. From the Header panel, click **Schedule Batch** and then click the **Schedule** tab.
2. Select the batch you want to pause from the list.
3. For the batch you have selected to pause, select the scheduling type (Once, Daily, Weekly, Monthly, or Cron Expression).
4. Click **View Schedules**.

   The scheduled batches are displayed on the right pane.

Figure 7-3 Scheduled Batches



5. Click the pause icon for the batch you want to pause.
6. In the **Manage Pause** screen, click **Add Pause**.
7. Select the **Pause Start** and **End Date**, add comments if required, and click +.

   The Pausing schedule appears in the Pause Summary.
8. Repeat the steps if you want to add multiple pausing schedules for the same batch.

7.1.5.2.7 Re-start a Batch/Batch Group

You can restart a Batch which has not been executed successfully or which has been explicitly interrupted, or cancelled, or put on hold during the execution process. By restarting a Batch, you can continue Batch execution directly from the point of interruption or failure and complete executing the remaining tasks.

To re-start a batch:

1. Click **Schedule Batch**from the Header panel.

   The **Schedule** window is displayed.
2. From the **Schedule** window, select **Re-start** tab.
3. Select the Batch Group you want to re-start from the Select Name drop down menu.
4. Select the Batch Run ID.
5. Click **Re-start**.

7.1.5.2.8 Re-run a Batch/Batch Group

You can re-run a Batch which has previously been executed. Rerun Batch facilitates you to run the Batch irrespective of the previous execution state. A new Batch Run ID is generated during the Rerun process and the Batch is executed as similar to the new Batch Run.

To re-run a batch:

1. Click **Schedule Batch** from the Header panel.

   The **Schedule Batch** window is displayed.
2. In the **Schedule Batch** window, select **Re-run** tab.
3. Select the Batch or Batch Name you want to re-run from the **Select Name** list.
4. Select the **Batch Run ID**.
5. Click **Re-run**.

7.1.5.2.9 Edit Dynamic Parameters

Dynamic Parameters facilitate modification of dynamic parameters for the batch. You can change the param value and save the changes to the batch. The **Edit Dynamic Parameters** option is available in all the tab in the *Schedule Batch* window.

To edit the dynamic parameters for a batch/batch group:

1. In the **Schedule Batch** window, click **Edit Dynamic Parameters**.
2. In the **Edit Dynamic Params** window, modify the values as required.
3. Click **Save**.

   The modified parameters are applied to the Batch.

7.1.5.2.10 Task Definitions of a Batch/Batch Group

You can modify the task definition state in the Batch Execution window to exclude or hold the defined task in a Batch from execution. The excluded tasks are therefore assumed to have completed execution and get excluded during the Batch Run.

While executing or scheduling a batch group from the Schedule Batch window, you can:

- Exclude a task or include the excluded task.
- Hold a task or release the held task.

7.1.5.2.10.1 Exclude or Include Tasks

You can exclude tasks or include the excluded tasks during Batch/Batch Group Execution. The excluded task components are executed in the normal process assuming that the excluded tasks have completed execution.

To exclude/include tasks:

1. In the **Schedule Batch** window, click **Exclude Tasks**.

   The **Select Tasks** window is displayed.
2. To exclude tasks:
   a. Select one or multiple tasks from the Included Tasks list and click > or >> respectively. Press the Ctrl key and click to select multiple items.

3. To include the excluded tasks:
   a. Select one or multiple tasks from the **Excluded Tasks** list and click < or << respectively. Press the Ctrl key and click to select multiple items.
4. Click **Save**.

7.1.5.2.10.2 Hold or Release Tasks

You can hold tasks or release the held tasks during Batch Group Execution. The tasks which are on hold along with the defined components are skipped during execution. However, at least one task should be available in a Batch without being held or excluded for Batch execution.

To hold/release tasks:

1. In the **Schedule Batch** window, click **Hold Tasks**.

   The **Select Tasks** window is displayed.
2. To hold tasks:
   a. Select the required task from the **Released Tasks** list and click **Move**. You can press Ctrl key for multiple selections.
   b. To hold all tasks, click **Move All**.
3. To release held tasks:
   a. Select the required task from the **Held Tasks** list and click **Remove**. You can press Ctrl key for multiple selections.
   b. To release all held tasks, click **Remove All**.
4. Click **Save**.

7.1.5.2.10.3 Pre-Conditions for a Batch Group

You can schedule the batches and set the pre-conditions within a Batch group with frequency as Weekly, Monthly, or based on an Interval for scheduling the batches. The batch that satisfies the configured pre-conditions are executed as part of the schedule.

> **Note:**
>
> Pre-Conditions can only be applied when using the Schedule option in the Schedule Batch window.

Weekly

To set the pre-conditions to the batches in a batch group weekly, perform the following steps:

1. In the **Schedule Batch** window, you can select either Once, Daily,Weekly, Monthly, or Cron Expression option based on the schedule that you want to run the batch group.
2. Select the Batch Group you want from the Select drop down menu.
3. Enter a Schedule Name.
4. Specify the other details displayed when you are selecting Once, Daily,Weekly, Monthly, or Cron Expression.
5. Click Pre-Conditions.
6. In the Pre-Conditions window, specify the Batch from the drop down and from the Frequency drop down and select Weekly.

7. Select the days from the Select Days drop down that you want to schedule the batch run within the selected week.
8. Click **Add** to add the specified entry.

> **Note:**
>
> Pre-Conditions can be added only to one batch at a time.

9. Click **Save**.
   The batch is executed based on the configured pre-conditions.

Monthly

To set the pre-conditions to the batches in a batch group monthly, perform the following steps:

1. In the **Schedule Batch** window, you can select either Once, Daily,Weekly, Monthly, or Cron Expression option based on the schedule that you want to run the batch group.
2. Select the Batch Group you want from the Select drop down menu.
3. Enter a Schedule Name.
4. Specify the other details displayed when you are selecting Once, Daily,Weekly, Monthly, or Cron Expression.
5. Click Pre-Conditions.
6. In the Pre-Conditions window, specify the Batch from the drop down and from the Frequency drop down and select Monthly.
7. Select the days from the Select Days drop down that you want to schedule the batch run within the selected week.
8. Click **Add** to add the specified entry.

> **Note:**
>
> Pre-Conditions can be added only to one batch at a time.

9. Click **Save**.
   The batch is executed based on the configured pre-conditions.

Interval

To set the pre-conditions to the batches in a batch group based on an interval, perform the following steps:

1. In the **Schedule Batch** window, you can select either Once, Daily,Weekly, Monthly, or Cron Expression option based on the schedule that you want to run the batch group.
2. Select the Batch Group you want from the Select drop down menu.
3. Enter a Schedule Name.
4. Specify the other details displayed when you are selecting Once, Daily,Weekly, Monthly, or Cron Expression.
5. Click Pre-Conditions.
6. In the Pre-Conditions window, specify the Batch from the drop down and from the Frequency drop down and select Interval.

7. Select the interval from the Custom Recurrence (Repeat every) Days drop down that you want to schedule the batch run within the selected week.

> **Note:**
>
> The Custom Recurrence can be set maximum to 60 days.

8. Click **Add** to add the specified entry.

> **Note:**
>
> Pre-Conditions can be added only to one batch at a time.

9. Click **Save**.
10. The batch is executed based on the configured pre-conditions.

7.1.5.3 Edit Dynamic Parameters

Dynamic Parameters facilitates you to the modify the dynamic parameters for the batch. You can change the param value from the Edit Dynamic Params window and save the changes to the Batch. The Edit Dynamic Parameters option is available in all the tab in the **Schedule Batch** window.

To edit the dynamic parameters for a batch group, perform the following steps:

1. In the **Schedule Batch** window, click **Edit Dynamic Parameters**.

   The **Edit Dynamic Params**window is displayed.
2. In the **Edit Dynamic Params** window, modify the values as required.
3. Click **Save**.

   The modified parameters are applied to the Batch.

7.1.5.3.1 Accessing Scheduler Optional Parameters Inside Studio Paragraphs

The Scheduler Service automates the running of models in the application. The Scheduler Service contains a graphical user interface and a single control point for defining and monitoring background executions.

For more details on the Scheduler Service, see Scheduler Service section.

To access the optional parameters passed from the Scheduler during execution of any models, run the following script:

```
To print
print('BATCHRUNID value is : ${$BATCHRUNID$}')
print('TASKID value is : ${$TASKID$}')
print('FICMISDATE value is : ${$FICMISDATE$}')
```

To access the optional parameters passed during execution of any models, run the following scripts:

```
%python
print('threshold value is : ${threshold}')
```

> **Note:**
>
> Threshold is the optional parameter passed during the model execution.

From scheduler task/edit dynamic parameters, the aprameter set to a model can be passed in optional parameter textbox as parameterSets=Set1 where parameterSets is the key for paramter sets and Set1 is the actual set value.

From EICS it is as below:

```
./EICSchedulerService.sh -o 1 -u mmgadmin -w CS -c ofsauser -s secret -x 232 -
b BatchModel -t rest -v '{"batchParams":
{"$FICMISDATE$":"2021-11-30","$BATCHRUNID$":"BATCHRUNID"},"taskRuntimeParams":
{"Task1":{"optionalparams":"parameterSets=Set1"},"Task2":
{"optionalparams":"parameterSets=Set2"}}
```

7.1.6 Monitor Batch or Batch Groups

Use the Monitor Batch feature to view the status of executed batch/batch group along with the task's details. You can track the issues if any, on regular intervals and ensure smoother batch/ batch group execution. A visual representation as well as tabular view of the status of each task in the batch/batch group is available.

You can monitor operations for batches and batch groups using the Monitor Batch feature.

To monitor a batch/batch group:

1. Click **Monitor Batch** from the Header panel.
   The **Monitor** window is displayed.
2. Select the batch/batch group from the **Select** drop-down and then select the **Batch Run ID** from the **Run ID** list.
3. Click **Start Monitor**.
   The result is displayed in **Visualization** and **List View** tabs.
   - On the Visualization tab are charts and in List View are details in a tabular format with the following details:
     a. i. Batch Status: Displays the batch status—NOT-STARTED, ON-GOING, SUCCESSFUL, FAILED, INTERRUPTED, EXCLUDED, HELD, and UNDEFINED.
     b. Batch Start Time: The batch start time.
     c. Batch End Time: The batch end time.
     d. Task Details: Mouseover the task to display task status and additional details.
     e. More Information: The message returned by the Rest Service.

4. Select **Stop Monitor** if you wish to stop monitoring. You can also specify the Start and Stop Monitor options along with refresh interval by providing the Refresh every seconds and minutes information.

> **Note:**
>
> - You can select the refresh interval and the duration of the auto refresh. The default refresh interval is 5 seconds and default duration 5 minutes. That is, data is refreshed every 5 seconds for the next 5 minutes.
> - The interval input range must be between 5 to 60 seconds and the duration input range between 5 to 180 minutes.
> - You can use the **Stop Monitor** button to stop the auto refresh.

5. To restart, rerun, or interrupt the monitoring, select the Restart, Rerun, or Interrupt buttons, respectively.
6. To view log information, click the log icon in the **List View** tab.
   The Log Viewer window appears displaying the log details.
7. Click the download icon to download the log, or click the close icon to close the log viewer.

> **Note:**
>
> Downloads PDF button has been introduced in the batch monitor.
> Downloaded PDF is available in: `<ftpsharepath>printservice/pdfs/scheduler/<DATE>`

7.1.6.1 Batch Group

Batch Group is a process of grouping the batches that are required to be execute together for execution of the Date and time-based background tasks based on a defined period during which the resources were available for batch processing.

To monitor a batch group:

1. Click **Monitor Batch** from the Header panel.

   The **Monitor** window is displayed.
2. Select the **Batch Group from the Select** drop-down and then select the **Batch Run ID from the Run ID** drop-down.
3. Click **Start Monitor**.

   The result is displayed in Visualization and List View tabs. Details of these tabs are as follows:

   - The Visualization tab displays the details in the form of a chart represented with the following details:
     - Batch Status: Displays the batch status, the different batch status are NOT-STARTED, ON-GOING, SUCCESSFUL, FAILED, INTERRUPTED, EXCLUDED, HELD, and UNDEFINED.
     - Batch Start Time: Displays the batch start time details.

  - Batch End Time: Displays the batch end time details.
  - Batch Details: Mouseover the task to display its status and details.
- The **List View** tab displays the details in a tabular form with the following details:
  - Batch Status: Displays the batch status, the different batch status are NOT-STARTED, ON-GOING, SUCCESSFUL, FAILED, INTERRUPTED, EXCLUDED, HELD, and UNDEFINED.
  - Batch Start Time: Displays the batch start time details.
  - Batch End Time: Displays the batch end time details.
  - Batch Details: Mouseover the task to display its status and details.
  - More Information: The message returned by the Rest Service.

4. If you wish to stop the monitoring, select Stop Monitor. You can also specify the Start and Stop Monitor options along with refresh interval in the Refresh every seconds and minutes fields.

> **Note:**
>
> - You can select the refresh interval and the duration for the auto refresh. The refresh interval is defaulted to 5 seconds and duration is defaulted to 5 minutes. That is the refresh happens every 5 seconds for next 5 minutes.
> - Range of interval input must be between 5 to 60 seconds and range of duration.
> - Input should be between 5 to 180 minutes.
> - You can use the Stop Monitor button to stop the auto refresh.

5. To restart the Batch Group, select **Restart**.
6. To rerun the Batch Group, select **Rerun**.
7. To interrupt the Batch Group, select **Interrupt**.

   To view the log information about the batch group, click **View Log** in the **List View**tab.

7.1.7 External Interface Component Scheduler Service

Use the **EICSchedulerService.sh** utility located in the `<installed path> mmg-home/bin` folder to perform basic scheduler operations.

Prerequisites:

- Ensure public and private keys are present in the conf folder and the corresponding path is present in `application.properties` of `mmg UI service`.
- You should be aware of the clientId/secret which will be used for basic authentication. (Currently, these values are stored in `application.properties` as `token.clientid` and token.secret).

Table 7-3 List of possible flagged arguments to be passed while executing EICSchedulerService.sh

Flag	Description	Comments
-o	Operation Type: The operation the user wants to perform.	A mandatory argument for all type of requests. Possible Values are: 1 – To trigger an object 2 – To get the current status of a run. 3 – To re-start an execution. 4 – To re-run an execution. 5 – To interrupt an execution.
-u	ofs_remote_user	A mandatory argument for all the type of requests
-w	Workspace Id	A mandatory argument for all the type of requests
-c	Client Id for authentication	A mandatory argument for all the type of requests Currently, in mmg-ui application.properties the parameter token.clientid is set to ofsauser
-s	Client Secret for authentication	A mandatory argument for all the type of requests Currently, in mmg-ui application.properties the parameter token.secret is set to secret
-b	Object Name	Batch or Batch Group Name.
-t	Object Type	For Batch, it is set to rest and for Batch Group, it is set to group.
-r	Batch Execution Id	Execution ID of the Batch or Batch Group.
-x	External Unique Id	The unique ID for every execution. Mandatory for execute and rerun operations.
-i	Included Jobs	In case of trigger, included jobs can be provided as comma separated values using this flag.
-e	Excluded Jobs	In case of trigger, excluded jobs can be provided as comma separated values using this flag.
-h	Held Jobs	In case of trigger, held jobs can be provided as comma separated values using this flag.
-v	Dynamic parameter list	This is an optional parameter to support passing of dynamic runtime parameters for Batch/ Task.

Sample commands to perform Scheduler Operations

- **Trigger/Execute**
  Execute batch “batch1”:

  Syntax: ./EICSchedulerService.sh -o 1 –u < ofs_remote_user > -w < Workspace Id > -c < Client Id for authentication > -s < Client Secret for authentication > -x < External Unique Id > -b < Object Name > -t < Object Type>

  Example: ./EICSchedulerService.sh -o 1 -u scheduser -w WS1 -c ofsauser -s secret -x batch1_1001 -b batch1 -t rests

  Execute batch “batch1” with Dynamic Parameter list

  Syntax: ./EICSchedulerService.sh -o 1 –u < ofs_remote_user > -w < Workspace Id > -c < Client Id for authentication > -s < Client Secret for authentication > -x < External Unique Id > -b < Object Name > -t < Object Type> -v <Dynamic Runtime parameter list>

  Example: ./EICSchedulerService.sh -o 1 -u mmgadmin -w CS -c ofsauser -s secret -x 232 -b BatchModel -t rest -v '{"batchParams": {"$FICMISDATE$":"2021-11-30","$BATCHRUNID$":"BATCHRUNID","param_1":"qwert","param_2":"abcde"},"taskRuntimeParams":{"Task1": {"p1":"11102","p2":"22202","abc":"123342","extraP":"wert2y"},"Task2": {"abc":"1233","p3":"3333","extP2":"0983"}}}'

  > **Note:**
  >
  > The Task/Batch parameters should be the actual parameter key.

  Execute batch “batch1” with “task1” and “task2” included:

  Syntax: ./EICSchedulerService.sh -o 1 –u < ofs_remote_user > -w < Workspace Id > -c < Client Id for authentication > -s < Client Secret for authentication > -x < External Unique Id > -b < Object Name > -t < Object Type> -I < Included Jobs with comma separated values>

  Example: ./EICSchedulerService.sh -o 1 -u scheduser -w WS1 -c ofsauser -s secret -x batch1_1001 -b batch1 -t rest -i task1,task2

  Execute batch “batch1” with “task1” as excluded and “task2” as held:

  Syntax:

  ./EICSchedulerService.sh -o 1 –u < ofs_remote_user > -w < Workspace Id > -c < Client Id for authentication > -s < Client Secret for authentication > -x < External Unique Id > -b < Object Name > -t < Object Type> -e < Excluded Jobs> -h < Held Jobs>

  Example:

  ./EICSchedulerService.sh -o 1 -u scheduser -w WS1 -c ofsauser -s secret -x batch1_1001 -b batch1 -t rest -e task1 -h task2

- **Status**
  Get status of execution having runid = MMG_R1_2022-09-07_1662557327886_1:

  Syntax:

  ./EICSchedulerService.sh -o 2–u < ofs_remote_user > -w < Workspace Id > -c < Client Id for authentication > -s < Client Secret for authentication > -r < Batch Execution Id>

  Example:

  ./EICSchedulerService.sh -o 2 -u scheduser -w WS1 -c ofsauser -s secret -r MMG_R1_2022-09-07_1662557327886_1

- **Re-Start**

Restart execution for batch “MMG_R1” having runid = MMG_R1_2022-09-07_1662557327886_1: Syntax: ./EICSchedulerService.sh -o 3–u < ofs_remote_user > -w < Workspace Id > -c < Client Id for authentication > -s < Client Secret for authentication > -b < Object Name> -r < Batch Execution Id>

Example:

./EICSchedulerService.sh -o 3 -u scheduser -w WS1 -c ofsauser -s secret -b MMG_R1 -r MMG_R1_2022-09-07_1662557327886_1

- **Re-Run**
  Rerun execution for batch “batch1” having runid = batch1 -r batch1_2022-09-01_1662011065557_1:

  Syntax:

  ./EICSchedulerService.sh -o 4–u < ofs_remote_user > -w < Workspace Id > -c < Client Id for authentication > -s < Client Secret for authentication > -x < External Unique Id > -b < Object Name> -r < Batch Execution Id>

  Example:

  ./EICSchedulerService.sh -o 4 -u scheduser -w WS1 -c ofsauser -s secret -x batch1_1002 -b batch1 -r batch1_2022-09-01_1662011065557_1

- **Interrupt**
  Interrupt execution for batch “batch1” having runid = batch1_2022-09-07_1662530601924_1:

  Syntax:

  ./EICSchedulerService.sh -o 5–u < ofs_remote_user > -w < Workspace Id > -c < Client Id for authentication > -s < Client Secret for authentication > -b < Object Name> -r < Batch Execution Id>

  Example:

  ./EICSchedulerService.sh -o 5 -u scheduser -w WS1 -c ofsauser -s secret -b batch1 -r batch1_2022-09-07_1662530601924_1

Note:

- On successful request for execute/rerun, the batchrunid corresponding to external unique id will be stored in the AAICL_SS_EXT_BATCH_RUN_ID_MAPPING table.
- External Unique ID is required only in case of execute/rerun request.

8
More

More feature covers:

- **Export Objects**: Utilize Object Migration feature to efficiently export and save objects.
- **Import Objects**: Import previously exported objects into the system.
- **Model Actions**: Review and approve model deployments in bulk.
- **Audit Trail**: Monitor all system activities and changes.
- **Data Pipelines**: Create and manage data pipelines for efficient data flow and processing.

8.1 Object Migration

Object Migration is the process of migrating or moving objects between environments.

You may want to migrate objects for reasons such as managing global deployments on multiple environments or creating multiple environments so that you can separate the development, testing, and production processes.

Prerequisites to Migrating Objects

In order to migrate objects, users must be mapped to the **Object Migration Admin Group**.

> **Note:**
>
> Identity Administrator Group users cannot migrate objects if they are not mapped to the **Object Migration Admin Group**.

Migration Object Types

You can migrate (import/export) the following object types by clicking the Object Migration icon on the left menu.

- **Schedule**: Provides the instruction to schedule the execution of defined processes. When a schedule is migrated, the associated batch is also migrated.
- **Batch**: A group of jobs that are scheduled to automatically execute at a preset interval of time without any user intervention. When a batch is migrated, the batch and the associated pipeline information are also migrated.
- **Batch Group**: A set of individual batches are consolidated to form a single Batch Group. When you migrate a batch group all the batches, tasks, and pipeline information associated with that batch group are also migrated.
- **Pipeline**: A pipeline is an embedded data processing engine that filters, transforms, and migrates data on-the-fly. Pipelines consist of a set of data processing elements called widgets connected in series, where the output of one widget is the input to the next element.

- **Threshold**: The threshold limit associated with values of set variables for scenarios in FCCM Cloud Service. These threshold values are set when scenarios are created or installed and can be changed, if required.
- **Job**: Jobs provide a set of instructions to execute Workflow Pipelines, based on the set threshold values.
- **PMF Process**: PMF Processes are defined to sequence the Workflow Pipelines, the applications, and to design the artifacts that participate in the pipelines, to implement the pipelines. When exporting a PMF Process, dependent metadata such as data fields, transition rules associated with the PMF process are taken care of.
- **Roles**: Roles are used to map functions to a defined set of groups to ensure user access system security.
- **Groups**: Groups are used to map Roles. Specific User Groups can perform only a set of functions associated with that group.

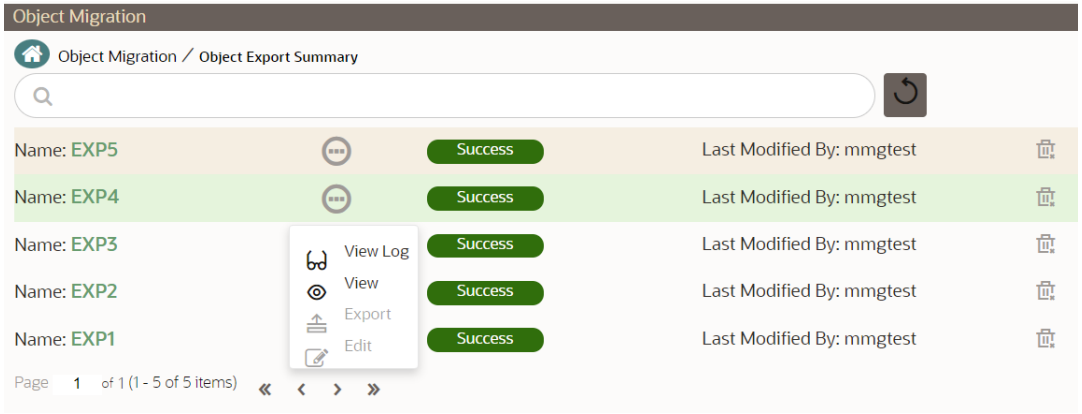
8.1.1 Export Objects

To access the Object Export Summary page:

1. Click **Launch Workspace** next to corresponding Workspace to display the **Dashboard** window with application configuration and model creation menu.
2. Click the **Object Migration** icon and select **Export Objects**.

   The Object Export Summary page containing the records is displayed with the following details.

Figure 8-1 Object Export Summary page



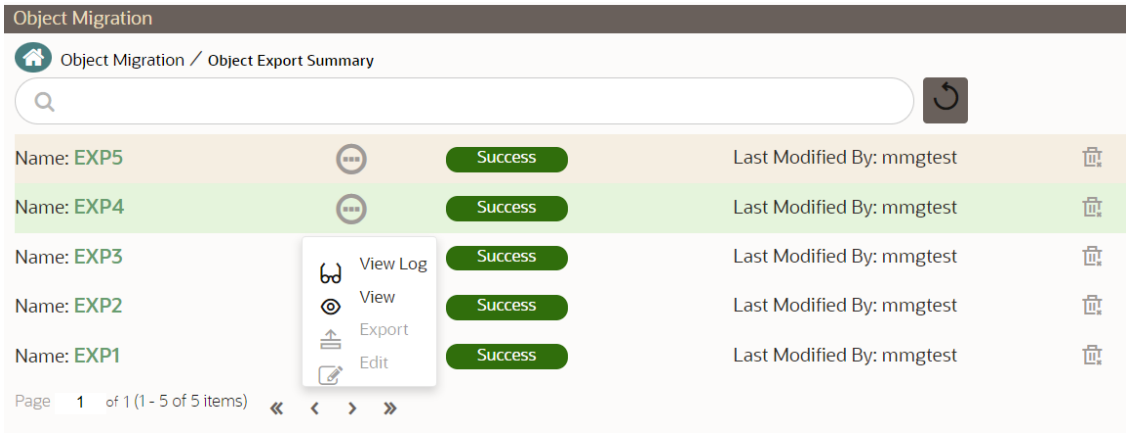
- **Name**: The unique migration name assigned to the collection when the migration definition was created.
- **Object Migration Status**: The migration status of the record corresponding to the specified Definition Name. The three migration status values are as follows:
  - **Success** - Set to Success, when the object export is completed successfully.
  - **Saved** - Set to Saved, when the migration definition is ready for export and needs to import.
  - **Failed** - Set to failed, when the migration definition is not exported successfully.

- **Last Modified By**: The ID of the Last Modified by user who has modified the record.

8.1.1.1 Search Objects to Export

To search for a specific migration definition, type the first few letters of the user name that you want to search in the **Search** box and click the **Search** icon.

Figure 8-2 Search Objects to Export



Use the controls at the bottom of the page, to set the number of entries displayed per page and to navigate between pages.

8.1.1.2 Create Migration Export Object Definitions

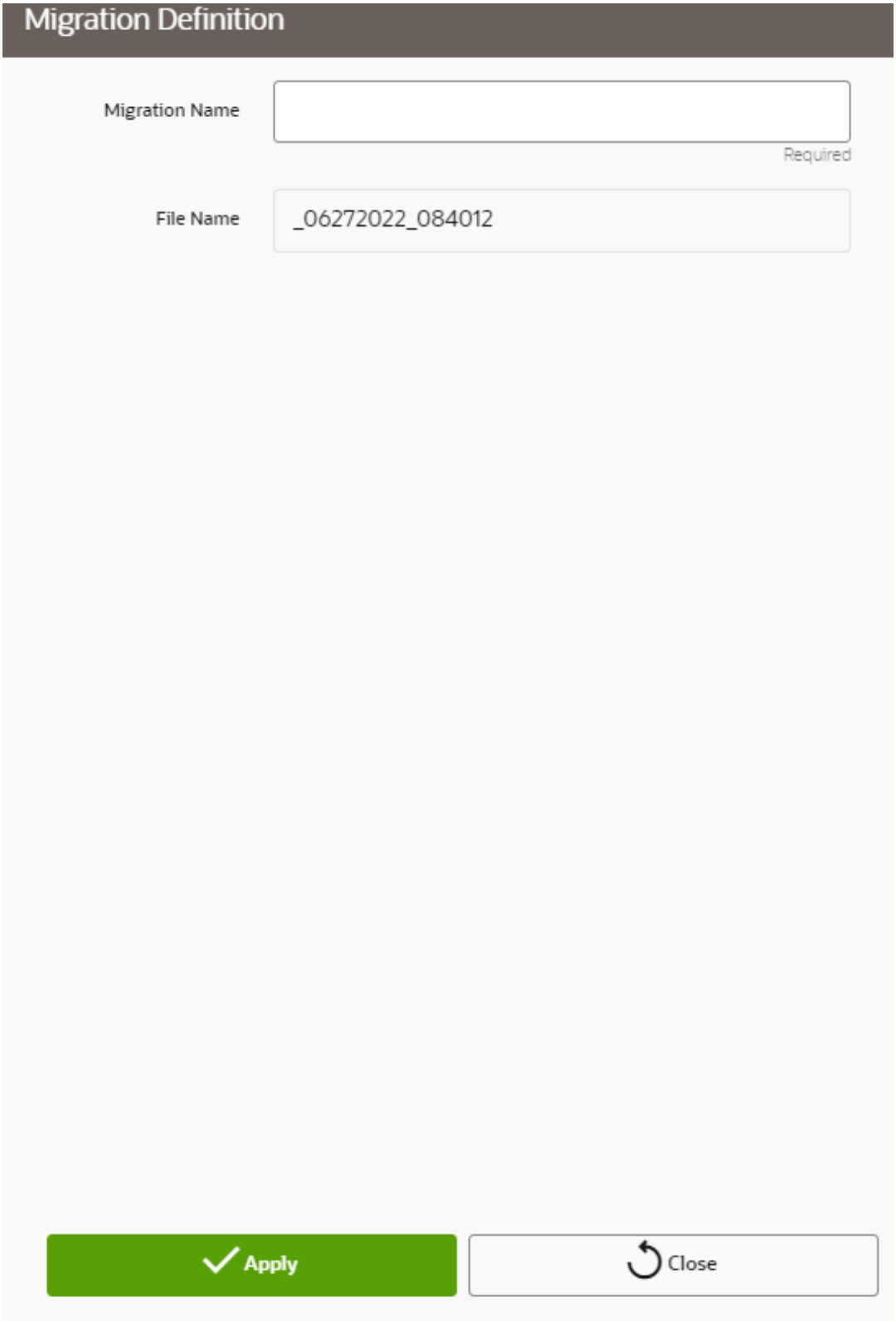
You can create Migration Export object definitions for the following object types.

- Batch Group
- Threshold
- PMF Process
- Batch
- Pipeline
- Schedule
- Job

To create a definition for export of objects to be migrated, do as follows:

1. Click **Add** in the **Object Export Summary** page to display the **Migration Definition** window.

Figure 8-3 Migration Definition window



2. In the **Migration Definition** window, enter the details for the following:
   - **Migration Name**: Enter the code of the export of objects to be migrated definition. This is a unique identifier.

- **File Name**: The system auto-creates the file name of the objects that can be used to export the definition in the following format:
  - **For Business Objects**: Migration Name_BO_Time Stamp (MMDDYYY HHMMSS)
  - **For Identity Objects**: Migration Name_IDM_Time Stamp (MMDDYYY HHMMYY)

3. Click **Apply** to save the details.

   The Object Selection Page is displayed.

4. In the Object Selection page, select the required Object Type from the Object Types drop-down list. The object types listed for the System Configuration tab are, Schedule, Batch group, Batch, Pipeline, Threshold, Job and PMF process. For more information about the object types, refer Migration Object Types.

   You can also enter the first few letters in Search to add a particular object from a selected Object type.

   The list of objects is displayed.

5. Select the objects to be added to the Migrate Definition.

   The selected objects are added to the respective object type branch.

6. Click **Save**to create the Migrate Definition.

   A confirmation message is displayed, when the definition is saved successfully.

   The new migration definition is listed in the Object Export Summary page and the status is set to Saved.

> **Note:**
>
> If the migration definition object is not created successfully and the status is set to Failed. Contact My Oracle Support of more information.

8.1.1.3 View Migration Objects to be Exported

You can view and edit the migration objects from the **Object Export Summary** page.

To view the list of migration export objects associated with a migration definition, highlight the migration definition and click **Menu**button.

You can:

Table 8-1 View Migration Export Objects

Field	Description
View Log	View the migration log details of the selected Migration Definition. For more information, refer to View Log of Migration Export Objects.
View	View the Object Details for a specific Migration Definition. Click an object to view more details. For more information, refer to Viewing Export Migration Object Details.

Table 8-1 (Cont.) View Migration Export Objects

Field	Description
Export	Click Export to initiate Object Migration (Export) for a specific Migration Definition. When the migration is completed, the status will change from Saved to Success. For more information, refer to Exporting Migration Objects.
Edit	Click Edit to view and edit the objects linked to a Migration Definition. For more information, refer to Editing Migration Export Definitions.

8.1.1.3.1 View Log of Migration Objects Exported

To view the log details of object with migration status Success or Failed, follow these steps. The view log facilitates you to view the log information of the definition for export of objects to be migrated with its status.

1. Highlight the migration definition and click Menu button and select View Log.

   The View Log page is displayed.

   The status of the export migration is displayed with the following details.

Table 8-2 View Log of Migration Export Objects

Field	Description
Object Migration ID	The migration ID associated with the export object.
Object Type	The object type of the export object.
Object Code	The object code associated with the export object.
Creation Date	The date of creation of the export object.
Created By	The User Id of the User who created the export object.
Status	The migration status of the export object. • Success - Indicates that the export migration was completed successfully. • Failed - Indicates that the export migration did not complete

> **Note:**
>
> The View Log page for a migration object with status Saved will be empty.

2. Click **OK** to close the window.

8.1.1.3.2 View Details of Exported Migration Objects

To view the list of objects added to a specific Migration Definition:

1. Highlight the migration definition and click **Menu** button and select **View**.

   The list of migration objects added to the definition is displayed.

2. Double-click an object to view the object attribute details.

8.1.1.4 Export Migration Objects

To export the list of objects added to a specific Migration Definition:

- Highlight the migration definition and click **Menu** button and select **Export**.

  The status is indicated as either Success or Failed.

> **Note:**
>
> Use Object Migration (export) to export a set of objects within the same setup or across different setups.

8.1.1.5 Edit Exported Migration Definitions

You can edit the migration export objects that are not exported and their status is Saved or Failed. If the object is already exported and the status is set to Success, you cannot edit the object details.

To edit a record of the definition of export of objects to be migrated, follow these steps. You can add more objects to export or remove existing objects.

1. Highlight the migration definition and click **Menu** button and select **Edit**.

   The Object Selection window is displayed.

2. Update the required details.

3. In the Object Selection window, select the required Object Type from the Object Types drop-down list.

   The object types listed for the System Configuration tab are, Batch_Group, PMF_Process, Batch and Schedule. For more information about the object types, refer Migration Object Types. You can also enter the first few letters in Search to add a particular object from a selected Object type.

   The list of objects is displayed.

4. Select the objects to be added to the Migrate Definition or to be deleted from the Migration Definition.

   The selected objects are added/deleted to the respective object type branch.

5. Click **Save** to edit the Migrate Definition.

   A confirmation message is displayed, when the definition is saved successfully.

   The edited migration definition is listed in the Object Export Summary page and the status is set to Saved.

6. Click **Save** to update the changes.

8.1.1.6 Delete Exported Migration Definitions

You can only delete records in the Saved or Failed state; records in Success state cannot be deleted.

To delete a migrate export object definition, highlight the record to be deleted and click the delete button. Click Yes to confirm and proceed with the deletion.

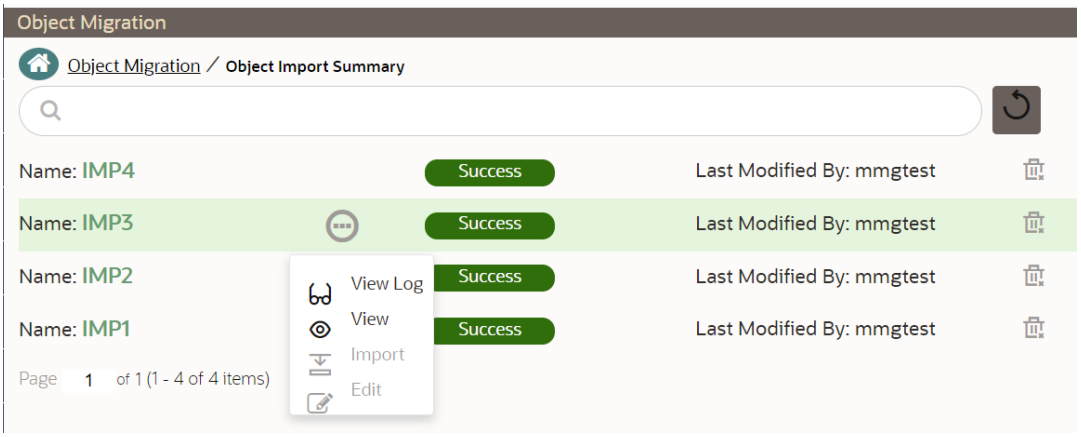
8.1.2 Import Objects

To access the Object, Import Summary page:

1. Click **Launch Workspace** next to corresponding Workspace to display the **Dashboard** window with application configuration and model creation menu.
2. Click the **Object Migration** icon and select **Import Objects**.

   The Object Import Summary page containing the records is displayed with the following details.

Figure 8-4 Object Import Summary page

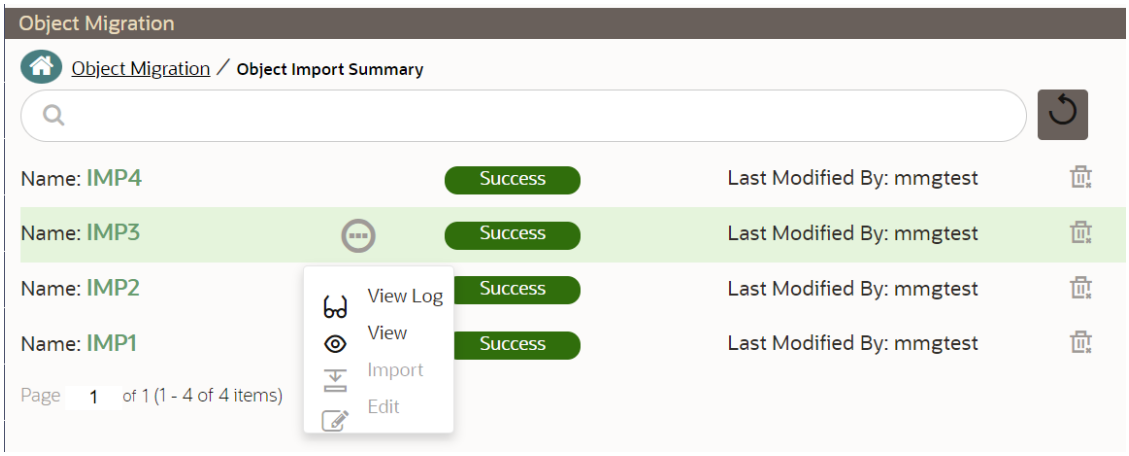


- **Name**: The unique migration name assigned to the collection when the migration definition was created.
- **Object Migration Status**: The migration status of the record corresponding to the specified Definition Name. The three migration status values are as follows:
  - **Success** - Set to Success, when the object import is completed successfully.
  - **Saved** - Set to Saved, when the migration definition is ready for import and needs to import.
  - **Failed** - Set to failed, when the migration definition is not imported successfully.
- **Last Modified By**: The ID of the Last Modified by user who has modified the record.

8.1.2.1 Search Objects to Import

To search for a specific migration definition, type the first few letters of the user name that you want to search in the **Search**box and click search icon. The search results display the names that consist of your search string.

Figure 8-5 Object Migration



At the bottom of the page, you have controls to navigate the search entries and also specify the number of entries available on a single page.

8.1.2.2 Create Import Object Definitions for Migrations

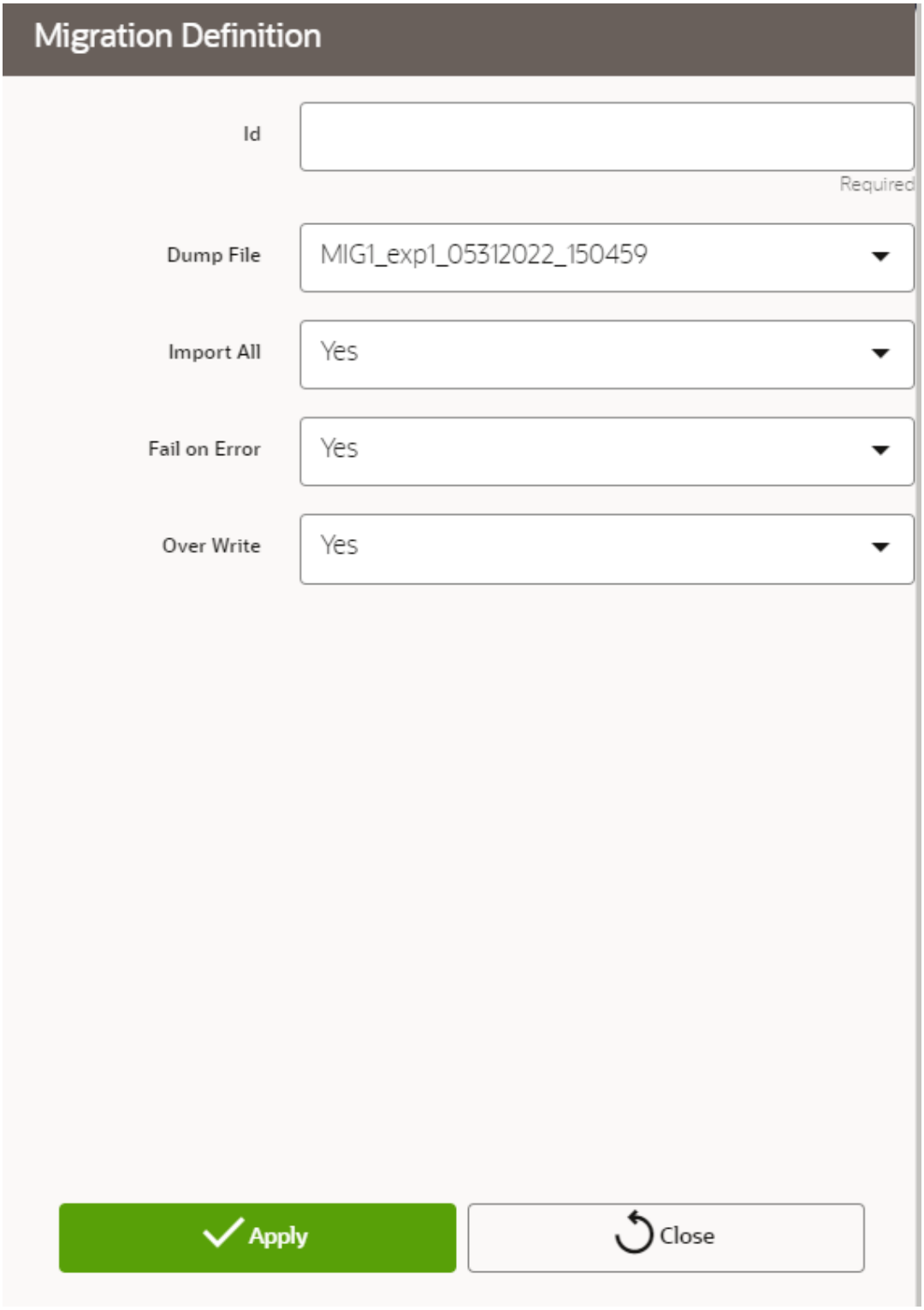
You can create Import Object definitions for migration for the following object types.

- Batch Group
- Threshold
- PMF Process
- Batch
- Pipeline
- Schedule
- Job

To create a definition for import of objects to be migrated, do as follows:

1. Click **Add** in the **Object Import Summary**page to display the **Migration Definition window**.

Figure 8-6 Migration Definition window



2. In the **Migration Definition window**, enter the details for the following:
   - **ID**: Enter a valid name for the new migration import definition. This is a unique identifier.
   - **Dump File**: Select the dump file to be utilized for creating the Migration Import Definition.

3. **Import All**: Select an option to import the nodes that are associated with the selected object type.

   You can edit this option if required, in the **Object Selection**page.

Figure 8-7 Object Selection page

   - **Yes** - Imports all the nodes that are included in the dump file.
   - **No** - Imports only those nodes that you can select in the Object Selection page.

4. **Fail on Error**: Select an option to proceed with the definition creation in case of an error.

   You can edit this option if required, in the **Object Selection**page.

   - **Yes** - Stops the creation process, if error is generated.
   - **No** - Creates the import definition even when error is generated. The node with the error is not included in the object creation.

5. **Overwrite**: Select an option to overwrite the existing definition file.

   You can edit this option if required, in the **Object Selection**page.

   - **Yes** - Replaces the existing Import definition.
   - **No** - Creates a new Import definition.

6. Click **Apply** to save the details.

   The **Object Selection** page is displayed.

7. In the **Object Selection** page, click **Add Members**, to add objects associated with specific object types.

8. In the **Object Selection** page, select the required **Object Type** from the **Object Types** drop-down list.

   The object types listed are, Schedule, Batch group, Batch, Pipeline, Threshold, Job and PMF process. For more information about the object types, refer Migration Object Types.

   You can also enter the first few letters in **Search** to add a particular object from a selected Object type.

   The list of objects is displayed.

9. Select the objects to be added to the Migrate Definition.

   You can select the objects only if you select **No** for **Import All**.

   The selected objects are added to the respective object type branch.

10. Click **Save** to create the Migrate Definition.

    A confirmation message is displayed, when the definition is saved successfully. The new migration definition is listed in the Object Import Summary page and the status is set to **Saved**.

> **Note:**
>
> If the migration definition object is not created successfully and the status is set to **Failed**, contact My Oracle Support of more information.

8.1.2.3 View Migration Objects to be Imported

You can view and edit the migration objects from the Object Import Summary Page.

To view the list of imported migration objects associated with a migration definition, highlight the migration definition and click **Menu**button.

The following options are displayed.

Table 8-3 Viewing Migration Import Objects

Field	Description
View Log	View the migration log details of the selected Migration Definition. For more information, refer to View Log of Migration Import Objects.
View	View the Object Details for a specific Migration Definition. Click an object to view more details. For more information, refer to Viewing Import Migration Object Details.
Export	Click **Export** to initiate Object Migration (Import) for a specific Migration Definition. When the migration is completed, the status will change from Saved to Success. For more information, refer to Importing Migration Objects.
Edit	Click **Edit** to view and edit the objects linked to a Migration Definition. For more information, refer to Editing Migration Import Definitions.

8.1.2.3.1 View Log of Migration Objects Imported

To view the log information of a record of the definition for import of objects to be migrated, follow these steps. The view log facilitates you to view the log information of the definition for import of objects to be migrated with its status.

1. Highlight the migration definition and click **Menu** button and select **View Log**.

   The **View Log** page is displayed.

   The export migration status with the following details is displayed.

Table 8-4 View Log of Migration Import Objects

Field	Description
Object Migration ID	The migration ID associated with the export object.

Table 8-4 (Cont.) View Log of Migration Import Objects

Field	Description
Object Type	The object type of the export object.
Object Code	The object code associated with the export object.
Creation Date	The date of creation of the export object.
Created By	The User Id of the User who created the export object.
Status	The migration status of the export object. • Success - Indicates that the export migration was completed successfully. • Failed - Indicates that the export migration did not complete

> **Note:**
>
> The View Log Page for a migration object with status Saved will be empty

2. Click **OK** to close the window.

8.1.2.3.2 View Details of Imported Migration Objects

To view the list of objects added to a specific Migration Definition:

1. Highlight the migration definition and click **Menu**button and select **View**.

   The list of migration objects added to the definition is displayed.

2. Double-click an object to view the object attribute details.

8.1.2.4 Import Migration Objects

To export the list of objects added to a specific Migration Definition:

- Highlight the migration definition and click **Menu**button and select **Export**.

  After you export, the following types of status are displayed:

  - **Success**- When the object is migrated successfully.
  - **Failed**- When the object migration didn't complete.

> **Note:**
>
> Object Migration (export) facilitates you to export a set of objects within the same setup or across different setups.

> **Note:**
>
> Currently, if the user give production workspace a name during the import, the utility migrates the dump to the production workspace as well. This should be restricted, as the models in sandbox may or may not be promoted and the invalid status will be shown in the production.

8.1.2.5 Edit Imported Migration Definitions

You can edit the migration export objects that are not exported and their status is Saved or Failed. If the object is already exported and the status is set to Success, you cannot edit the object details.

To edit a record of the definition of export of objects to be migrated, follow these steps. You can add more objects to export or remove existing objects.

1. Highlight the migration definition and click **Menu** button and select **Edit**.

   The **Object Selection** window is displayed.
2. Update the required details.
3. In the **Object Selection** window, select the required **Object Type** from the **Object Types**drop-down list.

   The object types listed for the System Configuration tab are, Batch_Group, PMF_Process, Batch and Schedule. For more information about the object types, refer Migration Object Types.

   You can also enter the first few letters in **Search** to add a particular object from a selected Object type.

   The list of objects is displayed.
4. Select the objects to be added to the Migrate Definition or to be deleted from the Migration Definition.

   The selected objects are added/deleted to the respective object type branch.
5. Click **Save** to edit the Migrate Definition.

   A confirmation message is displayed, when the definition is saved successfully.

   The edited migration definition is listed in the Object Export Summary page and the status is set to Saved.
6. Click **Save** to update the changes.

8.1.2.6 Delete Imported Migration Object Definitions

You can delete only a record that is set to Saved or Failed status and not a record that is in Success status.

To delete a migrate export object definition:

1. Highlight the record to be deleted and click the **Delete**button.
2. Click **Yes**to confirm and proceed with the deletion.

8.1.3 Prerequisites to Migrate Objects

In order to migrate the objects, the users must be mapped to the **Object Migration Admin Group**.

> **Note:**
>
> The Identity Administrator Group Users cannot migrate objects if they are not mapped to the **Object Migration Admin Group**.

8.1.4 Migration Object Types

You can migrate (import/export) the following Object Types by clicking the Object Migration icon on the left menu.

- Schedule
- Batch
- Batch Group
- Pipeline
- Threshold
- Job
- PMF_Process
- Roles
- Groups

8.1.4.1 Schedule

Schedule provides the instruction to schedule the execution of defined processes. When a schedule is migrated, the associated batch is also migrated.

8.1.4.2 Batch

A batch is a group of jobs that are scheduled to automatically execute at a preset interval of time, without any user's intervention. When a batch is migrated, the batch and the associated pipeline information are also migrated.

8.1.4.3 Batch Group

A set of individual batches are consolidated to form a single Batch Group. When we migrate a Batch Group all the batches, tasks and pipeline information associated with that Batch Group are also migrated.

8.1.4.4 Pipeline

A pipeline is an embedded data processing engine that runs inside the application to filter, transform, and migrate data on-the-fly. Pipelines are a set of data processing elements called widgets connected in series, where the output of one widget is the input to the next element.

8.1.4.5 Threshold

The threshold limit associated with set variables values for scenarios in FCCM Cloud Service. These threshold values are set when scenarios are created or installed and can be changed, if required.

8.1.4.6 Job

Jobs provide set of instructions to execute Workflow Pipelines, based on the set threshold values

8.1.4.7 PMF_Process

PMF_Processes are defined to sequence the Workflow Pipelines the applications, and to design the artifacts that participate in the Pipelines, to implement the Pipelines. Export of PMF_Process will take care of dependent metadata, such as data fields, transition rules associated to the PMF process, that are defined in PMF.

8.1.4.8 Roles

Roles are used to map functions to a defined set of groups to ensure user access system security.

8.1.4.9 Groups

Groups are used to map Roles. Specific User Groups can perform only set of functions associated with that group.

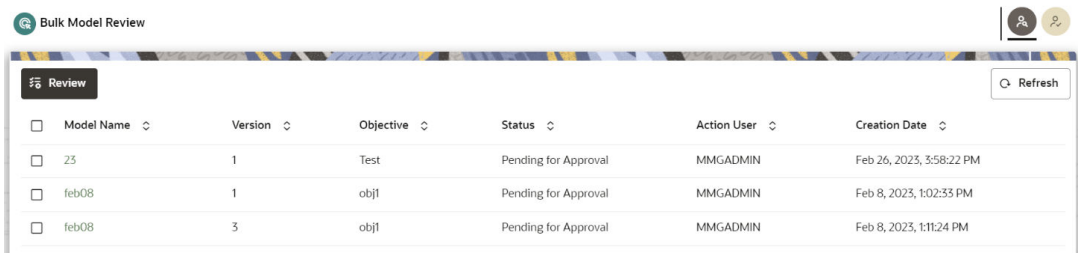
8.2 Model Actions

The Model Actions window allows you to view the list of models for which the Workflow action has been taken which requires review or approve across the workspace.

For example, when any Model User sends and Model for review or approval, the user receives the bulk Models in Model Actions window for review or approval. This helps when lot of models are received for review before deployment or some actions need to be performed on many models, such as making multiple models champion locally or globally, and so on.

1. Click **Launch Workspace** next to the corresponding Workspace to display the **Dashboard** window with application configuration and model creation menu.
2. Click Model Actions to view the list models pending an action such as approval.

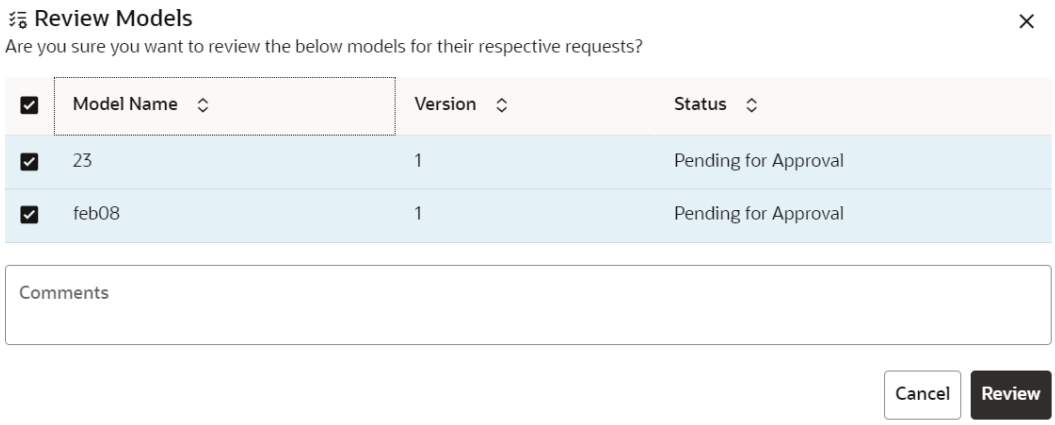
Figure 8-8 Review Models screen



If you are a reviewer, Review Models screen is displayed and if you are an approver, Approve Models screen is displayed. This can be toggled in the header of the screen using Reviewer and Approver icon.

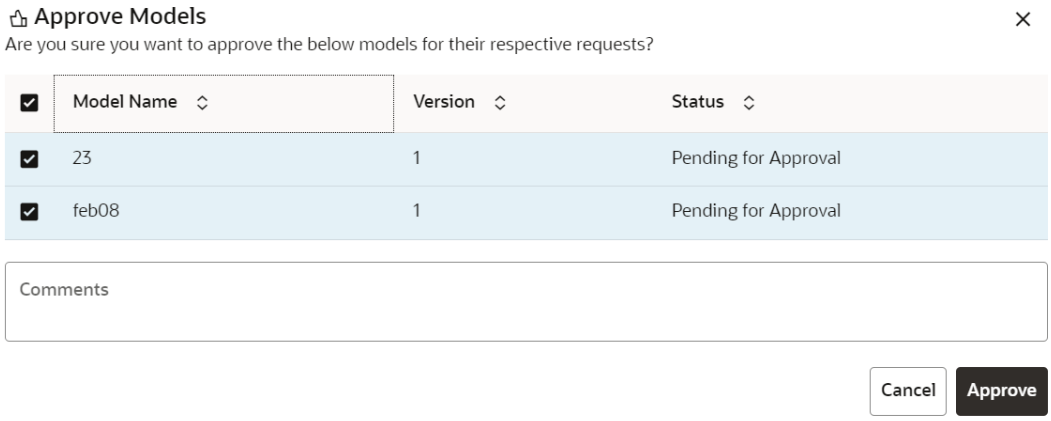
3. To review the model, follow these steps:
   a. Select the Model using the corresponding check box and click **Review**.
   b. Enter **Comments**, if any, and click **Review**.

Figure 8-9 Review Models window



4. To approve or reject the models, follow these steps:
   a. Select the Model using the corresponding check box and click **Approve** or **Reject**.
   b. Enter the comments in Comments field of the **Approve Models** window or **Reject Models** window and click **Approve** or **Reject**.

Figure 8-10 Approve Models screen



8.3 Audit Trail

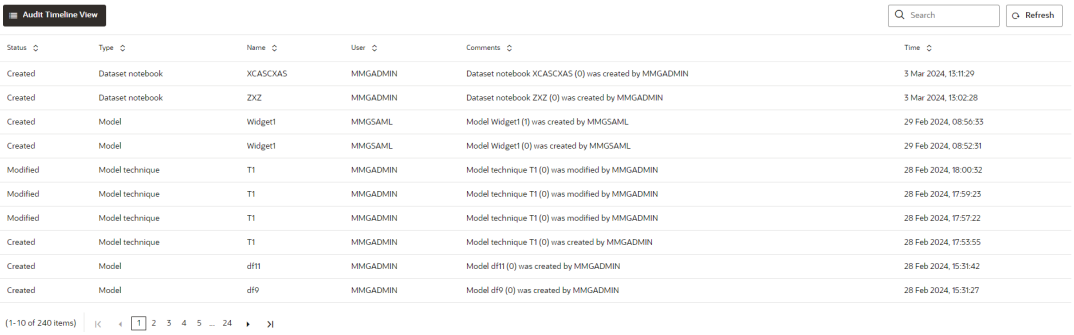
At any time, you can audit the models from the Audit Trail window. The Audit Trail window provides the complete details of model. This shows the information such as, when Model was created, who created the Model, workflow of Model, when this Model became champion or deployed, and so on.

The sequence of actions performed in the model lifecycle is listed in the table view and the timeline view (graphical representation).

To audit models:

1. Click **Launch Workspace** next to the corresponding Workspace to display the **Dashboard** window with application configuration and model creation menu.
2. In the Mega menu, click **More > Audit Trail** to view the time of the various actions performed on the model in the **Audit Trail** window.

Figure 8-11 Audit Trail window

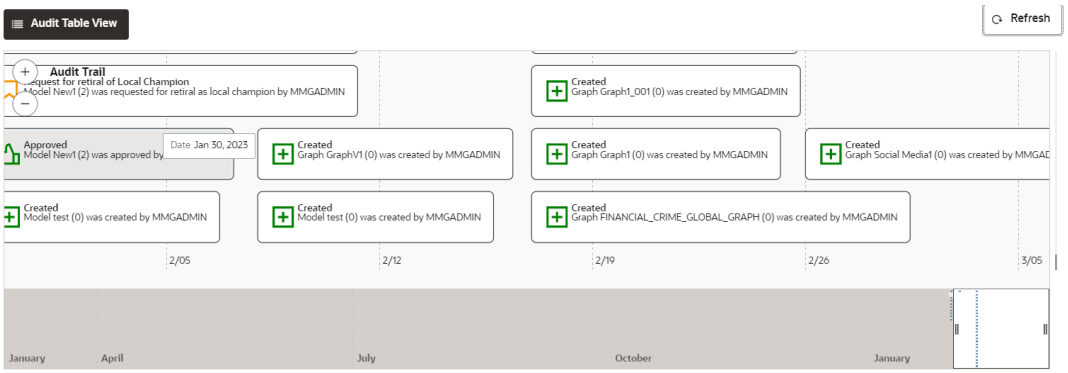


3. Click for Timeline View.

   Click to switch to the regular view. Click to refresh the Audit Trail window.

   You can search the models by entering the name in the search box.

Figure 8-12 Audit Trail Window Timeline View with Horizontal Time Axis



9
Using Compliance Studio Workspace

The Compliance Studio Workspace provides core capabilities that are not model or sandbox-specific.

9.1 Compliance Studio Global Graph

Compliance Studio provides a pre-defined Financial Crime graph that can be pre-loaded for most Financial Crime use cases, including Investigation Toolkit.

It also provides an intuitive way for creating graphs used in notebooks, where you can load graphs from external sources or create custom graphs. Compliance Studio includes the PGX, a fast, parallel, in-memory graph analytic framework developed by Oracle. Using PGX, you can load multiple graphs into a notebook and create PGQL queries against different graphs. The result obtained from running a paragraph in a notebook can be used as input to other paragraphs in the notebook. The results of analytics algorithms are stored as the graph's nodes and edges properties. Pattern matching can then be used against these properties.

The Financial Crime Graph is created by an ETL process that brings in data from our Financial Crime Data Model and freely available data from the International Consortium of Investigative Journalists by default but can be easily configured to include other data sources. As part of the ETL process, matching is run on nodes within the graph and creates similarity edges that link nodes based on name, address, and other attributes. For more information, see the Creating Match/Merge Ruleset section.

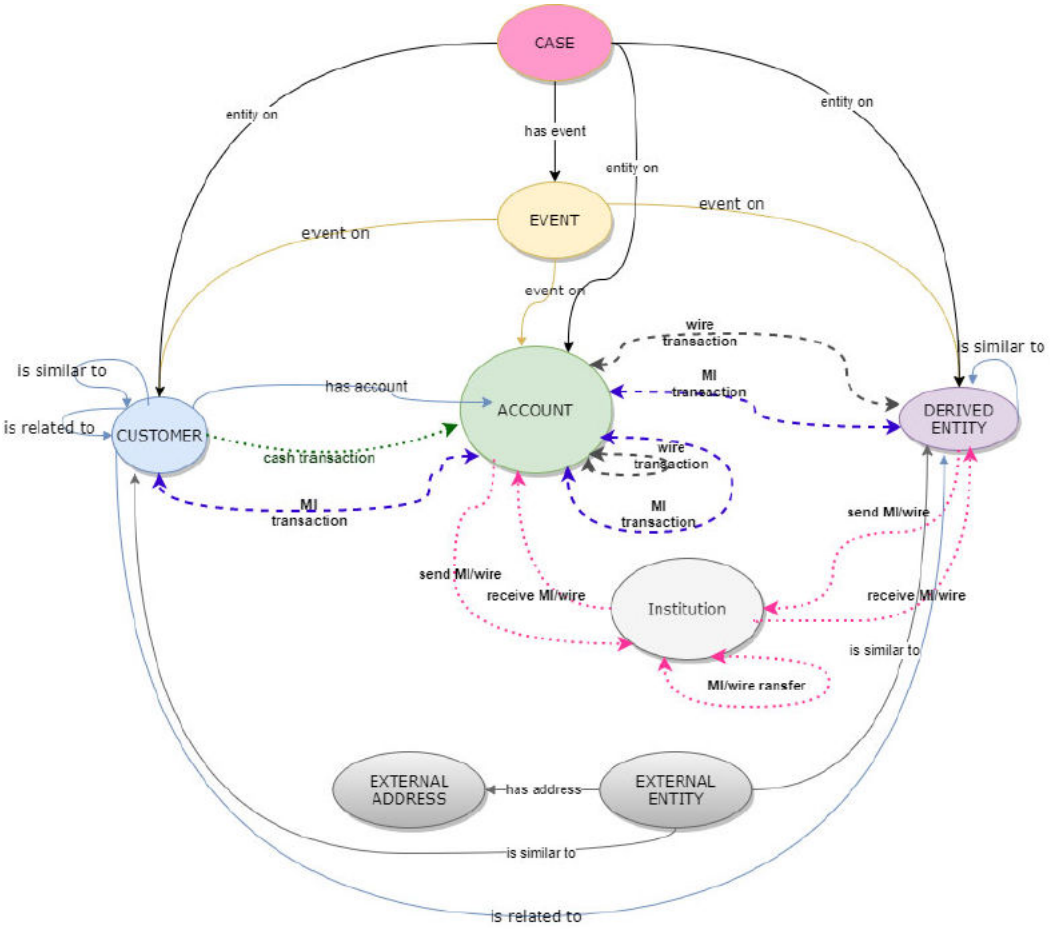
ETL can be run in two ways:

1. Through graph pipelines, for more information see Graphs section.
2. To configure and execute the ETL, follow these steps:

> **Note:**
>
> This step is deprecated in the current release and will be removed in the future release.

   a. For more information, see the **Configure ETL** section in the OFS Compliance Studio Administration and Configuration Guide.
   b. Through legacy ETL on hive, for more information see the **Execute ETL** section in the OFS Compliance Studio Administration and Configuration Guide.

Figure 9-1 Oracle Financial Crime Graph Model



> **Note:**
>
> The Case node in this Financial Crime Graph Model is loaded only when loading the FCDM data from Enterprise Case Management (ECM). The graph includes “CASE” nodes and “has event” edges when data is loaded from ECM.

9.1.1 Load Data into ICIJ Tables

To load data into ICIJ tables, see the **Load Data into ICIJ Tables** section in the OFS Compliance Studio Administration and Configuration Guide.

After installing the Compliance Studio, you need to run the script. For more details, see the **Importing OOB Graph Definition and other related Metadata** section in the OFS Compliance Studio Installation Guide.

9.1.2 Compliance Studio OOB Graph

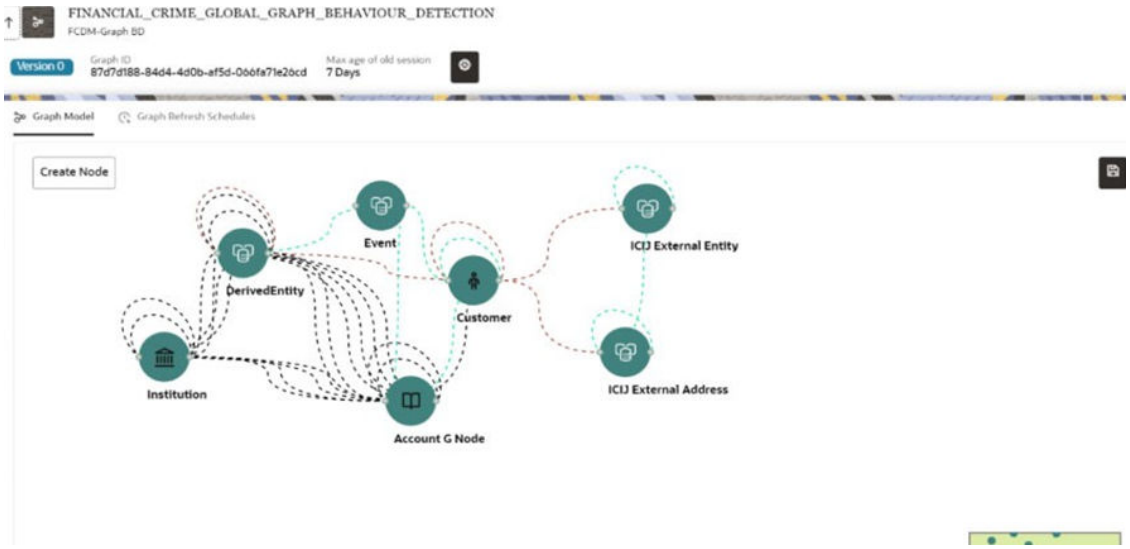
Compliance Studio provides a pre-defined Financial Crime graph that can be pre-loaded for most Financial Crime use cases, including Investigation Toolkit.

It also provides an intuitive way to create graphs used in notebooks, where you can load graphs from external sources or create custom graphs that are then loaded into PGX. PGX is a fast, parallel, in-memory graph analytic framework developed by Oracle. Using PGX, you can load multiple graphs into a notebook and create PGQL queries against different graphs. Following are the pre-defined global financial crime graph for Compliance Studio:

- FINANCIAL_CRIME_GLOBAL_GRAPH
- FINANCIAL_CRIME_GLOBAL_GRAPH_BEHAVIOUR_DETECTION

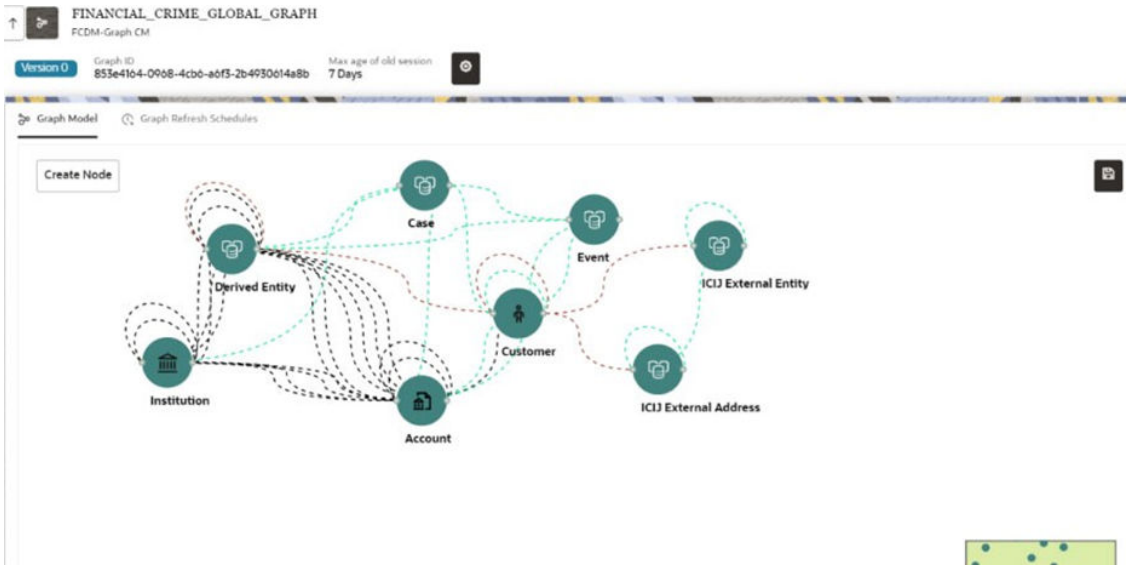
The following figure shows the out-of-the-box Global Graph Model for BD.

Figure 9-2 OOB Global Graph Model for BD



The following figure shows the out-of-the-box Global Graph Model for ECM.

Figure 9-3 OOB Global Graph Model for ECM



9.1.3 Pre-configured Graph Model (FCGM)

This topic provides information about nodes, edges, similarity edges and indexes for BD and ECM graphs.

List of Nodes for BD

Table 9-1 Node Details for BD

Name	Description	Pipeline	Data Pipeline Name	Is Active
Customer	Customer	Data Pipeline	Customer	Yes
Account G Node	Account	Data Pipeline	AccountNodeGDP	Yes
DerivedEntity	Derived Entities	Data Pipeline	DerivedEntityDP	Yes
Institution	Institutions	Data Pipeline	Institution	Yes
Event	Events	Data Pipeline	Event	Yes
ICIJ External Entity	ICIJ External Entity	Data Pipeline	ICIJ External Entity	Yes
ICIJ External Address	ICIJ External Address	Data Pipeline	ICIJ External Address	Yes

List of Nodes for ECM

Table 9-2 Node Details for ECM

Name	Description	Pipeline	Data Pipeline Name	Is Active
Customer	Customer	Data Pipeline	Customer CM	Yes
Account	Account	Data Pipeline	Account CM	Yes
Event	Events	Data Pipeline	Event CM	Yes
DerivedEntity	Derived Entities	Data Pipeline	Derived Entity CM	Yes
Institution	Institutions	Data Pipeline	Institution CM	Yes
ICIJ External Entity	ICIJ External Entity	Data Pipeline	ICIJ External Entity CM	Yes
ICIJ External Address	ICIJ External Address	Data Pipeline	ICIJ External Address CM	Yes
Case	Case	Data Pipeline	Case CM	Yes

List of Edges for BD

Table 9-3 Edge Details for BD

Name	Description	Pipeline	Data Pipeline Name	Is Active
Cash Trxn Cust To Acct Edge	Cash Transactions between Customer and Account	Data Pipeline	Cash Trxn Cust To Acct Edge	Yes
Cust Has Account	Customer Has Account(s)	Data Pipeline	Cust Has Acct EdgeDP	Yes

Table 9-3 (Cont.) Edge Details for BD

Name	Description	Pipeline	Data Pipeline Name	Is Active
Cust Is Related To Cust Edge	Customer is Related to any other Customers	Data Pipeline	Cust Is Related To Cust Edge	Yes
Cash Trxn EE Id	Cash Transactions between Account and Derived Entity	Data Pipeline	Cash Trxn EE Id	No
Cash Trxn EE Name	Cash Transactions between Account and Derived Entity	Data Pipeline	Cash Trxn for EE Name	No
End To End Mi EE to Acct	End to End Mi Transactions from Derived Entity to Account	Data Pipeline	End To End Mi EE to Acct	No
End to End Wire Acct to EE	End to End Wire Transactions from Account to External Entity	Data Pipeline	End to End Wire Acct to EE	No
End To End Mi Acct to EE	End to End Mi Transactions from Account to External Entity	Data Pipeline	End To End Mi Acct to EE	No
End to End Wire EE Acct	End to End Wire Transactions from External Entity to Account	Data Pipeline	End to End Wire EE to Acct	No
End to End Mi EE to EE	End to End Mi Transactions from External Entity to External Entity	Data Pipeline	End to End Mi EE to EE	No
End to End Wire EE to EE	End to End Wire Transactions from External Entity to External Entity	Data Pipeline	End to End Wire EE to EE	No
End to End MI Acct to Acct	End to End Mi Transactions from Account to Account	Data Pipeline	End to End MI Acct to Acct	No
End to End Wire Acct to Acct	End to End Wire Transactions from Account to Account	Data Pipeline	End to End Wire Acct to Acct	No
Event On Cust	Any events on Customers	Data Pipeline	Event On Cust	No
Event On EE	Any Events on External Entity	Data Pipeline	Event On EE	No
Event On Acct	Any events on Account	Data Pipeline	Event On Acct	No
Send Wire Acct to Instn	Sending Wire Transaction from Acct to Institution	Data Pipeline	Send Wire Acct to Instn	No

Table 9-3 (Cont.) Edge Details for BD

Name	Description	Pipeline	Data Pipeline Name	Is Active
Send Wire EE to Instn	Sending Wire Transaction from Derived Entity to Institution	Data Pipeline	Send Wire EE to Instn	No
Send Mi Acct to Instn Edge	Sending Mi Transaction from Account to Institution	Data Pipeline	Send Mi Acct to Instn	No
Send Mi EE to Instn	Sending Mi Transaction from Derived Entity to Institution	Data Pipeline	Send Mi EE to Instn	No
Receive Mi Instn to Acct	Receiving Mi Transaction from Institution to Account	Data Pipeline	Receive Mi Instn to Acct	No
Receive Mi Instn to DE	Receiving Mi Transaction from Institution to Derived entity	Data Pipeline	Receive Mi Instn to EE	No
Receive Wire Instn to Acct	Receiving Wire Transaction from Institution to Account	Data Pipeline	Receive Wire Instn to Acct	No
Receive Wire Instn to DE	Receiving Wire Transaction from Institution to Derived Entity	Data Pipeline	Receive Wire Instn to EE	No
Transfer Mi Instn to Instn	Transfer an Intermediary Mi Transaction from Institution to Institution	Data Pipeline	Transfer Mi Instn to Instn	No
Transfer Wire Instn to Instn	Transfer an Intermediary Wire Transaction from Institution to Institution	Data Pipeline	Transfer Wire Instn to Instn	No
ICIJ EE Is Related To EE	ICIJ External Entity is Related To External Entity	Data Pipeline	ICIJ EE Is Related To EE	No
ICIJ Address Of EE To EA	ICIJ Address Of External Entity To External Account	Data Pipeline	ICIJ Address Of EE To EA	No
ICIJ EA Is Related To EA	ICIJ External Account Is Related To External Entity	Data Pipeline	ICIJ EA Is Related To EA	No

List of Edges for ECM

Table 9-4 Edge Details for ECM

Name	Description	Pipeline	Data Pipeline Name	Is Active
Cust Has Account	Customer Has Account(s)	Data Pipeline	Cust Has Acct CM	Yes
Cust Is Related To Cust	Customer is Relatedany other Customers	Data Pipeline	Cust Is Related To Cust CM	Yes
Cash Trxn Cust To Acct	Cash Transactions between Customer and Account	Data Pipeline	Cash Trxn Cust To Acct CM	Yes
Cash Trxn EE Id	Cash Transactions between Account and Derived Entity	Data Pipeline	Cash Trxn EE Id CM	No
Cash Trxn EE Name	Cash Transactions between Account and Derived Entity	Data Pipeline	Cash Trxn EE Name CM	No
End To End Mi Trxn Acct To Acct	End to End Mi Transactions from Derived Entity to Account	Data Pipeline	End To End Mi Trxn Acct To Acct CM	No
End To End Wire Trxn Acct To Acct	End to End Wire Transactions from Account to External Entity	Data Pipeline	End To End Wire Trxn Acct To Acct CM	No
End To End Mi Trxn EE To EE	End to End Mi Transactions from External Entity to External Entity	Data Pipeline	End To End Mi Trxn EE To EE CM	No
End To End Wire Trxn EE To EE	End to End Wire Transactions from External Entity to External Entity	Data Pipeline	End To End Wire Trxn EE To EE CM	No
End To End Mi Trxn EE To Acct	End to End Mi Transactions from External Entity to Account	Data Pipeline	End To End Mi Trxn EE To Acct CM	No
End To End Wire Trxn EE To Acct	End to End Wire Transactions from External Entity to Account	Data Pipeline	End To End Wire Trxn EE To Acct CM	No
End To End Mi Trxn Acct To EE	End to End Mi Transactions from Account to External Entity	Data Pipeline	End To End Mi Trxn Acct To EE CM	No
End To End Wire Trxn Acct To EE	End to End Wire Transactions from Account to External Entity	Data Pipeline	End To End Wire Trxn Acct To EE CM	No

Table 9-4 (Cont.) Edge Details for ECM

Name	Description	Pipeline	Data Pipeline Name	Is Active
Send Mi Trxn Acct To Instn	Sending Mi Transaction from Account to Institution	Data Pipeline	Send Mi Trxn Acct To Instn CM	No
Send Mi Trxn EE To Instn	Sending Wire Transaction from Acct to Institution	Data Pipeline	Send Mi Trxn EE To Instn CM	No
Send Wire Trxn EE to Instn	Sending Wire Transaction from Derived Entity to Institution	Data Pipeline	Send Wire Trxn EE To Instn CM	No
Receive Mi Trxn From Instn To Acct	Receiving Mi Transaction from Institution to Account	Data Pipeline	Receive Mi Trxn From Instn To Acct CM	No
Receive Wire Trxn From Instn To Acct	Receiving Wire Transaction from Institution to Account	Data Pipeline	Receive Wire Trxn From Instn To Acct	No
Receive Mi Trxn From Instn To EE	Receiving Mi Transaction from Institution to Derived entity	Data Pipeline	Receive Mi Trxn From Instn To EE CM	No
Receive Wire Trxn From Instn To EE	Receiving Wire Transaction from Institution to Derived Entity	Data Pipeline	Receive Wire Trxn From Instn To EE CM	No
Transfer Mi Trxn From Instn To Instn	Transfer an Intermediary Mi Transaction from Institution to Institution	Data Pipeline	Transfer Mi Trxn from Instn To Instn CM	No
Transfer Wire Trxn From Instn To Instn	Transfer an Intermediary Wire Transaction from Institution to Institution	Data Pipeline	Transfer Wire Trxn From Instn To Instn CM	No
Event On Cust	Any events on Customers	Data Pipeline	Event On Cust CM	No
Event On EE	Any Events on External Entity	Data Pipeline	Event On EE CM	No
Event On Acct	Any events on Account	Data Pipeline	Event On Acct CM	No
ICIJ EE Is Related To EE	ICIJ External Entity is Related To External Entity	Data Pipeline	ICIJ EE Is Related To EE CM	No
ICIJ Address Of EE To EA	ICIJ Address Of External Entity To External Account	Data Pipeline	ICIJ Address Of EE To EA CM	No

Table 9-4 (Cont.) Edge Details for ECM

Name	Description	Pipeline	Data Pipeline Name	Is Active
ICIJ EA Is	ICIJ External Account Is Related To External Entity	Data Pipeline	ICIJ EA Is Related To EA CM	No
Related To EA				
Case On EE	Case On External Entity	Data Pipeline	Case On EE CM	No
Case On Acct	Case On Account	Data Pipeline	Case On Acct CM	No
Case On Cust	Case On Customer	Data Pipeline	Case On Cust CM	No
Case On	Case On Institution	Data Pipeline	Case On Institution CM	No
Institution	Case Has Event	Data Pipeline	Case Has Event CM	No
Case Has Event	Customer Has Account(s)	Data Pipeline	Cust Has Acct CM	No

List of Similarity Edges for BD

Table 9-5 Similarity Edge Details for BD

Name	Description	Pipeline	Match Ruleset Name	Is Active
Cust is similar to Cust	Is Customer is similar to customer	Data Pipeline	Customer to Customer Match BDG	Yes
Cust is similar to Derived Entity	Is Customer similar to Derived Entity	Data Pipeline	Customer to Derived Entity Match BDG	Yes
Derived Entity is similar to Derived Entity	Is Derived Entity is similar to Derived Entity	Data Pipeline	Derived Entity to Derived Entity Match BDG	Yes
Cust is similar to External Entity ICIJ	Customer is similar to External Entity ICIJ	Data Pipeline	Customer to External Entity ICIJ Match BDG	Yes
Cust is similar to External Address ICIJ	Customer is similar to External Address ICIJ	Data Pipeline	Customer to External Address ICIJ Match BDG	Yes

List of Similarity Edges for ECM

Table 9-6 Similarity Edge Details for ECM

Name	Description	Pipeline	Match Ruleset Name	Is Active
Cust is similar to Cust	Is Customer is similar to customer	Data Pipeline	Customer to Customer Match CM	Yes

Table 9-6 (Cont.) Similarity Edge Details for ECM

Name	Description	Pipeline	Match Ruleset Name	Is Active
Cust is similar to Derived Entity	Is Customer similar to Derived Entity	Data Pipeline	Customer to Derived Entity Match CM	Yes
Derived Entity is similar to Derived Entity	Is Derived Entity is similar to Derived Entity	Data Pipeline	Derived Entity to Derived Entity Match CM	Yes
Cust is similar to External Entity ICIJ	Customer is similar to External Entity ICIJ	Data Pipeline	Customer to External Entity ICIJ Match CM	Yes
Cust is similar to External Address ICIJ	Customer is similar to External Address ICIJ	Data Pipeline	Customer to External Address ICIJ Match CM	Yes

List of Indexes for BD

This table describes the Out-of-the-box indexes for the match rules that are defined as similarity edge (Table 9-3) in the FCGM.

Table 9-7 Pre-seeded Indexes for BD

Index Name	Name	Display Name	Type
Customer Graph	address	Address	Text
	alias	Alias	Text
	business_domain	Business Domain	Text
	city	City	Text
	concat_name	Concatenated Name	Text
	country	Country	Text
	family_name	Last Name	Text
	given_name	First Name	Text
	industry	Industry	Text
	middle_name	Middle Name	Text
	name	Name	Text
	state	State	Text
	dob	DOB	Date
	d_date	Date	Date
	email	Email	Text
	entity_type	Entity Type	Text
	industry	Industry	Text
	jurisdiction	Jurisdiction	Text
	label	Label	Text
	naics_code	Naics Code	Text
	original_id	Original ID	Text
	phone	Phone	Text
	risk	Risk	Number
	source	Source	Text

Table 9-7 (Cont.) Pre-seeded Indexes for BD

Index Name	Name	Display Name	Type
	status	Status	Text
	tax_id	Tax ID	Text
Derived Entity Graph	label	Label	Text
	source	Source	Text
	original_id	Original ID	Text
	jurisdiction	Jurisdiction	Text
	name	Name	Text
	risk	Risk	Number
	d_date	Date	Date
	business_domain	Business Domain	Text
	address	Address	Text
	city	City	Text
	state	State	Text
	country	Country	Text
	status	Status	Text
	external_id	External Id	Text
External Address Graph	address	Address	Text
	jurisdiction	Jurisdiction	Text
	country	Country	Text
	source	Source	Text
	label	Label	Text
External Entity Graph	label	Label	Text
	source	Source	Text
	original_id	Original ID	Text
	jurisdiction	Jurisdiction	Text
	name	Name	Text
	address	Address	Text
	city	City	Text
	state	State	Text
	country	Country	Text
	business_domain	Business Domain	Text
Event Graph	jurisdiction	Jurisdiction	Text
	name	Name	Text
	risk	Risk	Number
	d_date	Date	Date
	status	Status	Text
	reason	Reason	Text
	rule_name	Rule Name	Text
	business_domain	Business Domain	Text
	label	Label	Text
	source	Source	Text
	original_id	Original ID	Text
Institution Graph	label	Label	Text

Table 9-7 (Cont.) Pre-seeded Indexes for BD

Index Name	Name	Display Name	Type
	risk	Risk	Number
	country	Country	Text
	state	State	Text
	d_date	Date	Date
	jurisdiction	Jurisdiction	Text
	original_id	Original ID	Text
	source	Source	Text
	name	Name	Text
Account Graph	tax_id	Tax ID	Text
	d_date	Date	Date
	address	Address	Text
	city	City	Text
	state	State	Text
	country	Country	Text
	status	Status	Text
	risk	Risk	Number
	business_domain	Business Domain	Text
	source	Source	Text
	original_id	Orignal ID	Text
	jurisdiction	Jurisdiction	Text
	name	Name	Text
	entity_type	Entity Type	Text
	label	Label	Text

List of Indexes for ECM

This table describes the Out-of-the-box indexes for the match rules that are defined as similarity edge (Table 9-4) in the FCGM.

Table 9-8 Pre-seeded Indexes for ECM

Index Name	Name	Display Name	Type
Customer_Graph_CM	address	Address	Text
	alias	Alias	Text
	business_domain	Business Domain	Text
	city	City	Text
	concat_name	Concatenated Name	Text
	country	Country	Text
	family_name	Last Name	Text
	given_name	First Name	Text
	industry	Industry	Text
	middle_name	Middle Name	Text
	name	Name	Text
	state	State	Text

Table 9-8 (Cont.) Pre-seeded Indexes for ECM

Index Name	Name	Display Name	Type
	dob	DOB	Date
	d_date	Date	Date
	email	Email	Text
	entity_type	Entity Type	Text
	industry	Industry	Text
	jurisdiction	Jurisdiction	Text
	label	Label	Text
	naics_code	Naics Code	Text
	original_id	Original ID	Text
	phone	Phone	Text
	risk	Risk	Number
	source	Source	Text
	status	Status	Text
	tax_id	Tax ID	Text
Derived_Entity_Graph_CM	label	Label	Text
	source	Source	Text
	original_id	Original ID	Text
	jurisdiction	Jurisdiction	Text
	name	Name	Text
	risk	Risk	Number
	d_date	Date	Date
	business_domain	Business Domain	Text
	address	Address	Text
	city	City	Text
	state	State	Text
	country	Country	Text
	status	Status	Text
	external_id	External Id	Text
External_Address_Graph_CM	address	Address	Text
	jurisdiction	Jurisdiction	Text
	country	Country	Text
	source	Source	Text
	label	Label	Text
External_Entity_Graph_CM	label	Label	Text
	source	Source	Text
	original_id	Original ID	Text
	jurisdiction	Jurisdiction	Text
	name	Name	Text
	address	Address	Text
	city	City	Text
	state	State	Text
	country	Country	Text
	business_domain	Business Domain	Text

Table 9-8 (Cont.) Pre-seeded Indexes for ECM

Index Name	Name	Display Name	Type
Account_Graph_CM	tax_id	Tax ID	Text
	d_date	Date	Date
	address	Address	Text
	city	City	Text
	state	State	Text
	country	Country	Text
	status	Status	Text
	risk	Risk	Number
	business_domain	Business Domain	Text
	source	Source	Text
	original_id	Orignal ID	Text
	jurisdiction	Jurisdiction	Text
	name	Name	Text
	entity_type	Entity Type	Text
	label	Label	Text

Graph Model

The Oracle Financial Crime Graph Model serves as a window into the financial crimes data lake. It collates disparate data sets into an enterprise-wide global graph, enabling an entirely new set of financial crime use cases. The Graph model enables to accelerate financial crime investigation use cases.

For information on the node and edge properties of the Oracle Financial Crime Graph Model, see the Data Model Guide.

9.2 Entity Resolution

Compliance Studio allows firms to get unified views of both customers and external entities by consolidating duplicate representations of any customer or external entity into one consolidated Global party ID.

OFS Compliance Studio Entity Resolution is a configurable process that allows data to be matched and merged to create contextual links in the global graph or resolve relational party records to a global party record as part of ingestion. OFS Compliance Studio has pre-built configurations supporting matching (or linking) in the FCGM and resolving entities in CSA for data being loaded into Financial Services Data Foundation (FSDF).

Comparison for Delta Processing

The first time Entity Resolution runs, it operates on the full data set. This means the initial run will take longer than subsequent runs after the initial processing where deltas are calculated (regardless of whether full or delta data is populated in the input tables) so that matching happens only on new and changed records for improved performance.

Candidate Selection

Scoring on all pairs of records is not performant, so the Entity Resolution process first finds candidates with similar attributes and only scores on those pairs of records. Candidate Selection can either be run using Oracle OpenSearch or in the database using Oracle Text.

Matching

Matching rules are used to compare entities to identify pairs that refer to the same entity. It creates a probable link between entities by comparing the attributes of the entities.

For example, de-duplicating customers, resolving derived entities, or linking customers or derived entities to external data such as Panama papers or sanctions lists with different rules and thresholds.

For more information on scoring methods, see the OFS Compliance Studio Matching Guide.

For more information on how to create the Matching Rules, see the Creating Match/Merge Ruleset section.

Grouping

It is used to Group (entity Ids or Customer Ids) entities based on similarity links between them using matching rules and applying the merge rules on similarities. Once it is grouped, the system assigns the global party id to each Group.

> **Note:**
>
> Grouping is an automatic process. Grouping will be based on the match edges without any configuration.

Merge Rules

Merging rules are used to group multiple entities or customers into a single global party based on the merge ruleset.

For more information on how to create the Merging Rules, see the Creating Match/Merge Ruleset section.

Persisting

Records identified for merging will be collapsed into a single global party record, and a mapping from this global party record to the original party records will be created. Ongoing changes to the original party records may impact the global parties. For more details, see **Persisting the Data** section in the OFS Compliance Studio Administration and Configuration Guide.

Data Survival

When party records are identified for merging, a single output party record is created for the main or parent Dataset. Entity Resolution provides a mechanism to select the best view of the data from across the multiple-party records using attribute-by-attribute selection functions like Most Common or Maximum. It also provides a mechanism for transforming the child records stored in related tables, such as an address, email, or document ids.

Merge and Split Global Parties

Entity Resolution provides a mechanism to manually merge,split, create, and rearrange the entities for Global parties. Whenever there is a manual action (merge, split, create, rearrange) to the entities of a global party, the same data survival logic will be applied. See Using Merge and Split Global Entities section on how to perform the actions.

> **Note:**
>
> - When the records are not matched and not merged, they pass straight through and have a one-to-one mapping with the global party.
> - Where Entity has been resolved/unresolved, data origin is set to **EntRes** for all the records.
> - The Data Survival job cannot override the manual actions to a global party in batch mode.

For more information on configuring the rules for attribute survival, see the Data Survival Rules section.

For further details on the end-to-end Entity Resolution Process, see the Entity Resolution section in the OFS Compliance Studio Administration and Configuration Guide.

9.3 Creating Rulesets

Rulesets define similarity edges in the graph and match entities as part of the Entity Resolution process.

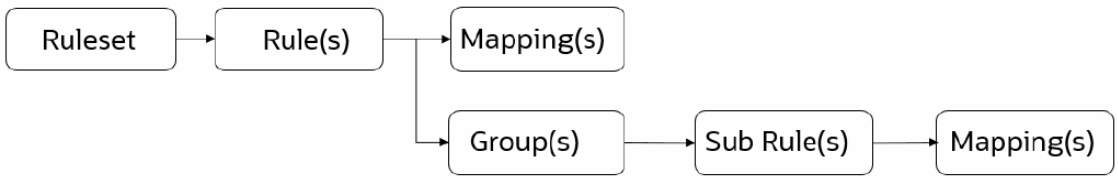
The Ruleset identifies the similarity between two entities (customer, account, etc.) and derives a match. A Ruleset is a set of rules applied to the defined source and target entities and compares the entities' attributes to derive a match.

Compliance Studio provides OOB rulesets; however, you can modify these rulesets or create your rulesets:

- Use Ruleset
- Create Ruleset

The process for creating Match and Merge Rule (s) is as follows:

Figure 9-4 Process for Creating Match and Merge Rulesets



> **Note:**
>
> - The sum of weightage for mappings and groups for each rule must be 1.
> - The sum of weightage for mappings for each sub-rule must be 1.

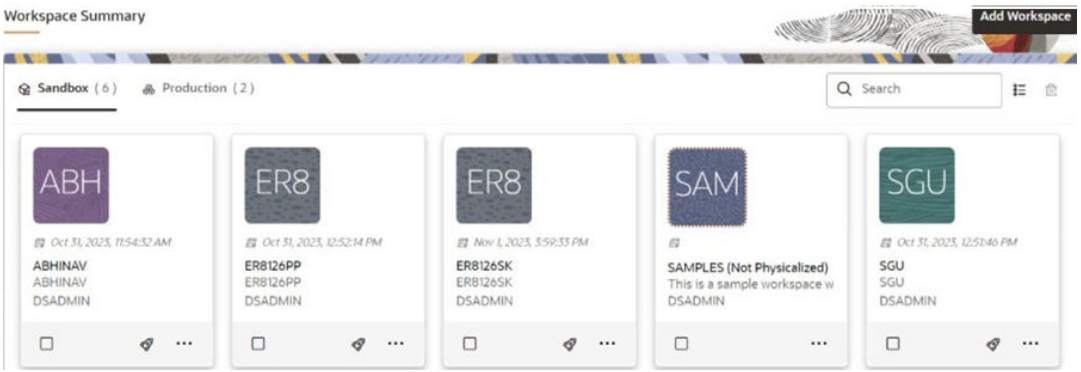
9.3.1 Using Match/Merge Ruleset

This topic uses to enable or disable the out-of-the-box rulesets.

You can also modify source or target entities and attributes for these rulesets and also remove rulesets from the existing list.
To use the existing ruleset, follow these steps:

1. Log in to the OFS Compliance Studio application. The Workspace Summary page is displayed.
2. On **Workspace Summary** page, click **Launch Workspace** to display the CS Production Workspace window with application configuration and model creation menu.

Figure 9-5 Workspace Summary



3. On **More** menu, Click **Match Rules/Merge Rules**. The Match Rules/Merge Ruleset window is displayed according to the selection.
4. On the details page, perform the following actions:
   a. To enable the Ruleset, select the Ruleset and move the toggle switch to the right.
   b. To add the Ruleset, click **Add**. For more information, see the Creating Match/Merge Ruleset section.
   c. To edit the Ruleset, click on the **Ruleset**, and the Ruleset page is displayed. On the Ruleset page, click **Edit**. The page displays Ruleset details.

      You can modify the Ruleset details except for the following fields:

      - Ruleset Name
      - Scoring Aggregation Type
      - Source Name
      - Target (It is applicable only for Matching)

   d. To delete the Ruleset(s), click **Delete**. The selected Ruleset(s) will be deleted from the list.

> **Note:**
>
> Fields marked with * are mandatory fields.

For more information on the fields in the Match/Merge Ruleset window, see the Creating Match/Merge Ruleset section.

9.3.2 Creating Match/Merge Ruleset

This topic describes the systematic instructions to create match and merge rulesets.

> **Note:**
>
> To create custom match rulesets, see the **Custom Rulesets for Matching** section in the OFS Compliance Studio Administration and Configuration Guide.

Each Ruleset comprises multiple rules. The Ruleset compares the attributes defined in the rules for the source entity with the target entity. For example, Customer to Customer, Customer to Derived Entity, Derived Entity to Derived Entity, etc.

Every entity has attributes such as email IDs, the date of birth, jurisdiction, etc. To derive a match and create a similarity edge on the graph or merge records in FSDF, you must apply the conditions for these attributes such as match type, scoring method, threshold score, and weightage.

To create a new Ruleset, follow these steps:

1. On the Ruleset page, click **Add**. The page displays as follows:
   - For Match Rules: Create Ruleset fields details
   - For Merge Rules: Create Merge Ruleset page details
2. Enter the ruleset field values as described in the following table.

Table 9-9 Create a Ruleset

Fields	Description
Ruleset Name	The name of the Ruleset. **Note**: Ruleset Name field accepts only 255 characters.
Scoring Aggregation Type	By default, the maximum scoring aggregation method is selected. This scoring aggregation considers the highest score obtained out of all the rules created for a ruleset.
Source	The source entity. For example, customer, account, address, etc.

Table 9-9 (Cont.) Create a Ruleset

Fields	Description
Source Filter	Source Filter icon is applicable for the Source field. You can filter the records further with Attributes corresponding to the selected Source type. It excludes the records by filtering out based on provided filter criteria and executes the matching rules on the filtered records. **Note**: It is an optional step. Multiple filter criteria are allowed in the filter. To add filter conditions, follow these steps a. Click **Source Filter** icon. The Add Source entity Filter window is displayed. b. Click **Add** icon. The Attributes and conditions are displayed. c. Select the attribute, **Like (equal to)**operator from **Attribute Operator** drop-down list and enter the corresponding value(s) in the **Value** text box, respectively. d. Click **Save** to save the filters. e. Click **Delete** icon to delete the corresponding filter. **Note**: • You can add multiple filter conditions based on the available attributes from the **Attribute** drop-down list. • You can enter multiple values for each **Attribute** in the **Value** text box with comma-separated. • The **Value** field accepts only a comma as a special character. If any of the attributes requires a special character other than a comma is not allowed to filter the records. For example, email ID, DOB, Date, Phone number (if it has a special character).
Target	Select the target entity. For example, customer, account, address, etc. **Note**: This field is applicable for Match Ruleset only.
Target Filter	Target Filter icon is applicable for the Target field. You can filter the records further with **Attributes** corresponding to the selected **Target** type. It excludes the records by filtering out based on provided filter criteria and executes the matching rules on the filtered records. **Note**: It is an optional step. Multiple filter criteria are allowed in the filter. How to add filter conditions is similar to the source filter. See Source Filter description. **Note**: If you select the Target as any one of the following, then the only Country attribute is applicable in the Target Filters: • Panama • Bahamas • Paradise • Offshore • ICIJ- Paradise Address • ICIJ- Panama Address • ICIJ- Offshore Address • ICIJ- Bahamas Address
Description	The description of the Ruleset.

Table 9-9 (Cont.) Create a Ruleset

Fields	Description
ER Type	The type of data that is to be matched. The values are as follows: • CSA_8128 • CSA_8129 • Graph (This is applicable if any graphs that are available in the workspace. Graphs could be defined via Graph pipeline) **Note**: It is recommended to use CSA_8129 pipeline with Compliance Studio 8129 and BD 8129 and the other pipeline is deprecated. The lower version pipeline is supported only if you are upgrading to CS 8.1.2.9.0.
Auto Threshold	The threshold value for a ruleset. This is cumulative of all the attribute values entered in the Rule for entity match. One stands for 100%. Enter values from 0 to <=1. Enter a matching percentage above which you want the matches to get into the graph automatically. The threshold scoring logic works in the following way: • A total threshold score of less than 0.7 (70%) must be discarded. • A total threshold score of more than or equal to 0.8 (80%) must create a similarity match automatically. This is configured using the Automatic Threshold in Ruleset. **Note**: The Threshold value must be less than or equal to 1.
Manual Threshold	**Note**: This is applicable only to Match Ruleset. Enter the threshold value for a ruleset. This is cumulative of all the attribute values entered in the Rule for entity match. One stands for 100%. Enter values from 0 to <=1. Enter a matching percentage above which you want the matches to get into the graph automatically. The threshold scoring logic works in the following way: • A total threshold score of less than 0.7 (70%) must be discarded. • The User must manually decide on a total threshold score between 0.7 (70%) and less than 0.8 (80%). This is configured using the Manual Threshold in the Ruleset. **Note**: The Manual Threshold value must be less than the Auto Threshold.

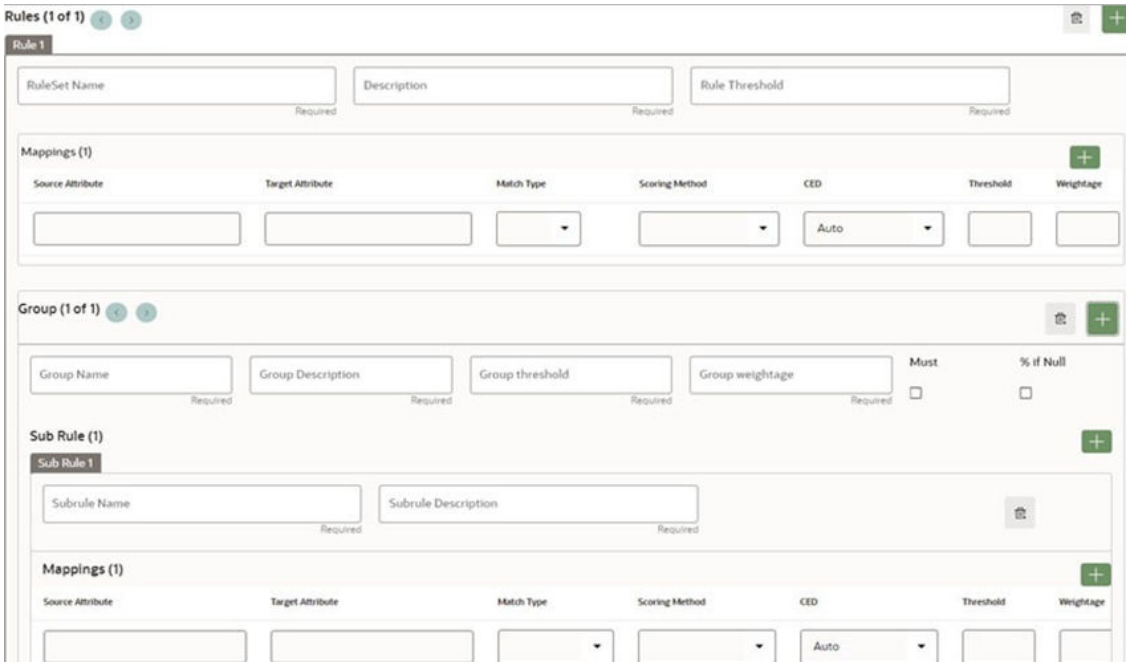
3. Click **Save** to save the Ruleset. You can save the Ruleset with or without creating Rules or click **Cancel** to cancel.

9.3.2.1 Creating Rules in Ruleset

This section describes about creating rules in ruleset.

Every Ruleset has a set of rules such as mapping, groups, rule threshold, scoring method (default or Jaro Winkler, ML Boosted Name, etc.), match type (exact or fuzzy), and source and target attribute (Country, Email, Date of birth, etc.). These attributes help to create a set of rules for a ruleset.

Figure 9-6 Creating Rules in Ruleset



Note:

- You can create multiple rules in the Ruleset, multiple mapping and groups in Rule, and multiple sub-rules and mappings within the Group.
- The sum of the Weightage of Group(s) should be equal to the Weightage of the Rule.
- You can navigate the Rules and Groups within the Rule using the Previous and Next arrow icons. The navigation icons will be enabled only when you create multiple rules and groups within the Rule.
- You can see the Add and Delete options on respective Mapping, Groups, and Sub Rule frames within the Rule tab to perform the operations add and delete, respectively.

To create Rule(s), follow these steps:

1. On the Ruleset Page, click **Add**. The page displays the Rule (n), where n is the number based on the order of the creation.
2. On Rule (n) tab, click **Add** to create Rule. Enter/Select the fields as mentioned in the following table.

Table 9-10 Fields and Description - Create Rule

Field	Description
Name	The name of the Rule. **Note**: Rule Name field accepts only 255 characters.

Table 9-10 (Cont.) Fields and Description - Create Rule

Field	Description
Description	The description of the Rule.
Rule Threshold	The threshold value for a rule. The cumulative threshold value of all attributes in this Rule must not exceed this value. This Rule contributes to the matching only when the maximum score obtained for a rule is equal to or higher than the threshold value. **Note**: The Threshold value must be less than or equal to 1, and it cannot be null.
Source Attribute	The source attribute. For example, date of birth, email ID, entity type, jurisdiction, tax ID, etc. **Note**: Attributes are displayed based on the source entity selected when creating a rule. For example, address, event, etc. It cannot be null.
Target Attribute	The target attribute. For example, entity type, jurisdiction, tax ID, etc. **Note**: Attributes are displayed based on the source entity selected when creating a rule. For example, customer, account, etc. It cannot be null.
Threshold	The threshold score. It indicates the score value provided to the attribute. The sum of the total threshold value of all attributes must be from 0 to <=1. It indicates that a score below the mentioned value does not generate a result from the matching engine.
Weightage	The weightage. It indicates the weightage given for the attributes in the Rule. **Note**: Cumulative of attributes' weightage must not exceed one, and it cannot be null.
Must	It indicates that this attribute cannot have a null value. This attribute must be populated and must return a value for the matching.
% if Null	It indicates that if a Group has null for one entity for all attributes and no scoring attributes. It considers weighted score for mapping is 50% of the weightage.
Match Type	Select one of the following match types: • **Exact**: To obtain the matches that are 100% perfect when finding the entities in a database. • **Fuzzy**: To obtain the matches that are less than or equal to 100% perfect when finding the entities in a database. For more information on the Matching Type, see the OFS Compliance Studio Matching Guide.
Scoring Method	The scoring methods are as follows: • Default • Jaro Winkler • ML-Boosted Name • ML-Boosted Address • Levenshtein • Individual Name • Entity Name For more information on the Scoring Method, see the OFS Compliance Studio Matching Guide.

Table 9-10 (Cont.) Fields and Description - Create Rule

Field	Description
CED	CED (Character Edit Distance) indicates the "fuzziness" for a fuzzy match. An edit distance is the number of one-character changes needed to turn one term into another. These changes can include: • Changing a character (box → fox) • Removing a character (black → lack) • Inserting a character (sic → sick) • Transposing two adjacent characters (act → cat) If no value is selected, CED/ fuzziness is set to AUTO wherein OpenSearch generates an edit distance based on the length of the term. AUTO should generally be the preferred value for fuzziness.
Group name	The name of the Group.
Group Description	The description of the Group.
Threshold	The threshold score. It indicates the score value provided to the attribute. The sum of the total threshold value of all attributes must be from 0 to <=1. It indicates that a score below the mentioned value does not generate a result from the matching engine.
Weightage	The weightage. It indicates the weightage given for the attributes in the sum of sub-rules. **Note**: Cumulative of attributes' weightage must not exceed 1. It cannot be null.
Must	This attribute must be populated and must return a value for the matching.
% if Null	It indicates that if a Group has null for one entity for all attributes and no scoring attributes. It considers weighted score for mapping is 50% of the weightage.
Name	The name of the Sub Rule.
Description	The description of the Sub Rule.

3. Click **Delete** to delete the Rule.

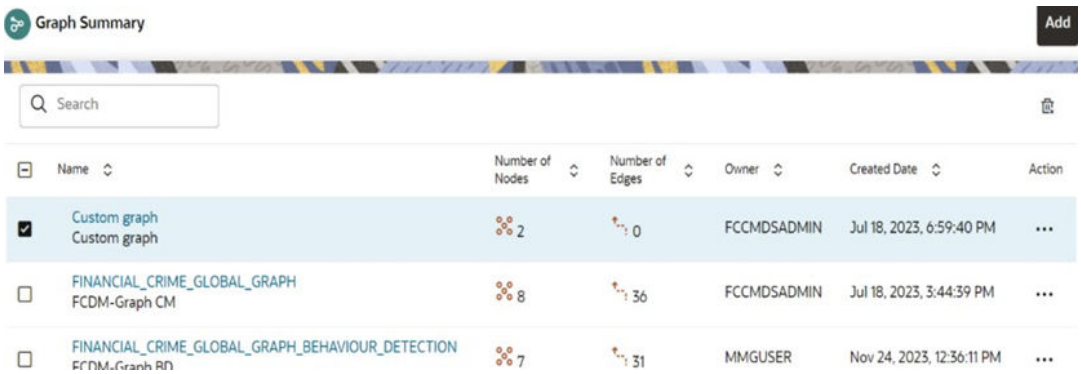
9.3.3 Cloning Match Ruleset

Use this topic to clone ruleset from an existing match rule.

You can also modify the ruleset name, description, source or target entities and attributes for these rulesets.
To clone an existing match rule, follow these steps:

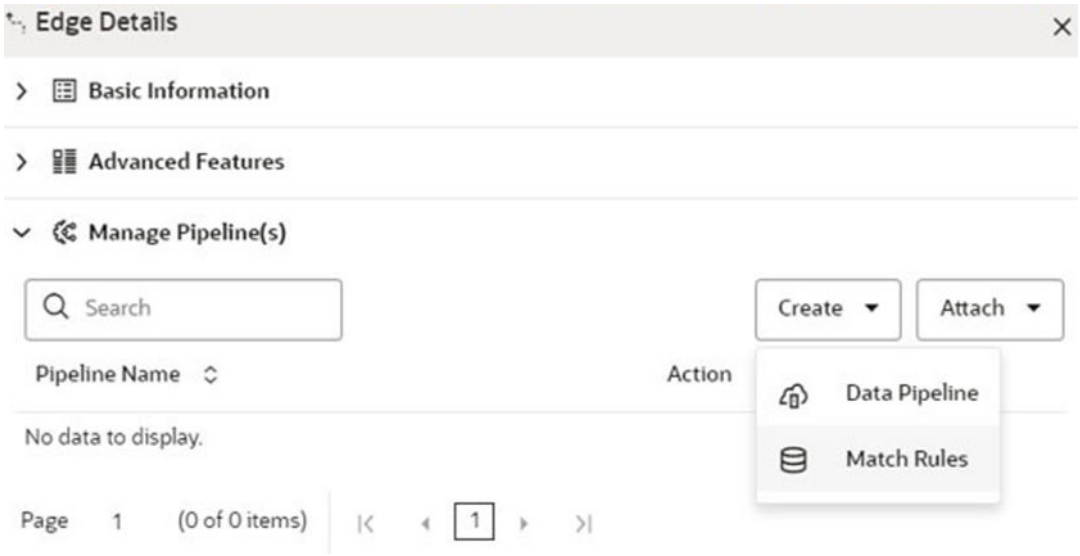
1. Click **Launch Workspace** next to corresponding Workspace to Launch Workspace to display the Dashboard window with application configuration and model creation menu.
2. On **Modeling** menu, click **Graphs**. The Graph page is displayed.

Figure 9-7 Graph Summary



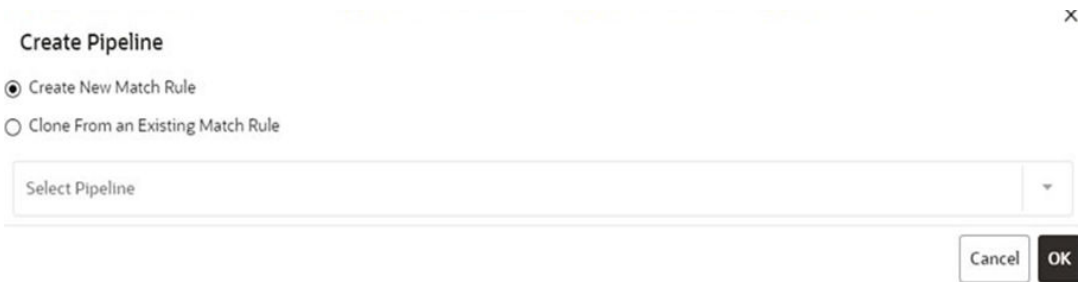
3. Hover over the graph pipeline and click edge on the graph. The Edge details page is displayed.

Figure 9-8 Edge Details



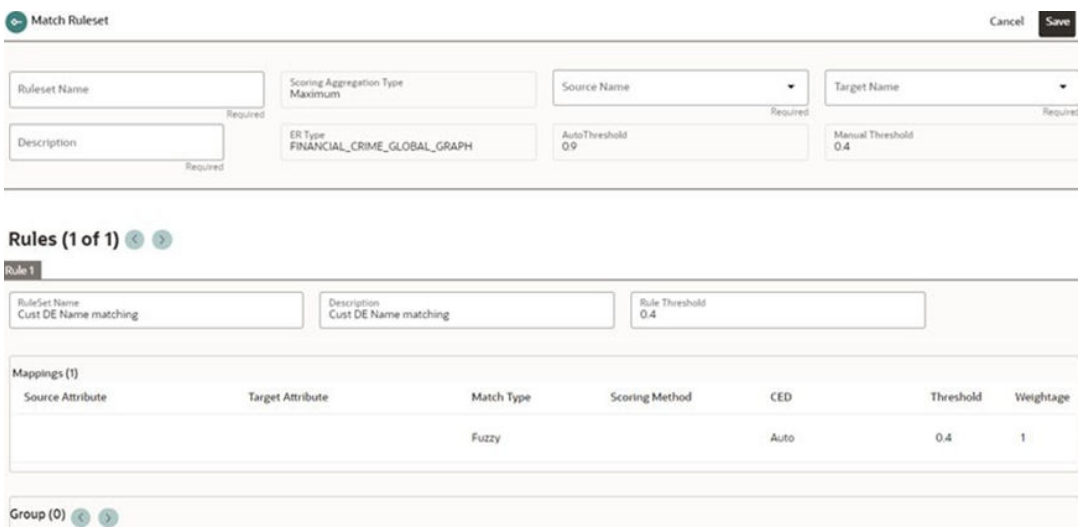
4. From the **Create** drop-down list, select the **Match Rules**. The Create Pipeline window is displayed.

Figure 9-9 Create Pipeline



5. Select **Clone From an Existing Match Rule** option and required pipeline from the **Select Pipeline** drop-down list and then click **OK**. The Match Ruleset window is displayed.

Figure 9-10 Match Ruleset Details



6. You can enter/select the Ruleset details for the following fields:
   - Ruleset Name
   - Source
   - Target
   - Description

   For information on the fields in the Ruleset Details window, see the Creating Match/Merge Ruleset section.

> **Note:**
>
> For cloning an existing match rule, the selected source and target attributes must be same as the source and target attributes of the existing match rule.

7. Click **Save**. The new match ruleset is cloned.

9.3.4 Data Survival Rules

This Rule facilitates storing the master record/final output tables based on the dataset survival rule stored against pipeline id.

Compliance Studio provides OOB rules; however, you can modify these rules or create your rules:

- Using Data Survival
- Creating Data Survival

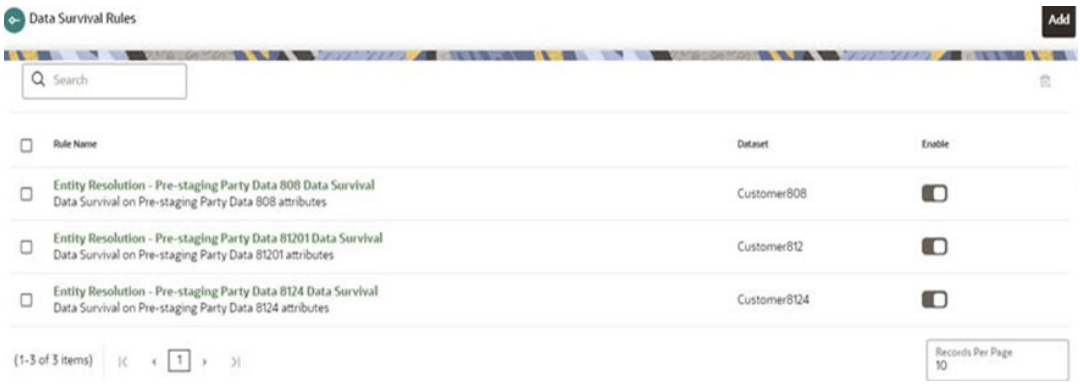
9.3.4.1 Using Data Survival Rules

Use this topic to enable or disable the OOB rules.

You can also modify this Rule's description and ER type attributes and remove rulesets from the existing list.
To use the existing Ruleset, follow these steps:

1. On the Workspace Summary page, click **Launch Workspace** to display the CS Production Workspace window with application configuration and model creation menu.
2. On **More** menu, click **Data Survival**. The Data Survival Rules page is displayed.

Figure 9-11 Data Survival Rules

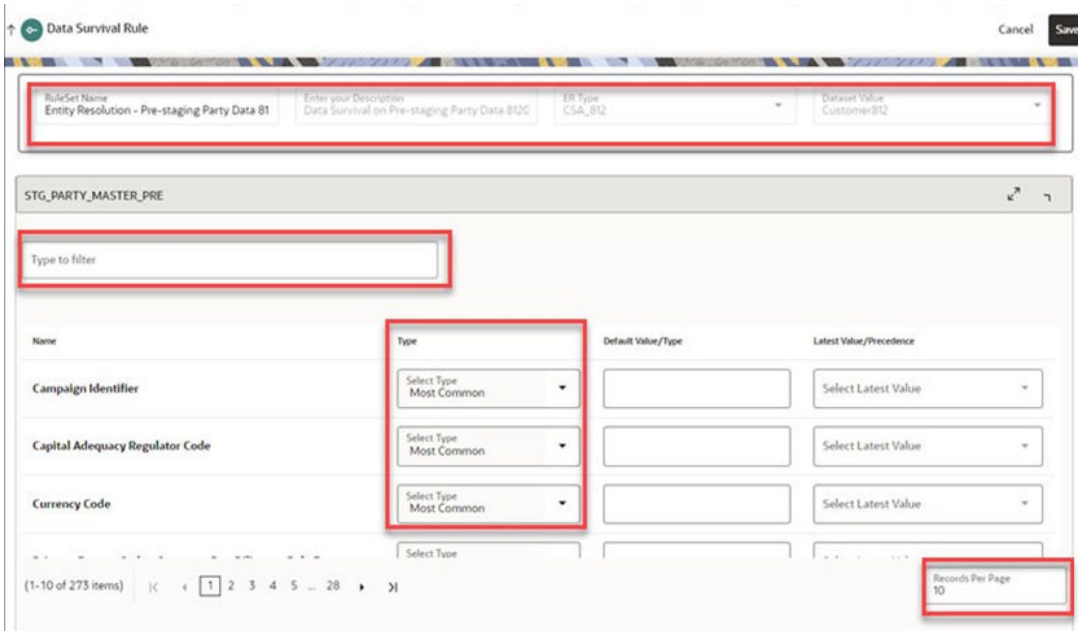


3. To enable the Rule, select the Rule and move the toggle switch to the right.
4. To add the Rule, click **Add**. For more details, see Creating Data Survival Rules section.
5. To delete the Rule(s), select the check box(es) corresponding to the rule(s) and click **Delete**.

   The confirmation message is displayed with details of Rule(s) to select again. You can check or uncheck check boxes for the required rule(s) and then click **Delete**.
6. To edit the Rule, click on the **Rule**, and the corresponding Rule page is displayed.
7. On the Rule page, click **Edit**. The page displays Rule details.
8. You can modify the Rule details except for the following fields:
   - RuleSet Name
   - Description
   - ER Type
   - Dataset Value

> **Note:**
>
> Fields marked with * are mandatory fields.

Figure 9-12 Edit Data Survival Rule



9. On this page, you can perform the following actions:
   - You can use the search filter to search for any particular attribute in each input table. The records will be displayed based on the specified value in the search filter (Type the filter).
   - You can sort the records.
   - Right-click on any of the columns and select **Sort** > **Sort Ascending** or **Sort Descending** for ascending and descending order, respectively. The records will be displayed according to sort order.
   - You can select the method type from the **Type** drop-down list.

> **Note:**
>
> You must sort the records again when changing the value in the **Records per Page** field.

10. Click **Save** to save the changes or click **Cancel** to revert the changes.

9.3.4.2 Creating Data Survival Rules

This topic describes the systematic instructions to create data survival rules for match ruleset.

To create data survival rules, follow these steps:

1. On the **Data Survival Rules** page, click **Add**. The page displays the attributes to create the Rule.

> **Note:**
>
> You should not create multiple data survival rulesets for a particular ER Type.

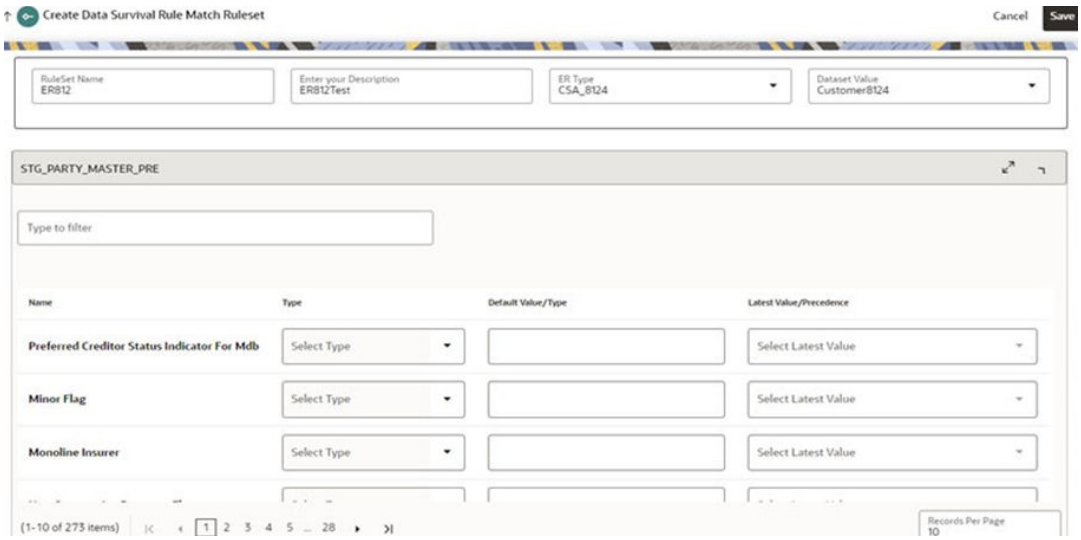
2. Enter the field values as described in the following table.

Table 9-11 Fields and Description - Data Survival Rules

Field	Description
Rule Name	The name of the Rule.
Description	The description of the Rule.
ER Type	The type of data that is to be matched. The values are as follows: • CSA_8128 • CSA_8129 **Note**: It is recommended to use CSA_8129 pipeline with Compliance Studio 8129 and BD 8129 and the other pipeline is deprecated. The lower version pipeline is supported only if you are upgrading to CS 8.1.2.9.0.
Dataset	The name of the output data structure to which this data is saved. The values are as follows: • Customer8128 • Customer8129

Once you select the Dataset, the input tables are displayed. The list of tables will change based on the Dataset selection. Each table displayed can be expanded to allow you to select the survival function for each attribute if needed.

Figure 9-13 Create Data Survival Rules



3. On the Create Data Survival Rule Match Ruleset page, perform the following:
   - Filter the records using the search filter (Type the filter)
   - Sort the records
     To sort the records, right-click on any of the columns and select **Sort** > **Sort Ascending** or **Sort Descending** for ascending and descending order, respectively. The records will be displayed according to sort order.
   - Select the method types. See the following step for method type details.
4. You can select the following method types used to store the final record.

- **Generate**: This is applicable for primary keys of the tables, where the system will generate the unique identification number for the record.
- **Generate Sequence**: This is applicable for child records, where the system will generate the identification number in sequential order of the particular parent key.
- **Distinct**: This will apply to child tables only. This will remove the duplicate values and allow the user to choose to save the unique record. This applies to columns that are marked as primary keys.

> **Note:**
>
> The user should not have Type **Distinct** and **All** together with other columns that return unique values in child tables.

- **Latest**: A date column like Data dump date against each row in the input tables will pick the latest date. For example, if the customer's Occupation is Teacher and Businessman and all the records have a Data Dump date against them, the latest record value will be taken.
  You can select the latest date attribute from the **Latest Values** drop-down list.
- **Longest**: It counts the number of characters and picks the most extended values.
- **Most Common**: This will apply to the most repeated value that will be picked for data survival. For example, if the Occupation of three customers is 1. Business, 2. Business, and 3. The teacher then Business occurs twice so that value will be picked.
- **All**: This case can only apply to child records, where you can store multiple values against customer id. For example, you have three customers to be merged, and all three have different email IDs, so all the email IDs will be stored against the merged Global entity.
- **Default**: This will be used for inserting a user-defined value for columns. For example – if the user wants to provide the default value of the **Branch_Code** column as **ER**, then ER will be inserted for all the rows.
- **Maximum**: This will apply to number columns. Users can choose this for storing the maximum number value. For example, if Annual Income is C1 - 2M, C2- 3M, and C3 - 5M, then 5 M will be saved for the global entity.
- **Minimum**: This will apply to number columns. Users can choose this for storing minimum number values. For example, if Annual Income is C1 - 2M, C2- 3M, and C3 - 5M, then 2M will be saved for the global entity.
- **Null**: The user can choose not to insert any value while persisting the record. So null will signify a null value for that particular column.
- **Metadata**: This will apply to attributes that should be selected based on metadata, for example, to select the occupation with the highest risk score.
  To enable this, select the **Type** as **Metadata**. You can select the metadata type in the **Default Value/Type** drop-down list.

  The types of Metadata are user-defined and can be set using an API call (See **Populate the Metadata for Data Survival in Studio Schema** section in the OFS Compliance Studio Administration and Configuration Guide). Each attribute value that appears in the metadata for the given type will be given a numerical value. In the UI, select the precedence value in the **Latest Values/Precedence** drop-down list as either Minimum or Maximum. This will select the attribute with the highest or lowest numerical score as part of the Data Survival logic.

> **Note:**
>
> – This is applicable only for String attributes.
>
> – Numerical scores have to be provided against each attribute value in the source data. Where a value does not exist in the metadata, then no numerical value will be assigned, and it will be excluded from selection if other attributes have values.
>
> – Where no strings are in the metadata, the data survival algorithm will select any one of the attributes for the output.

- **User Defined Method**: It displays the custom data survival method created by the user. To create a custom data survival method, see the **FCC_DATASURV_TYPE** table in the OFS Compliance Studio Administration and Configuration Guide.

> **Note:**
>
> If an invalid value is returned from the data survival logic defined in the custom method, then the job will be successful. However, the record is not survived in the Staging Output tables.

5. Click **Save** to save the Rule or click **Cancel** to revert the changes.

9.4 Using Manual Decisioning

The Manual Decisioning allows you to make manual decisions on similarity edges in the Global Graph and matches for Entity Resolution, where the score is between the manual and automatic thresholds set on the Match Ruleset page.

This flexibility provides an Ad-hoc experience to allow users to make decisions on borderline cases.

- **Entity Resolution** (CSA8xx): If the decision is approved, the match is added to the probable groups for merging in the Entity Resolution process. If the decision is rejected, it is not considered in probable groups.
- **Global Graph**: If the decision is approved, the match edge will be moved to Global Graph. The match edge will not be part of the Global Graph if the decision is rejected.

You can specify your Manual Threshold on the Match Ruleset Details page for Manual Decisioning.

To take the manual decision, follow these steps:

1. On **More** menu, click **Manual Decisioning**. The Manual Decisioning page is displayed.

Figure 9-14 Manual Decisioning

The role icons are at the top-right corner, the 1st is Request User, and the 2nd is Approver User. The number of icons will be based on user privileges. If both icons are shown, you can choose whether to make a decision as an Request User or Approver User.

2. You must select the following to display the records on the page:
   - Manual Decision Type
   - Select Pipeline / Select Graph
   - Data Source Type

Other than Graph, records will be displayed only when you select the Data Source Type for Entity Resolution (CSA8xx) (ER Type).

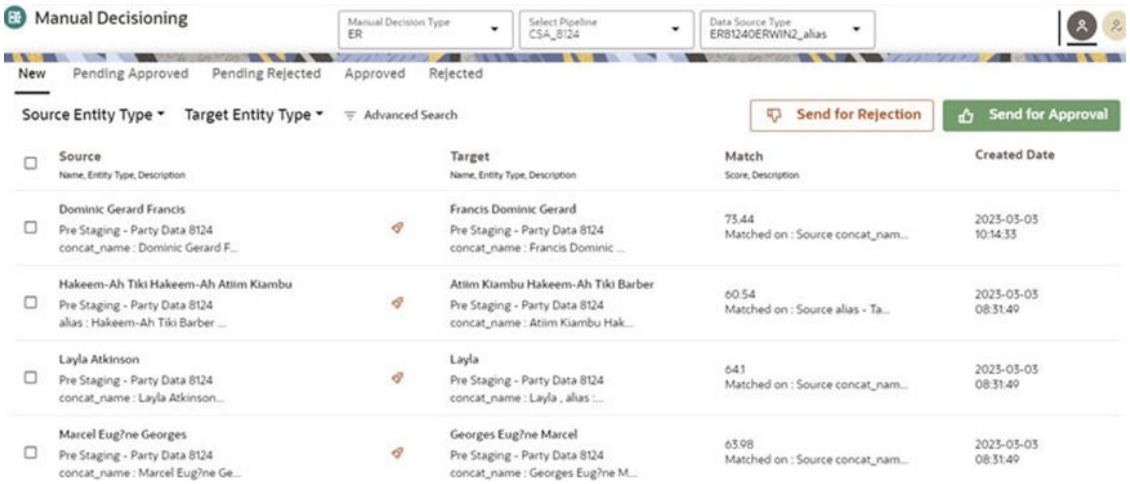
> **Note:**
>
> The records on the Manual Decisioning page will be displayed based on the selection of Manual Decision Type. The Data Source Type field is applicable only when ER Type is selected as Entity Resolution (CSA8xx).

Records are displayed on one of the following tabs, depending on where they are in the Manual Decisioning workflow:

> **Note:**
>
> Any manual actions performed in the UI will undergo the Data Survival rule and data will be persisted accordingly.

- **New**: Where no decision has yet been made
- **Pending Approved**: Where Request User has made an approve decision and this is to be confirmed or rejected by a Approver User
- **Pending Rejected**: Where Request User has made a reject decision and this is to be confirmed or rejected by a Approver User
- **Approved**: Where the decision has been approved
- **Rejected**: Where the decision has been rejected
  For more details, see the Workflow for Manual Decisioning section.

Figure 9-15 Manual Decisioning with Records



By default, ten edges or matches are displayed per page. Click the **Next** arrow to view the next set of edges or matches.

You can filter the edges using the following options:

- Source Entity Type (Entity Resolution (CSA8xx))/Source Node Type (Graph)
- Target Entity Type (Entity Resolution (CSA8xx))/Source Node Type (Graph)
- Advanced Search

Customizing View on the Manual Decisioning

To customize view on the page, follow these steps:

1. Select any column on the page.
2. Right-click and the following options are displayed:
   - **Columns**: You remove and add the column form list.
   - **Sort Ascending**: You can sort the edges on the page in Ascending order.
   - **Sort Descending**: You can sort the edges on the page in Descending order.
   - **Resize Column**: You can resize the column by specifying the column width.

Applying Filter on the Edges

To filter edges on the page, follow these steps:

- **Source Entity Type (Entity Resolution (CSA8xx))/Source Node Type (for Graph)**: Select the source from the list. It filters the edges based on the list under the Source column.
- **Target Entity Type (Entity Resolution (CSA8xx))/Source Node Type (for Graph)**: Select the target from the list. It filters the edges based on the list under the Target column.
- **Advanced Search**: You can filter in which the recodes search criteria as specified. For more details, see the Searching for Manual Decisions section.

> **Note:**
>
> You can apply all the filters in parallel. It considers source, target, and search criteria specified in Advanced Search. It uses AND logic. For example,
>
> – Source Node Type/Source Entity Type is the customer
> – Target Node Type/ Target Entity Type is customer
> – Advanced Search: The source Name is GRAF
> – The column names (Source and Target) are replaced with selected Source and Target values from the drop-down list.
> – You can drag and drop the columns on the page to change the order.

On the Manual Decisioning page, when you select the possible match from the **New** tab, the**Details** pane is populated on the right side of the screen, giving more detail on the possible match to allow the decision to be made.

> **Note:**
>
> The view edge is different for **Graph** and **Entity Resolution (CSA8xx)**.
> For more information, see the Viewing Edges from the Source section.

9.4.1 Workflow for Manual Decisioning

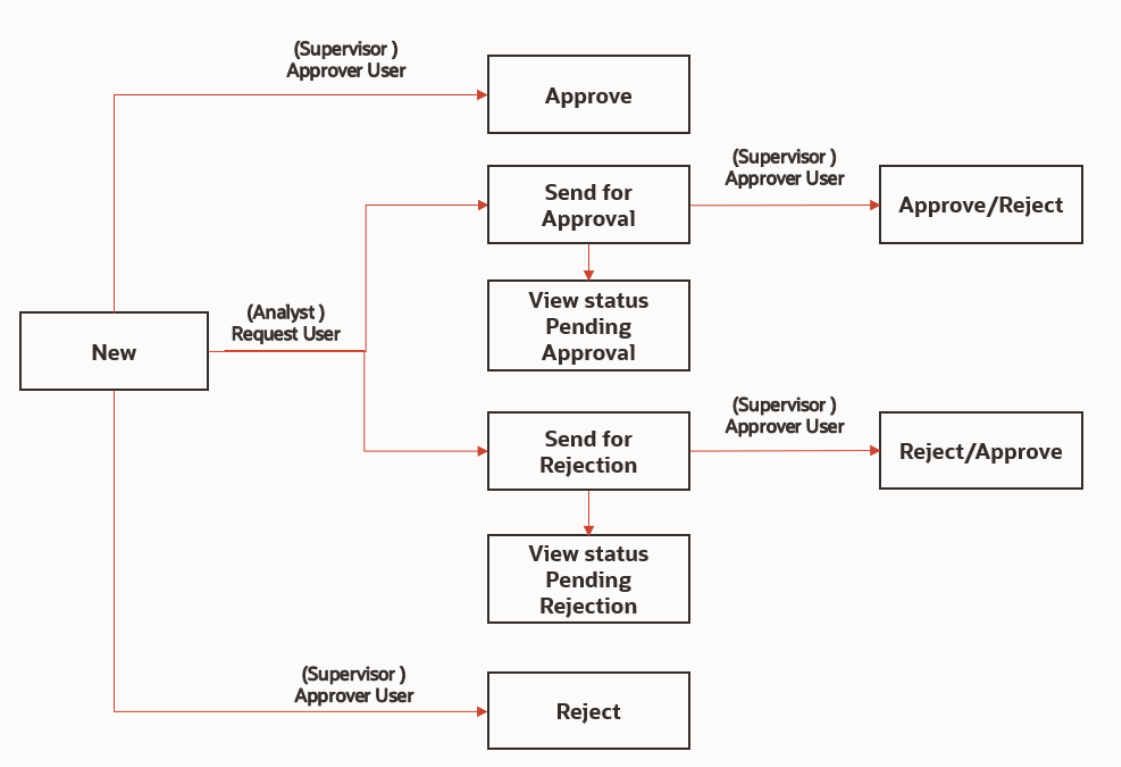
This topic describes workflow for the Manual Decisioning.

You can take the following decision as an **Analyst** (New tab) or **Supervisor** (New, Pending Approved, Pending Rejected tabs) on this page:

- Send for Approval
- Send for Rejection
- Approve
- Reject

The following figure illustrates the Manual Decisioning workflow.

Figure 9-16 Manual Decisioning workflow



9.4.2 Searching for Manual Decisions

This topic describes the systematic instructions to search the manual decision actions.

You can specify the Advanced Search using AND logic. It will display the edges only when they satisfy the search criteria.
To filter edges in the Advance Search, follow these steps:

1. Click **Advanced Search**. The Advanced Search window is displayed.
2. Select the fields from the **Select Fields** drop-down list. You can add multiple search criteria.
3. You can perform the following actions:
   a. Click **Add** icon to add the field value for the selected field.
   b. Click **Edit**icon to edit the field value for the selected field.
   c. Click **Delete** icon to delete the field value for the selected field, if required.
   d. Click **Apply** to display edges per search condition or click Cancel to Cancel.
   e. Click **Reset** to clear the specified values.

9.4.3 New, Pending Approved, and Pending Rejected Tabs

This topic describes different action tabs in the Manual Decisioning.

The action tabs contain a list of the similarity edges or Entity Resolution matches that require a decision to be made or approved. These tabs display the **Source** (Name, Node/Entity Type,

Description), **Target** (Name, Node/Entity Type, Description), **Match** (Score, Description), and **Created Date** information that can be **Rejected** and **Approved** to initially find the match. Additional information can be seen in the Details Panel.

Comments: The supervisor can also review the analyst's comments on Pending Approved and Pending Rejected tabs and take further action (Approve or Reject).

For more details, see the Workflow for Manual Decisioning section.

9.4.4 Approved and Rejected Tabs

The Approved and Rejected tab contain a list of the similarity edges or Entity Resolution matches that have been approved or rejected.

These tabs display the sources, Target, Match Score description, Created Date, and Comments Approved or Rejected that are sent by the Analyst and reviewed by Supervisor. No further actions can be taken from these tab.

9.4.5 Taking Decision on Similarity Edges

This topic describes the systematic instructions to take decision on the similarity edges.

To take the decision as an Analyst or Supervisor, follow these steps:

1. Select the edge and perform the following:
   a. Analyst
      - Click on **Send for Approval** to send for approval.
      - Click on **Send for Rejection** to send for rejection.

   b. Supervisor
      - Click on **Approve** to approve.
        In the case of **Graph**, the match edge will be moved to Global Graph.

        In the case of **Entity Resolution (CSA8xx)**, probable groups will be created for the selected records, and the new Global party IDs will be generated based on the merge rule.
      - Click on **Reject** to reject.
        In **Graph**, the match edge will not be part of the Global Graph.

        In the case of **Entity Resolution (CSA8xx)**, probable groups will be created for the selected records, and the same Global party IDs will be retained as-is.

2. Report Action Pop-up window appears, provide comments in the **Comments** text box or select **Standard Comments** from the drop-down list for selected edge(s):
   - Sending for Approval or Approve
   - Sending for Rejection or Reject

   You can provide free-form comments for the selected similarity edge using the **Comments** text box to provide your own comments or select from the Select **Standard Comments** drop-down list. The available options are as follows:
   - Additional identifier indicates the same entity
   - Address not similar enough
   - DOB missing or conflicting
   - Name is not similar enough

- No additional identifiers are available
- No transaction activity indicating the same entity
- Transactional behavior indicates same entity

> **Note:**
>
> Send for Rejection/Reject and Send for Approval/Approve can be performed for single or multiple edges.
> You can provide comments in the **Comments** text box or Select Standard Comments. Once you select Comments, the other one will be disabled automatically and vice versa.

3. Click **Save** to save the change or click **Cancel** to revert the changes.

9.4.6 Viewing Edges from the Source

This topic provides information about viewing edges in the Details pane and validate it to include in your graph.

On the Manual Decisioning page, when you select the edge from the action tabs (i.e, **New**, **Pending Approved**, **Pending Rejected**, **Approved**, and **Rejected**). The Details pane is populated on the right side of the screen.

You can view the edge in the Details pane of Graph and Entity Resolution (CSA8xx).

You can perform actions (**Approve**, **Reject**, **Send for Approval**, and **Send for Rejection**) on the View Edges pane for ER Type is Entity Resolution (CSA8xx).

> **Note:**
>
> - If you are viewing the paginated graph, increase the Page Size of the graph in Manual Decision UI to avoid the Paginated graph.
> - Navigate to **Settings** > **Visualization** > Increase the Page size to accommodate the graph on one page.
> - You need to unlock the Manual Decision notebook from the Notebook server UI to view the graph.

9.5 Using Merge and Split Global Entities

Merge and Split Global Entities allows you to manually break, merge, split, create, and re-arrange parties into global parties.

This flexibility provides an Ad-hoc experience to enable users to make decisions on borderline cases or where data issues or rules logic has grouped records incorrectly. It also allows users to approve system changes to global party membership where this has been configured.

On **More** menu, click **Merge and Split Global Entities**. The Merge and Split Global Entities page is displayed.

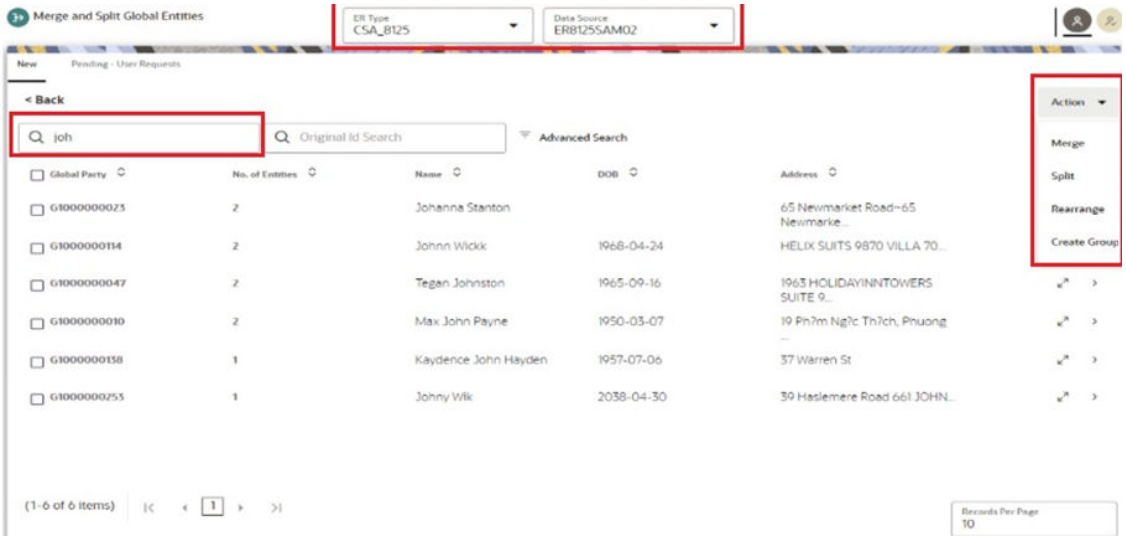
Figure 9-17 Merge and Split Global Entities

The role icons are at the top-right corner, the 1st is **Analyst**, and the 2nd is**Supervisor** . The number of icons will be based on the user privileges. If both icons are shown, you can choose whether to make a decision as an Analyst or Supervisor.

Unlike the Manual Decision screen, there are no records displayed initially. The user needs to search for entities to merge or split.

First, select the **ER Type** and **Data Source** drop-down list, and click the **Search** icon after entering the customer name or customer ID in the search filter.

Figure 9-18 Merge and Split Global Entities with Records



By default, the ten records are displayed per page. Click the **Next** arrow to view the next set of records. You can filter the records using the search filter based on a customer name in the **Name Search** field or Customer ID in the **Original Id Search** field and **Advanced Search** based on search criteria.

Note:

- Enter at least 3 characters for every token in the search field for better search results.
- The following special characters are not supported for search result: *&\%_,~-;$?[ ]{ }=( )>!|
- Advanced Search will further filter the results from the initial search (Name and Original ID).
- The maximum number of global party IDs is 100. The page will display only 100 global party IDs in the UI even if the results of the partial name search are more than 100 and the order of display will be random.

In addition, you sort the records by selecting each column.

Note:

- You can apply all the filters in parallel. It considers customer name or customer ID and search criteria specified in Advanced Search. It uses AND logic.
- For example, enter the name Lily in the search filter and select Country as the US in an Advanced Search.
- The Global Party records with customer details (Customer name and Country) are displayed.
- The records are displayed and refreshed based on the ER Type and Data Source for Entity Resolution (CSA8xx) selection on the New and Pending tabs along with the search filter.

When you initially select records on which to take action, you will see some details for the Entity:

- **Card View**: Displayed as a card with the customer ID, name, DOB, and address.

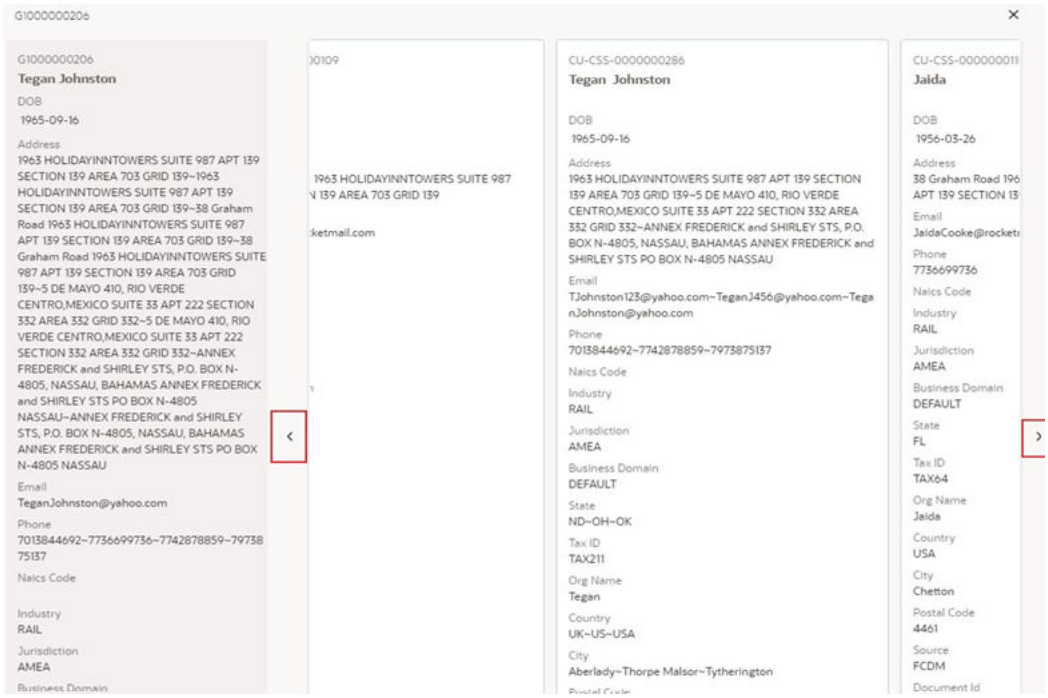
Figure 9-19 Card View

> **Note:**
>
> If entity does not belong to the User's business domain and jurisdiction, then all attributes are masked except Original ID.

To help in making a decision, you can view the Expand the records to Full view (to compare the customer records with Global Party details).

- **Full View**: Highlighted the comparison in the Customer details with respect to Global Party details. You can see next customer details by clicking the **Next** and **Previous** arrows.

Figure 9-20 Full View



Analysts and Supervisors can perform the following actions:

> **Note:**
>
> Any manual actions performed in the UI will undergo the Data Survival rule and data will be persisted accordingly.

- **Merge**: Customers in the existing Group have different global ids assigned. These customers are merged into a single group, and a new global id is assigned.
- **Split**: A single global id is assigned to all customers in the existing Group, and the customers are split into different groups with new global ids assigned to each.
- **Rearrange**: A number of parties exist in more than one global party with different global ids assigned. This action can be used to re-arrange the parties into different global parties. New global party ids will be assigned.
- **Create Group**: A new global party is created to which customers can be moved. New global party ids will be assigned.

- **Override Global Party (Reset to Source)**: The override flag is enabled only when manual action is taken on the particular global party id.

Additionally, supervisors can take the following actions on the Pending - User Requests and Pending - System Requests tab, where an Analyst has taken action already:

- **Approve**: The action taken by the analyst is approved, and the changes are committed.
- **Reject**: The initial state of the Global Party before manual action will be retained.

9.5.1 Workflow for Merge and Split Global Entities

This topic described workflow for the Merge and Split Global Entities.

You can take the following actions as an Analyst (New tab) or Supervisor (New tab)) on this page:

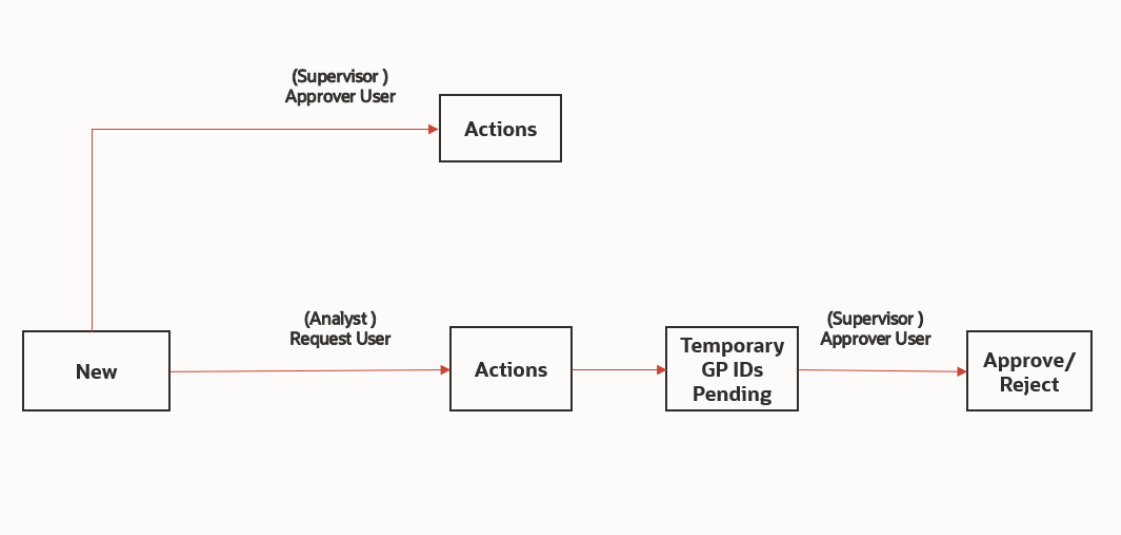
- Merge
- Split
- Rearrange
- Create Group

As a Supervisor, you can make a decision on the following actions taken by the Analyst.

- Approve
- Approve All
- Reject

The following figure illustrates the Merge and Split Global Entities workflow.

Figure 9-21 Merge and Split Global Entities workflow



9.5.2 Searching for Global Entities

This topic describes the systematic instructions to search global entities.

You can specify the Advanced Search using AND logic. It will display the records only when they satisfy the search criteria.
To filter the records in the Advanced Search, follow these steps:

1. Click **Advanced Search**. The Advanced Search window is displayed.
2. Select the fields from the **Select Fields** drop-down list. You can add multiple search criteria.
3. You can perform the following actions:
   - Click **Add** icon to add the field value for the selected field.
   - Click **Edit** icon to edit the field value for the selected field.
   - Click **Delete** icon to delete the field value for the selected field.
   - Click **Apply** to display edges per search condition or click **Cancel** to cancel.
   - Click **Reset** to clear the specified values.

9.5.3 New Tab

The results are displayed for each action (Merge/Split/Create/Create Group) in New/<Actions> tab with a new Global Party ID and the number of customers in No. of Entities.

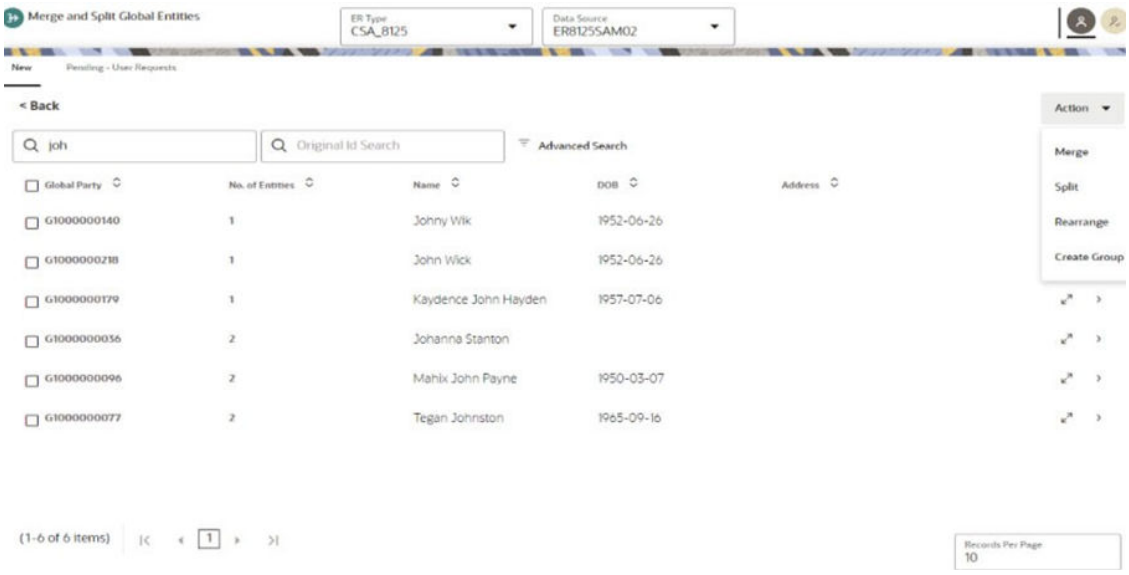
Enter the customer name or ID in the search filter and click the **Search** icon. The record information (**Global Party**, **No. of Entities**, **Name**, **DOB**, and **Address**) is displayed.
Use breadcrumb locator <Back link to navigate back to Home after checking the action results.

For Rearrange, you can select the records along with Global parties on the page and select the Global Party where you want to move all selected records. You must save the changes to reflect on the page.

> **Note:**
>
> You must select at least two Global Parties for Merge/Rearrange/Create Group. For Split, only one Global Party.

For more details, see the Workflow for Merge and Split Global Entities section.

Figure 9-22 Fields on New Tab



Analyst or Supervisor can perform the following actions:

1. Expand the records to Full view (to compare the customer records along with Global Party details) or Card view (to view customer name, id, and address).
2. Select the customer records and Global Party.
3. Click **Action** drop-down list and select any one of the actions according to your requirement:
   - **Merge**: You must select at least two Global Parties.

     Note:

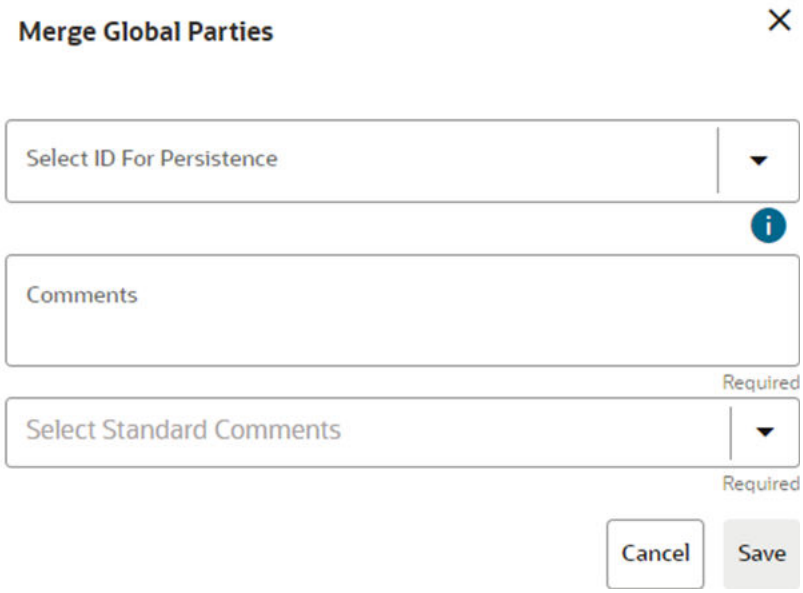
     – You can persist the Global Party ID using the **Select ID For Persistence** drop-down list.
     Persisting Global Party ID means one of the new parties retains the previous global id and hence retains the history of the previous global party.

     – It can only be done at the Global Party level, not for individual customer records in different Global Parties.

   - **Split**: You should select only one Global Party and multiple records to split.
   - **Rearrange**: You must select at least two Global Parties.
     – Select the Global Parties on which the user wants to perform Rearrange. Then the user will be moved to **New/Rearrange** breadcrumb, where you need to select customers and perform Rearrange.
     – Click **Save** to save the changes or click **Cancel** to revert the changes.
   - **Create Group**: The records and respective Global Parties must be selected.

4. After selecting the action, you can provide free-form comments for the selected action using the**Comments** text box or select comment from the **Select Standard Comments** drop-down list.

Figure 9-23 Merge Action pop-up



> **Note:**
>
> The **Select ID For Persistence** drop-down list is applicable for the following actions:
>
> - Merge
> - Split

Persisting Global Party ID means one of the new parties retains the previous global id and hence retains the history of the previous global party.

The available standard comments are as follows:

- Merge Entity as identifiers indicate the same entity
- Merge Entity as addresses indicate the same entity
- Split Entity as address not similar enough
- Split Entity as DOB missing or conflicting
- Split Entity as name not similar enough
- Split Entity as no additional identifiers available or not similar enough
- Remove any manual split or merge - fixed at source

> **Note:**
>
> – Actions can be performed for single or multiple Global Parties.
>
> – You can provide comments in the **Comments** text box or Select Standard Comments. Once you select Comments, the other one will be disabled automatically and vice versa.

The results are displayed to rearrange on New/<Action> tabs. The new Global Party ID has the selected records while creating the group, and existing the Global Party is updated with the number of records in No. of Entities.

9.5.4 Pending - User Requests Tab

The Pending - User Requests tab displays Cluster ID details for which Supervisor needs to take an action.

This tab shows **Cluster ID**, **Action**, **Created Date**, **Comment** and **Persisted Global ID** information.

Figure 9-24 Pending - User Requests Tab

Cluster

A Cluster is a group of global parties that collectively participates in an action. This action can involve various operations such as Split, Merge, or a combination of both Split and Merge.

For example, G1 has C1 and C2 entities, G2 has C3 and C4 entities. While performing a Split & Merge of G1 & G2, such that G1 and G2 are split, then C1, C3 are merged to be a single group and C2, C4 are merged to be another group; and then both G1 and G2 are grouped together to form a single cluster.

The global parties are processed as clusters in order to maintain the integrity and consistency of the resolution of entities.

To Approve or Reject as Supervisor on actions taken by the Analyst, follow these steps:

1. Select the record(s) and perform the following:
   - Click on **Approve** to approve. The match is resolved as per action which Analyst has taken.
   - Click on **Reject** to reject. The initial state of the Global Party before manual action will be retained.
2. A pop-up window is displayed for both Approve and Reject.

- For Approve
  a. You can provide free-form comments for the selected action using the **Comments** text box or select comment from the **Select Standard Comments** drop-down list.
  b. Click **Save** to approve the action or click **Cancel** to revert the changes.
- For Reject
  a. You can provide free-form comments for the selected action using the **Comments** text box or select comment from the **Select Standard Comments** drop-down list.
  b. Click **Save** to reject the action or click **Cancel** to revert the changes.

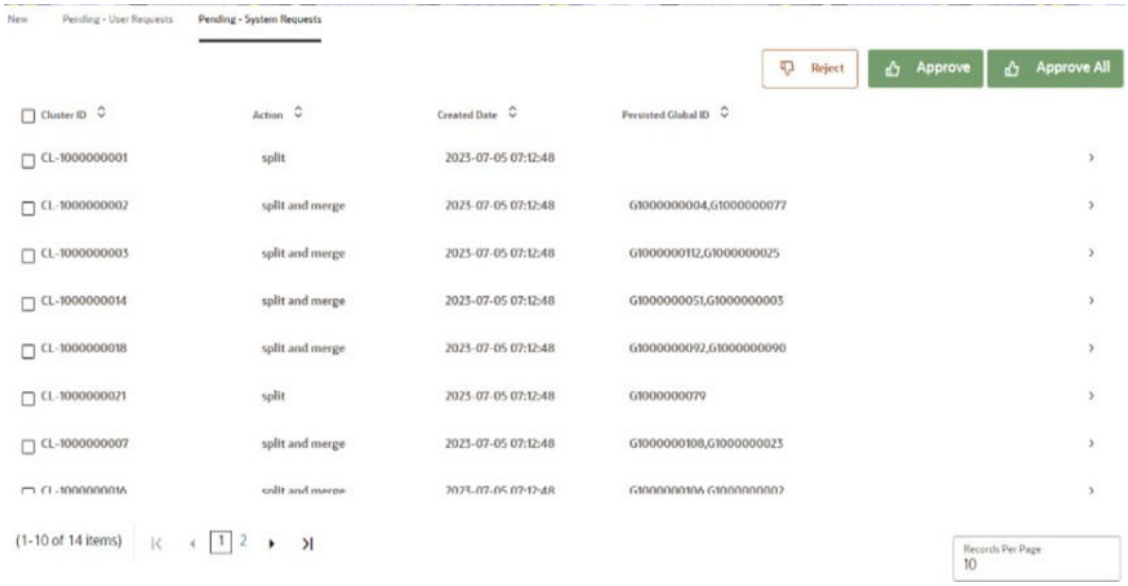
9.5.5 Pending - System Requests Tab

The Pending - System Requests tab displays Cluster ID details for which changes have been made by the system (Batch) and manual approval is required.

This tab shows **Cluster ID**, **Action**, **Created Date**, and **Persisted Global ID** information.

> **Note:**
>
> - The records will be displayed only if the GID Persistence and Manual Approval flags are set to true.
> - If both the flags are set to false, the GIDs are not persisted and a new global party will be created.
> - If GID persistence flag is set to true and manual approval flag is set to false, then the GID is persisted but the requests are approved automatically.
>   By default, the Global Party ID will be persisted that has most number of entities and if the number of entities are same between the groups, the global party with the lowest global party id will be persisted.
>
>   For more information about Global Party ID persistence, see the **Persisting the Data** section in the OFS Compliance Studio Administration and Configuration Guide.

Figure 9-25 Pending - System Requests Tab



For more information about the Cluster ID details, see the Cluster section.

To Approve/Approve All/Reject as Supervisor for the changes are done by the system on running the ER batches, perform the following:

1. Select the record(s) and perform the following:
   - Click on **Approve** to approve the selected records.

     > **Note:**
     >
     > Enable the **Cluster ID** check box and click **Approve** to approve all the records. All records will be approved only on the selected page. If you need to approve the remaining records, then navigate to the next page and perform the same action again to approve all records on the particular page.

   - Click on **Approve All** to approve all the records that are updated by the batch action.

> **Note:**
>
> The records need not to be selected when using the **Approve All** button.
>
> – If F_ER_DS_SUBSEQUENT_BATCH is set to False in the FCC_ER_CONFIG table in the ER Schema. After clicking **Approve All**, all the records will be approved only on the selected page. If you need to approve the remaining records, navigate to the next page and click **Approve All** again to approve all the records on the particular page.
>
> – If F_ER_DS_SUBSEQUENT_BATCH is set to True, then **Approve All** button approves all the records across all the pages.
> To configure F_ER_DS_SUBSEQUENT_BATCH and ER_DS_SYSTEM_PENDING_MAX_NO_REC, see the **Create Index and Load the Data** section in the OFS Compliance Studio Administration and Configuration Guide.

- Click on **Reject** to reject the selected records.

2. A pop-up window is displayed for **Approve**, **Approve All** and **Reject** actions. For example, if it is Approve action.
   a. You can provide free-form comments for the selected action using the **Comments** text box or select comment from the **Select Standard Comments** drop-down list.
   b. Click **Save** to approve the action or click **Cancel** to revert the changes.

9.5.5.1 Persisting Global Party ID through the Manual Action

This topic describes the systematic instructions to persist Global Party ID through the manual action.

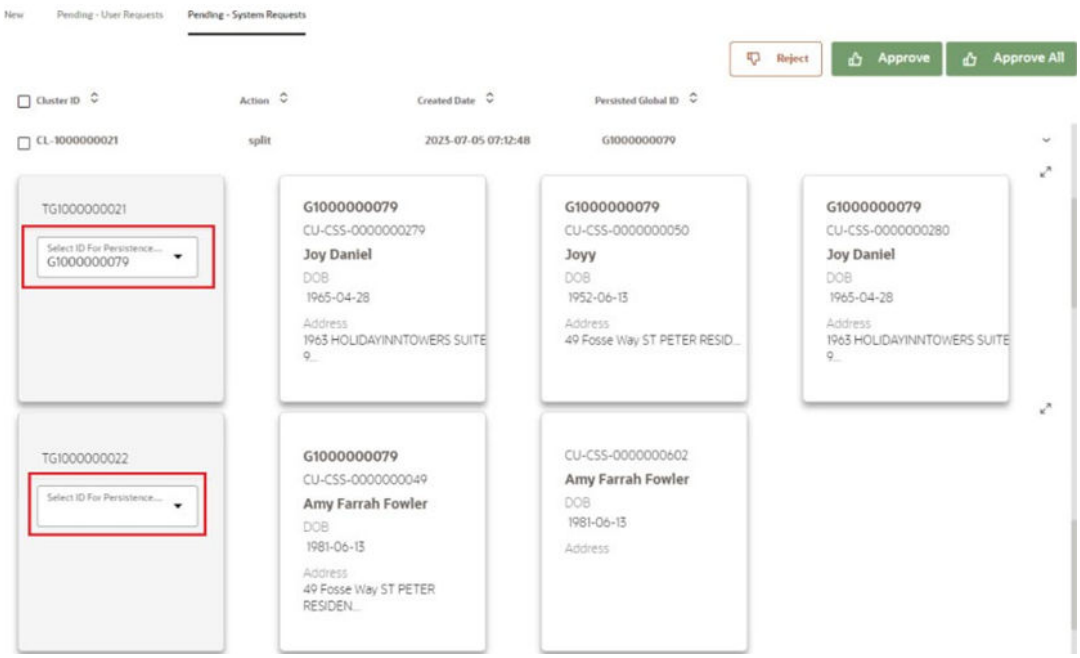
Supervisor can select which global party ids are persisted as part of the manual approval process.

> **Note:**
>
> **Approve/Reject**: Select the particular Cluster ID and click **Approve/Reject** to approve/reject the records.
>
> **Approve All**: Click **Approve All** to approve all the records that are updated by the batch action.

For example, if you want to split multiple records in one Global Party ID then follow these steps:

1. Select the Cluster ID and click **Expanded** arrow to view the cluster details.

Figure 9-26 Split Action



2. Select Global Party ID from the **Select ID for Persistence** drop-down list.

   You can persist Global Party ID in the following methods:

   - By default, the Global Party ID will be persisted for the group that has the most number of entities and if the number of entities are same between the groups, then the group with the least entity id gets the persisted GID. This is the behaviour if **Approve All** is selected and is also the default behaviour if manual approval is not configured.
   - If you want to persist the Global Party ID based on your requirement, then follow these steps:
     - Select blank Global Party ID from the **Select ID For Persistence** drop-down list which need not be persisted.
     - Select Global Party ID from the **Select ID For Persistence** drop-down list which needs to be persisted.

3. Click **Approve** to persist the Global Party ID for the selected cluster.

   A new Global Party ID will be created for the split cluster once the subsequent day's batch is executed.

9.6 Using Provided Notebooks

Compliance Studio has a collection of Notebooks that are available as part of the product.

These notebooks are deployed directly on the Notebook Server but can be exported and loaded as models if required.

9.6.1 Using Financial Crime Graph Patterns Notebook

Financial Crime Graph Pattern notebook provides sample graphs that are created from a synthetic data set of financial transactions.

There are seven different label vertices: CUSTOMER, ACCOUNT, DERIVED ENTITY, EXTERNAL ENTITY, EXTERNAL ADDRESS, CASE, EVENT, and INSTITUTION.

The graph edges are classified into the following categories:

- has event
- entity on
- event on
- is similar to
- is related to
- has account
- cash transaction
- MI transaction
- wire transaction
- send MI/wire
- receive MI/wire
- MI/wire transfer
- has address

This notebook provides patterns such as Transfer of money through an external entity, Transfer of money to an external entity, and High-risk accounts transacting with counterparties in high-risk geographies.

For example, the High-risk accounts transacting with counterparties in high-risk geographies patterns identify the following:

- Accounts that are alerted
- Accounts that have a high activity risk
- Accounts that are also located in high-risk areas

The high-risk geography determination is based on the geographic risk value of the addresses that the related to a counterparty.

9.6.2 Using Example Scenario Notebooks

The RMF scenario notebook is available for you to create notebooks that can be baselined based on these notebooks as an example and create different AML scenarios using different interpreters and be able to push events ECM.

RMF Account (SQL)

The notebook contains the Rapid Movements of Fund (RMF) - Account scenario. The scenario logic is defined in the Oracle SQL query. You can run the RMF Account scenario using the JDBC Interpreter. This interpreter uses Oracle (RDBMS) as the source and target database to create the required temporary tables. Temporary tables are used in the scenario development process.

9.6.3 Creating a Data Discovery Notebook or Model

This topic quickly walk you through building your notebook or model based on the common scenarios explained in the following sections and get a quick insight.

It is difficult to detect and discover the new patterns that criminals execute daily in the real world. These patterns are not just new; they could be used for years and are difficult to identify with traditional techniques.

9.6.3.1 Anti-Money Laundering Pattern Data Discovery

Anti-money laundering (AML) refers to the laws, regulations, and procedures intended to prevent criminals from disguising illegally obtained funds as legitimate income.

The AML Pattern Data Discovery is a scenario that helps you to see new financial crime patterns. These patterns enable you to easily explore any relationship among entities in real-time to extract valuable insights from your data.

This data discovery pattern quickly uncovers emerging, complex money laundering and terrorist financing threats with network and entity generation processes that automatically build network diagrams and reveal hidden relationships. The advanced graph analytics enables entity resolution by looking at multiple data sources and references to a customer, then accounting for inconsistencies, errors, abbreviations, and incomplete records to help determine whether they relate to the same entity.

The scenario includes the following actions that you can perform for the data discovery:

- Creating a Notebook or Model for Your Financial Crime Discovery
- Loading a Financial Crime Graph
- Insight to Customize and Arrange Data in a Graph
- Customizing the Nodes and Edges of the Graph

9.6.3.2 Creating a Notebook for Your Financial Crime Discovery

This topic describes the systematic instructions to create a notebook or model.

The first step in any scenario is to create a notebook or model. In this use case, you can create a notebook to discover your financial crime.

To create a new notebook, follow these steps:

1. Create a Notebook and specify the notebook details.
2. After the notebook is created, create a Paragraph using the PGX Interpreter.

Figure 9-27 Create Paragraph using the PGX Interpreter

In this example, the **PGX Interpreter** is used to create a **PGX query** to load the graph. You must specify the session information that you want to retrieve the data. The query is executed, and the graph details such as graph name, edges, nodes, or vertices are displayed. Further, to understand pattern detection, you must understand the node details, vertices, and edges within the displayed data.

3. For instance, you can specify the matrices, such as the vertices and edges that exist within the data.

Figure 9-28 Specify the Matrices

In this example, the parameters included are:

- SELECT: The `code v.entity,v.jrsdcn, count(*),max(v.rishnb) from g`, fetches the jurisdiction count and risk numbers within the graph **g**, which is built from the previous paragraph.
- MATCH: Matching the vertices.
- GROUP BY: Grouping the entities by vertices and jurisdiction.

4. Click **Execute Paragraph**. The entities' data, such as the jurisdiction, country, and risk numbers, are displayed.
5. You can click the Data Visualization icons within the paragraph to view the details in the visual form to make it more consumable for you.
6. You can hide the code and add a title to the data fetched to make it presentable, consumable, and sharable with others.

   The visualizations displayed are easy to consume, allowing you the flexibility to choose the type of visual. The visualization, in this case, is a **Pie Chart**, which enables you to see the entities associated by their type based on the percentage. You can use the other visualization option to see various types of visuals and different visualization formats.
7. After the first level execution, execute further to understand the relationships that exist between the edges by writing the relationship query.

   For example, the parameters included are:

   - SELECT: The `code e.relationship,e.type,count(*)` from g fetches the relationship count within the graph **g**, which is built from the first paragraph.
   - MATCH: Matching the vertices "v1" and "v2" and the edge is "e". You must write the vertices within parenthesis and edges within brackets to fetch the PGQL query.
   - GROUP BY: Grouping the entities by relationship and type.
8. Click **Execute Paragraph**. The entities' data, such as the jurisdiction, country, and risk numbers, are displayed.

Figure 9-29 Entities Data

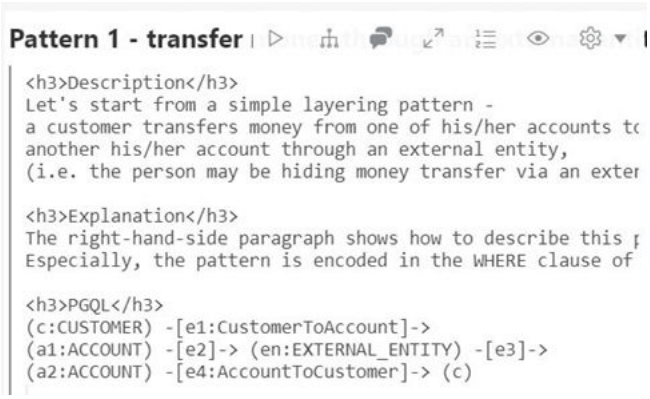
9. You can click the Data Visualization icons within the paragraph to view the details in the visual form to make it more consumable for you. Each paragraph executed is a rest service that allows you to execute based on the updated service. You can hide the code and add a title to the data that is fetched to make it presentable, consumable, and sharable with others.

9.6.3.3 Loading a Financial Crime Graph

To make your pattern even more readable, you can write plain text in a paragraph and then simultaneously write the query in another paragraph and seamlessly produce a readable graph insight. This pattern can be a pattern that an investor or an analyst has provided.
To create a plain text paragraph, follow these steps:

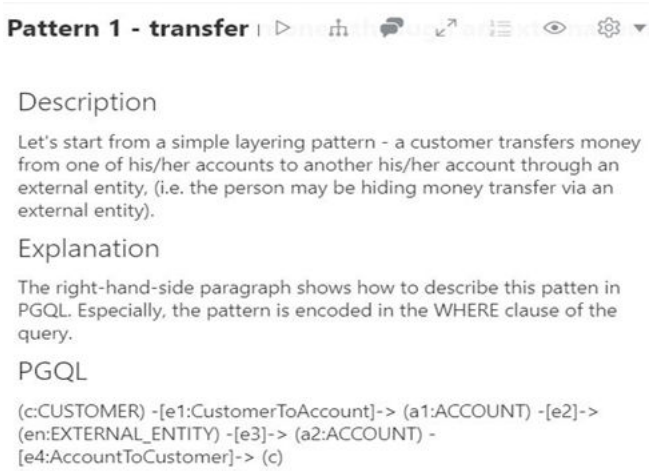
1. Create a paragraph and write the description of the analysis or your use case. For example, see the following figure.

Figure 9-30 Description of Use Case



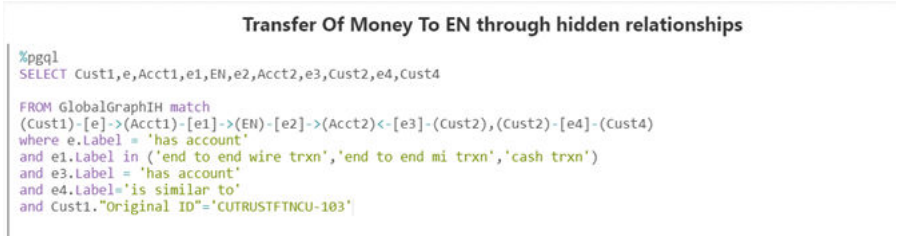
2. After executing the paragraph, the plain text is formatted and displayed as shown below.

Figure 9-31 Plain Text



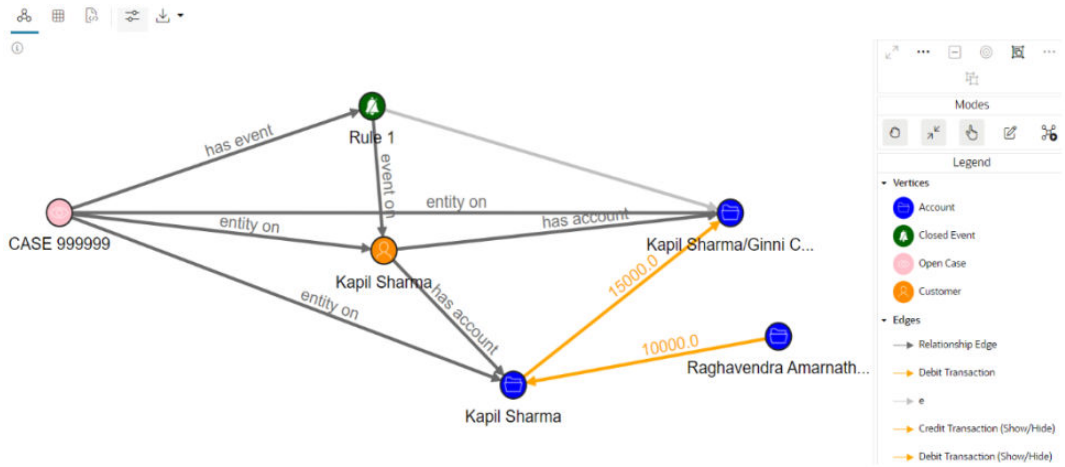
As your next step, create another paragraph to run a PGQL query based on the pattern that you described in the previous paragraph. The code snippet can be provided, as shown in the below figure.

Figure 9-32 Code to Create Another Paragraph



3. Click **Execute Paragraph** to visualize your pattern.

Figure 9-33 Graph Results



You can hide the code and add a title to the data that is fetched to make it presentable, consumable, and sharable with others. You can provide more highlights (such as labels and vertices), rearrange the nodes, and toggle the graph to enrich your graph, and you can open the paragraph in the iframe mode or expand the graph to make it more readable.

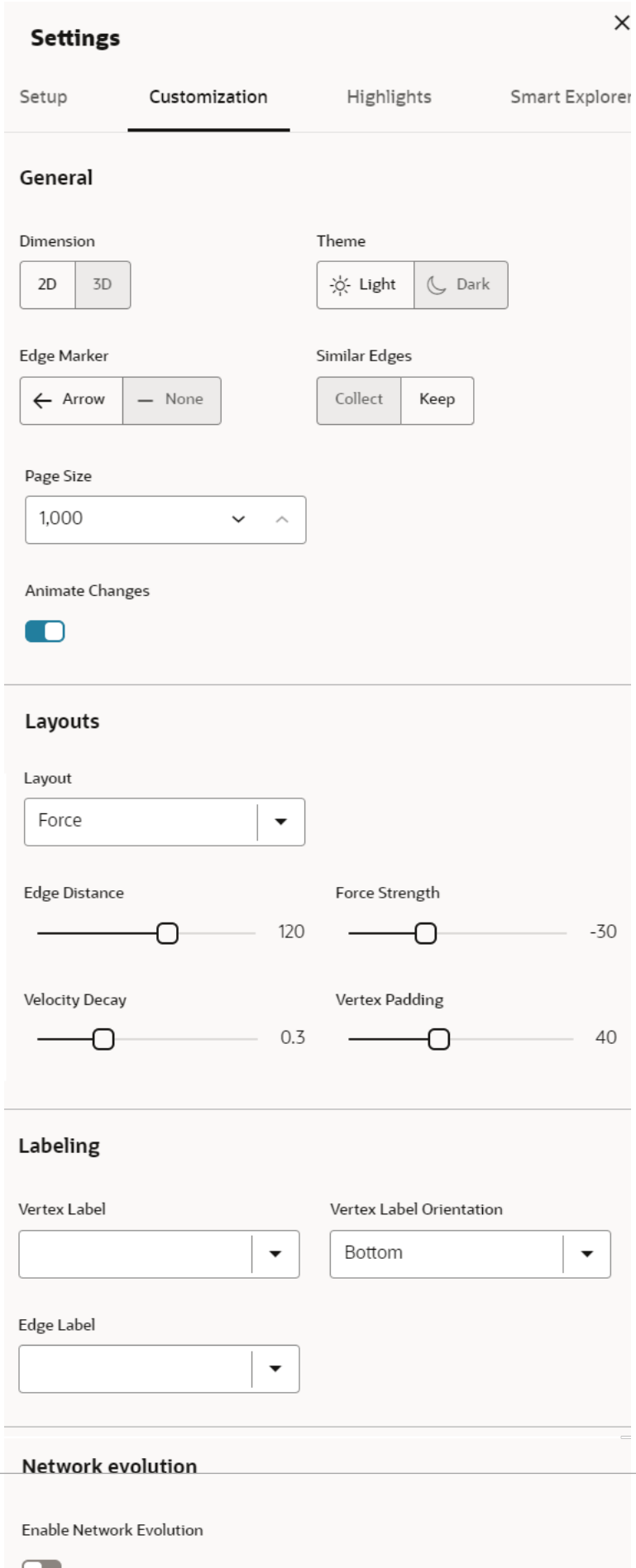
9.6.3.4 Insight to Customize and Arrange Data in a Graph

Using Graph Customization options, you can provide more highlights (such as labels and vertices) to enrich the executed graph.

The graph can be customized based on your visual needs. The graph customization and settings are based on the configuration in the graph settings, such as Setup, Customization, Highlights, and Smart Explorer.

The following figure is customized using Customization in the Graph Settings.

Figure 9-34 Graph Customization



9.6.4 Highlighting the Nodes and Edges of the Graph

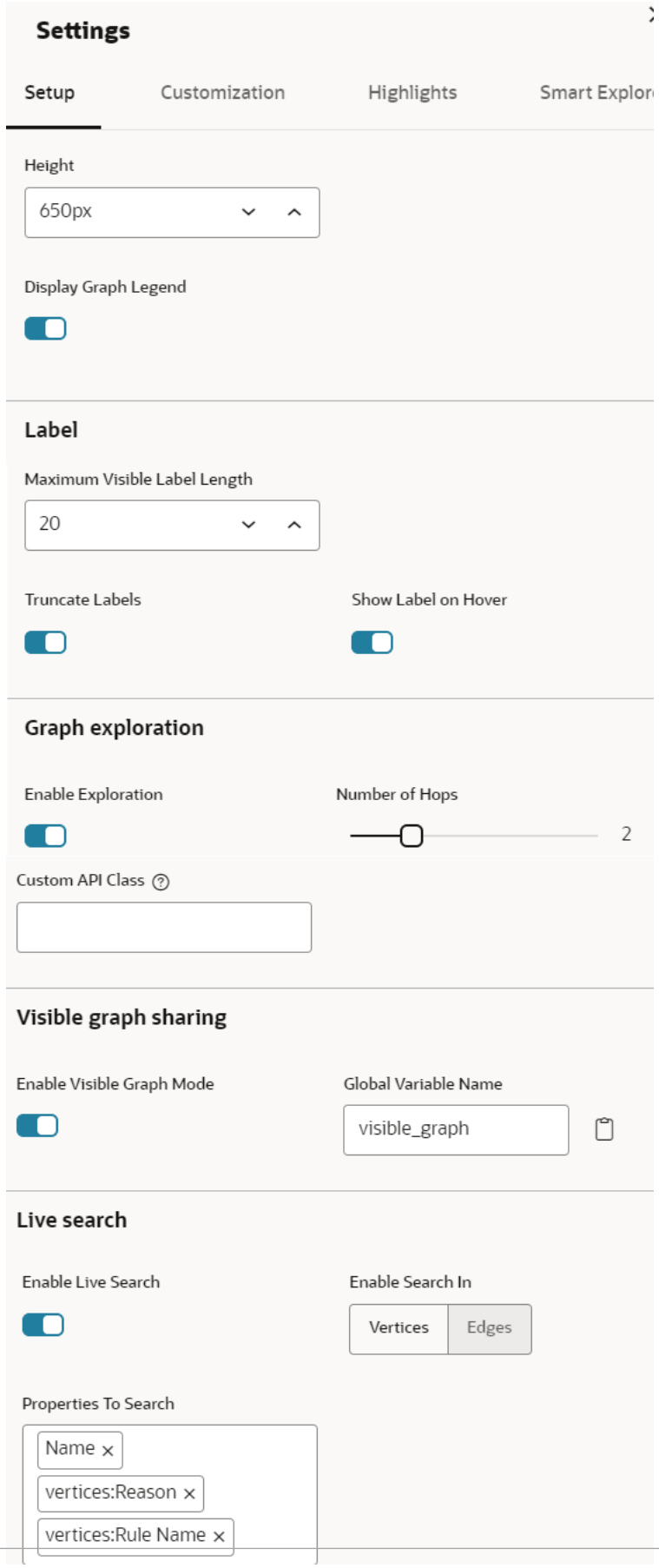
The Nodes and Edges in your graphs can be customized to visualize the data flow between the entities better.

To customize the Nodes and Edges, you can navigate to the graph setting and add new highlights in the **Highlights** tab.

Vertices

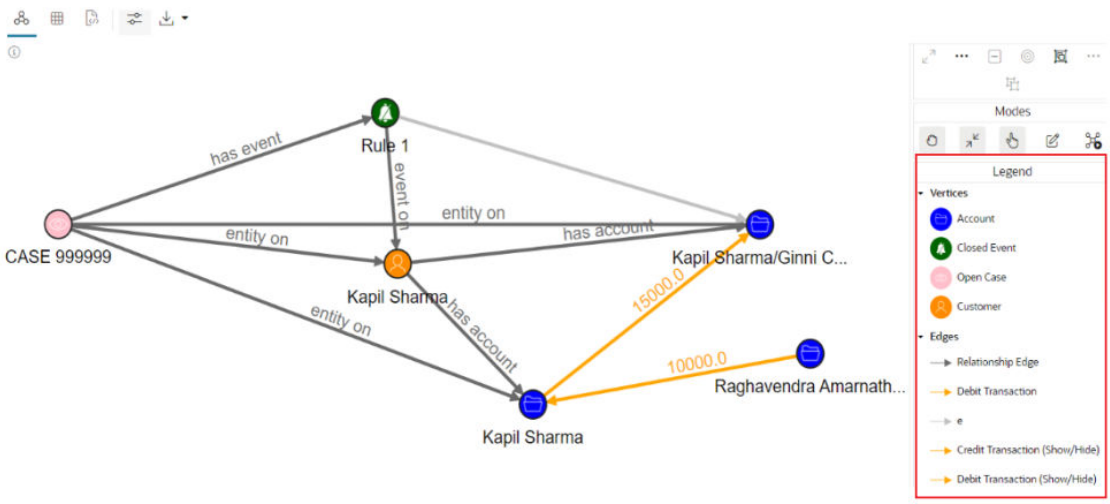
Vertices can be highlighted by clicking the **New Highlight** and specifying the parameters for the highlights. For example, the figure shows how you can differentiate the vertices with different entities such as Customer, Account, and Derived Entity.

Figure 9-35 Vertices Settings



Based on the applied highlights, the graph changes are displayed as follows.

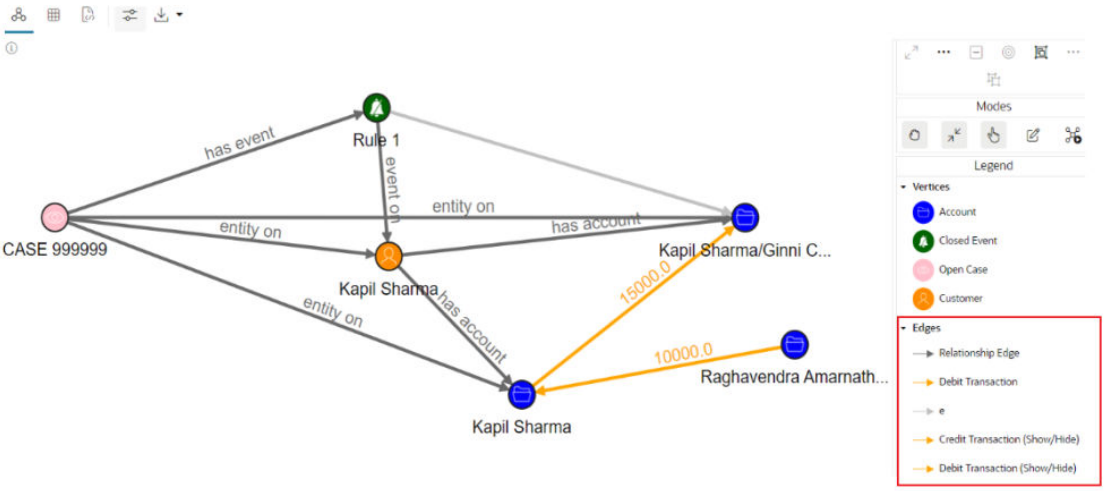
Figure 9-36 Graph Legend



Edges

Edge can be highlighted by clicking the **New Highlight** and specifying the highlights' parameters. Based on the applied highlights, the graph changes are displayed as follows.

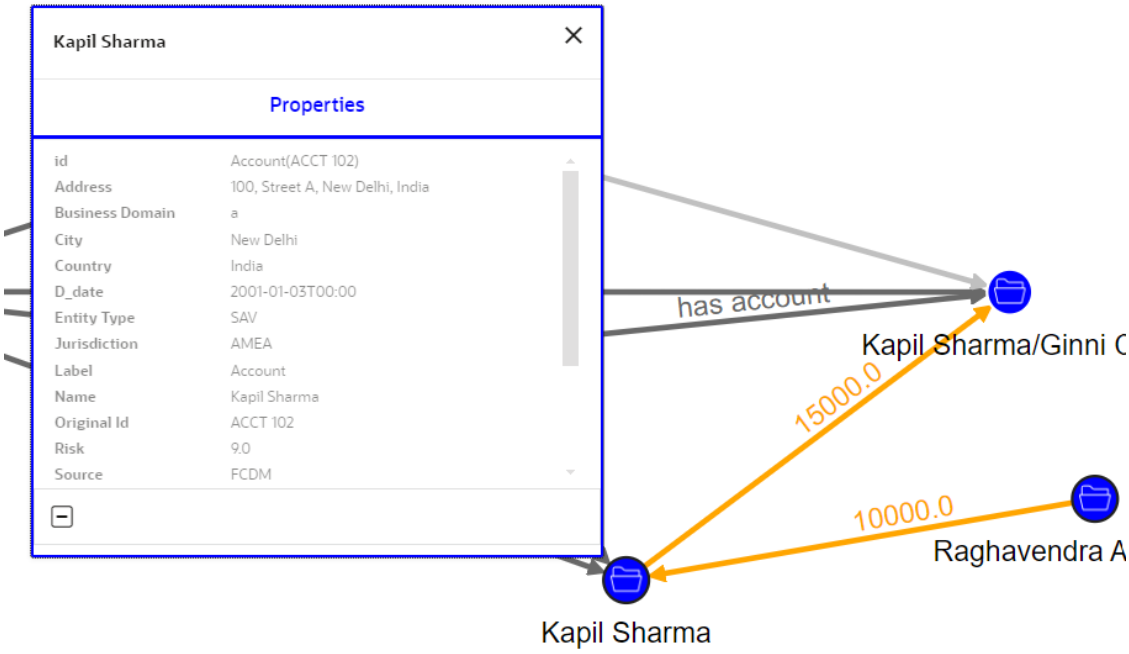
Figure 9-37 Graph Edges



The graphs illustrate a pattern, where there is a money flow from the bank to the different accounts from a customer, with the flow of money highlighted with different colors. Blue represents the money flow from the account to the bank, and red represents the money flow from the bank to the account.

You can also specify the risk numbers to this pattern and visualize your risk. For example, if you use the risk parameters as highlights.

Figure 9-38 Detailed View



By exploring the options available within the graph customization, you can streamline your analysis and present your reports based on your use case.

9.7 Create Event API

The Create Event API generates events at the end of a scenario Notebook or Model for transfer to the Enterprise Case Management Application for correlation into a Case.

It accepts inputs from two tables to generate alerts that are populated in the event tables. The syntax to call the Create Event API is as follows:

```
SELECT FUNC_CREATE_EVENT(,) FROM DUAL;
```

For information on Table_1 and Table_2, see Table 9-12 and Table 9-13.

Table 1 Details

Table 1 contains information related to the direct focus. This information is aggregated for focus. For example, the account number and the total transaction amount for that account will be stored in the Table 1. One record is inserted for each focus (account or customer).

Table 9-12 Table 1 Details

Column Name	Sample Values	Description
V_FOCUS_TYPE	ACCOUNT	The main focus of the scenario. The value can be obtained from the kdd_dataset_base table.
V_ENTITY_NAME	ANDERSON	Name of the focused entity, that is, account display name and customer display name.

Table 9-12 (Cont.) Table 1 Details

Column Name	Sample Values	Description
V_DATA_ORIGIN	AMLBIG	Data origin, that is, MAN, DLY, etc.
V_STATUS_CD	NEW	Event status code. Default value: New.
D_EVENT_CREATED_DATE	3/11/2019 15:02	Event creation date
V_JURISDICTION_CD	AMEA	Jurisdiction
V_BUSINESS_DOMAIN_CD	a	Business domain
V_COMMENTS	-	General comment for the scenario
V_SCENARIO_CLASS	AML	Class to which the scenario belongs to, that is, AML, TC, and BC
V_FOCUS_CD	ACMLTERR ORFINFAC- 005	Code for focus, that is, account internal ID in case of account
V_ENTITY_CD	832	Focus sequence ID, that is, account sequence ID in case of account focus
MAX_DATA_DUMP_DT	12/10/2015	Data Dump date for which business data is involved
MAX_NOT_EFT_TRANS	0	Specific for an account. This column can be ignored for other scenarios.
TOT_TRXN_AMT	10500	
TOT_TRXN_CNT	2	
EFFECTIVERISK	9	
ACTIVITYRISK	20	
RISK_CATEGORY	HR	
ACCT_TYPE	Seasoned	
TRUSTED_PRCTG	0	
PASS_THRU_PRCTG	0	
LRF_PRCTG	0	
HR_PRCTG	0	
ACT_TOT_CREDIT_TRXN_ AMT	5400	Threshold specific for an account. It can be different for other scenarios.
TH_TOT_MIN_CREDIT_TR XN_AMT	50	
TH_TOT_MAX_CREDIT_T RXN_AMT	10000000	
ACT_TOT_CREDIT_TRXN CT	2	
TH_TOT_MIN_CREDIT_TR XNCT	1	
TH_TOT_MAX_CREDIT_T RXNCT	20	
ACT_TOT_DEBIT_TRXNC T	2	
TH_TOT_MIN DEBIT_TR XNCT	1	
TH_TOT_MAX_DEBIT_TR XNCT	20	
ACT_TOT_DEBIT_TRXN_A MT	5100	
TH_MIN_DEBIT_TRXN_A MT	20	

Table 9-12 (Cont.) Table 1 Details

Column Name	Sample Values	Description
REMARK	-	
SCENARIO_CD	10000	
SCENARIO_MN	RMF Account(sql)	Name of scenario notebook
N_RUN_ID	84	The Create Event API automatically creates this value
N_EVENT_CD	366	-
V_ENTITY_CD	25291	Transaction sequence ID
V_FOCUS_CD	ACMLNOA AC- 302	Account Internal ID in case of account focus
N_ENTITY_ID	FOMLNOA AC- 301	Transaction Internal ID
TRXN_DATADUMP_DT	12/10/2015	Transaction Data Dump date
TRXN_PRDCT_TYPE	WIRE_TRXN	Transaction Product type. The value can be obtained from the kdd_dataset_base table.

Table 2 Details

Table 2 details contains information on all the transactions involved in alert generation.

Table 9-13 Table 2 Details

Columns	Sample Values	Description
V_ENTITY_C D	25291	Transaction sequence ID
V_FOCUS_C D	ACMLNOAAC-302	Account Internal ID in case of account focus
N_ENTITY_ID	FOMLNOAAC-301	Transaction Internal ID
TRXN_DATA DUMP_DT	12/10/2015	Transaction data dump date
TRXN_PRDC T_TYPE	WIRE_TRXN	Transaction Product type. The value can be taken from the kdd_dataset_base table.

9.8 Publishing your Notebooks from Notebook Server

Scenario Notebooks can be published from the Notebook interface and published as models using the Model Management Framework. Only one notebook can be in a published state for a given notebook ID. If another version of a notebook is published, the previously published version is replaced.

When a notebook is published:

- The original notebook is cloned, and a published notebook is created.
- Any changes made to the original notebook will have no impact on the published notebook.
- A new version of the published notebook is created whenever the original notebook is republished.
- The published notebook is in a read-only format.
- The published notebook can be run in a batch pipeline.

The notebook toolbar is extended by following conditional buttons after publishing the notebook:

- **Publish**
  Clicking this button will publish the given notebook. The published notebook is automatically loaded.
- **Published Notebook**
  If the notebook has a published version, this button appears and, on click, will redirect you to the published notebook.

The workspace view shows an additional Published column that reflects whether a notebook is published.

> **Note:**
>
> - When Publishing a Notebook, you must ensure there are no empty paragraphs.
> - Published notebooks do not appear in the overview. Published notebooks can only be seen by either knowing the direct link or navigating from a notebook that has a published notebook.

9.8.1 Publishing a Scenario Notebook Directly

After the notebook is approved, you can publish that notebook for your organization to use it. To publish a scenario notebook, follow these steps:

1. Navigate to a **Scenario Notebook** page.
2. Click **Publish Notebook** on the top left corner.

   The **Publish for Approval** dialog box is displayed with the Parameter Keys, and the corresponding Parameter values are added to the paragraphs.

   > **Note:**
   >
   > Ensure that the parameters with the same name must have the same values.

3. Click **Publish**.

   The Scenario Notebook is published for approval and listed in the Notebooks page with the For Approval tag. The published scenario notebook is shared with the user mapped to the DSBATCHGRP group in OFSAA, with the For Approval tag.

9.8.2 Approving a Scenario Notebook Directly

The notebook that is published must be approved by the group administrator.
To approve a notebook, follow these steps:

1. Log in to the Notebook Server application as a DSBATCHGRP user.
2. Navigate to the **Scenario Notebook** page that you want to approve.
3. Click **Approve Notebook** on the top left corner. The Approve Notebook dialog box is displayed.

4. Click **Approve**.

   A confirmation message is displayed to indicate that the notebook is approved. The following scenario and threshold tables are updated in the BD Atomic Schema: *KDD_SCNRO*.
   For an approved notebook, a scenario record is created in the KDD_SCNRO table, and the columns **scnro_id** and **cntry_id** are updated with ML and customer focus values () 113000004), respectively.

   - KDD_TSHLD_SET
   - KDD_TSHLD

   For an approved notebook, the parameters are captured in the *KDD_TSHLD* table. Publishing multiple notebooks creates multiple threshold sets in the *KDD_TSHLD* table: *SCNRO_NB_PUBLISH*.

> **Note:**
>
> In the current release, Notebook execution using Batch is deprecated and will be removed in the future release. It is recommended to use the scheduler to execute the notebook in Batch.
> You can see the example. For more details, see the **Example of Creating a batch in the Scheduler for Notebook Execution** section in the OFS Compliance Studio Administration and Configuration Guide.

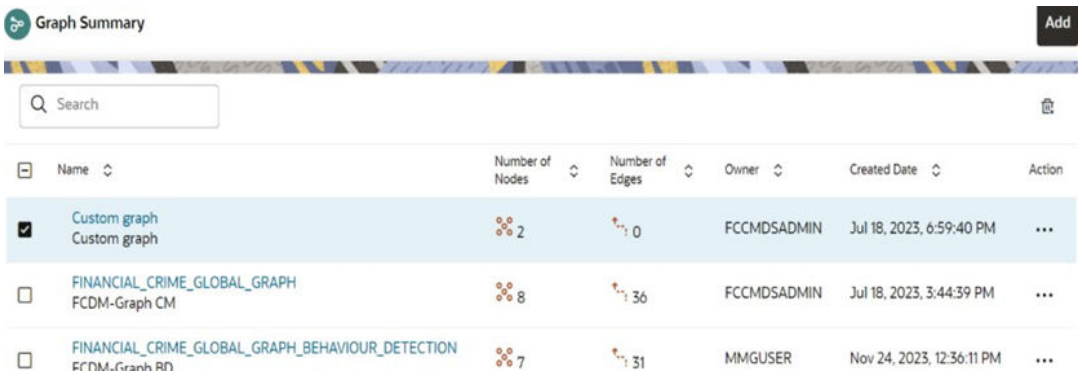
9.9 Updating Graph Definition

This topic describes the systematic instructions to update graph definition.

To update graph definition (run-time parameter and create the batch), follow these steps:

1. Log in to Compliance Studio and navigate to workspace summary.
2. On **Modeling** menu, click **Graphs**. The Graph Summary page is displayed.

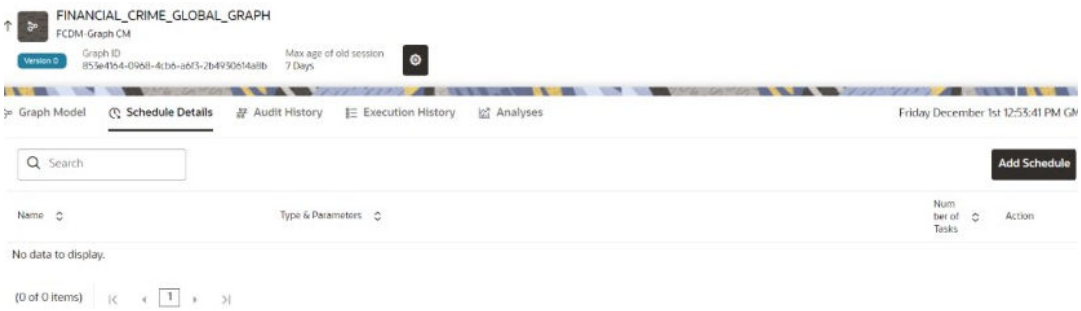
Figure 9-39 Graph Summary



> **Note:**
>
> - The derived entity nodes will be processed (new/updated) and moved to the graph schema only if the date of the graph pipeline execution (START_DATE parameter in the Run time parameters) and FIC_MIS_DATE for the derived entity are the same.
> - Whenever a node is added/deleted from graph pipeline, the population schedule needs to be re-created; so that the graph population schedule would be synchronized with the graph definition with the latest nodes and edges. This step is required to update the respective Scheduler batch-tasks definitions.
> - The Match Ruleset drop-down list will be displayed if it is a Similarity Edge. Any ruleset that is available on this drop-down list should not be edited from the Match Ruleset window.

3. Select the FINANCIAL_CRIME_GLOBAL_GRAPH, and then click **Action** icon.
4. Select **Edit** to view the graph.
5. Click **Schedule Details** tab.

Figure 9-40 Schedule Details



6. To edit the schedule of the graph, click **Edit** under Actions column.
7. Click **Add Schedule** to add a new schedule.

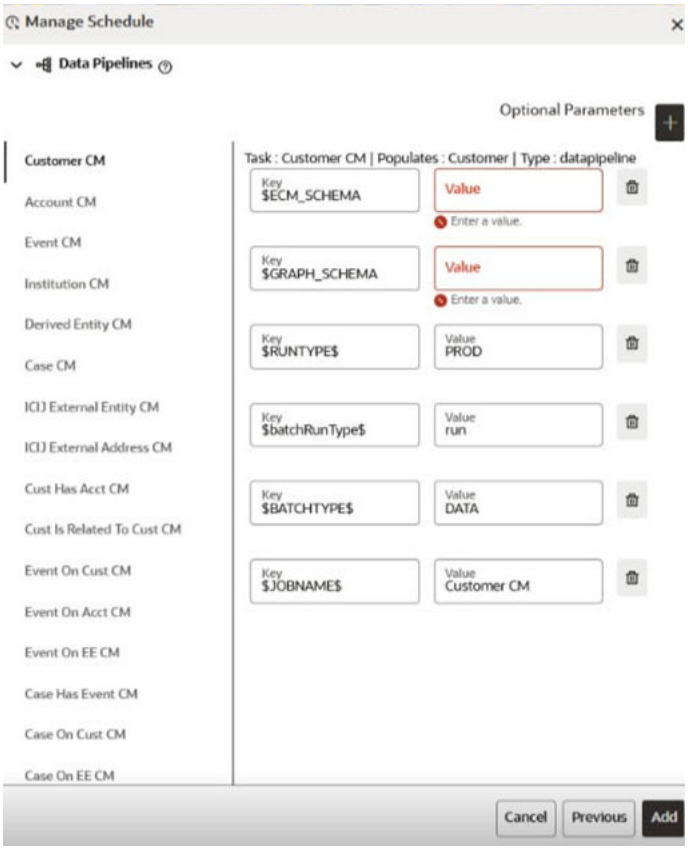
Figure 9-41 Add Schedule Window

8. On the **Add Schedule** screen, select all the tasks and click **Next**.

   For more information on how to select/enter the parameters, see Schedule Details section.

9. Update the value for “$BD_SCHEMA” or “$ECM_SCHEMA”, “$GRAPH_SCHEMA” in the selected task(s) where the value will be respective schema wallet alias. For “$GRAPH_DS”, where the value will be datasource of the graph schema that is attached to the Compliance Studio workspace.

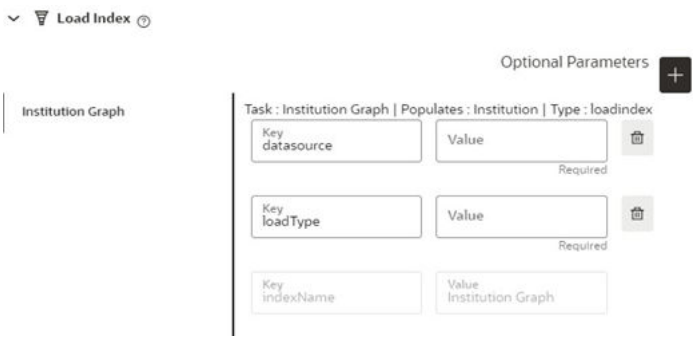
Figure 9-42 Data Pipeline Parameter



> **Note:**
>
> For BD graph, $BD_SCHEMA must be selected as the key value for the optional parameter. Similarly, for ECM graph, $ECM_SCHEMA must be selected as the key value for the optional parameter.

In **Load Index** group,

Figure 9-43 Load Index Parameter



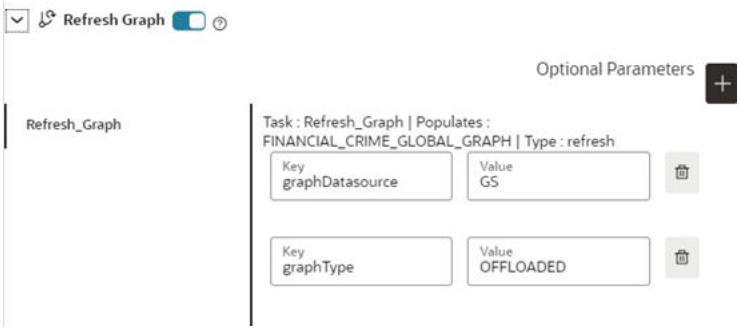
For the Pre-configured load index, respective Key fields will be displayed. You can configure the respective key-value in the Value field as required and specified in the following table.

Table 9-14 Parameter for Load Index

Name	Description	Default Value
datasource	Datasource of the graph schema that is attached to the Compliance Studio workspace.	-
loadType	Execution Load Type. Available types are full load and delta load. **Note**: In this release, only the delta load is supported.	DeltaLoad
indexName	Index name for the graph. **Note**: In this release, only the delta load is supported.	-

10. Enable the **Refresh Graph** option and click **Add** icon to refresh the Graph to add optional parameters.

Figure 9-44 Refresh Graph Parameter



You can configure the respective key-value in the Value field as required and specified in the following table.

Table 9-15 Parameter for Refresh Graph

Name	Description	Default Value
graphDatasource	Datasource of the graph schema that is attached to the Compliance Studio workspace.	-

Table 9-15 (Cont.) Parameter for Refresh Graph

Name	Description	Default Value
graphType	Execution Graph Type. Available types are OFFLOADED and IN-MEMORY. • **IN-MEMORY**: It collects all the data from the nodes and edge tables and it will be transformed into the Oracle Graph Analytics Engine (PGX). • **OFFLOADED**: It creates only metadata of the executed PGQL query instead of loading the entire graph into the PGX server. For more information about **offloaded** graph type, see the Using Off loaded Subgraph section.	OFFLOADED

11. Click **Add**. The schedule is added to respective graph. The batch will be executed based on the schedule type.

    The batch can be executed via Scheduler. For more information, see the Schedule Batch or Batch Groups section.

> **Note:**
>
> For each pipeline execution, inter and delta tables would be created. When the pipeline is executed for the second time, the tables from the first execution should be deleted which is not currently happening. However, the cleanup happens for the second execution on the next day.
> For example,
>
> Cleanup does not work for the first run on 1st- Aug and the second run on 1st Aug.
>
> Cleanup works only when the first run is on 1st- Aug and the second run on 2ndAug.

9.10 Managing Data Pipelines

Data Pipeline is required to transform the source data into the Node and edge that are required in Graph Schema.

The application uses data pipelines to prepare filtered data which can be used to create and run scenarios. Data pipelines prepare data by selecting and joining data sources to create virtual data tables, adding derived attributes to data, running derivations on the data to determine the risk associated with the entity, and so on.

9.10.1 Widgets in Data Pipelines

Depending on the pipeline type, specific widgets are available in the widgets pane of the pipeline.

Table 9-16 Widgets and Descriptions - Data Pipelines

Widget	Name	Description
	Dataset	Use this widget to add a Dataset. Datasets correspond to the contents of a single database table which can be a staging table, business table, or a table that a data pipeline has created. A data pipeline must always begin with a Dataset widget.
	Filter	Use this widget to filter the data in the pipeline to use a subset of the available data records. On applying a filter, all data matching the filter conditions are obtained. This allows you to search and analyze behaviors of interest.
	Join	Use this widget to combine or group multiple tables using various join operators.
	Persist	Use this widget to write data to database tables to be used in other pipelines.
	External Service	Use this widget to add an external service. External Services perform actions on the data, such as loading or moving the data or performing a virus scan.

9.10.1.1 Dataset Widget

The dataset widget enables you to select and filter data sources for use in the later stages of the pipeline.

A data pipeline must always begin with a dataset. Datasets correspond to the contents of a single database table which can be a staging table, business table, or a table that a data pipeline has created. Using the dataset widget, you can select any available staging table, name the dataset, perform DQ (data quality) checks on one, multiple, or all columns of the selected staging table, and filter the output by defining conditions for one, multiple, or all columns of the selected staging table using one of three methods: Expression Builder, Tables, or Text. When multiple columns are selected, the OR logic is applied to filter the outputs.
To create a dataset, follow these steps:

1. Navigate to the Pipeline Designer page.

2. Drag and drop the **Dataset** widget from the widgets pane in the upper-right corner of the designer pane.
3. Hover on the **Dataset** widget and click **Edit**. Provide the details as described in the following table.

Table 9-17 Fields and Description - Dataset

Field	Description
Name	Enter the name for your dataset.
Tables	Select a table from the Tables drop-down list. This list consists of all the available staging tables. The columns of the selected table are displayed in the Attributes pane. The attributes include the Logical Name, Column name, and Column Type.

Table 9-17 (Cont.) Fields and Description - Dataset

Field	Description
Enable DQ check	Select this option to enable the data quality check for the table. You can select each table column, specify checks such as range, length, LOV, and null check, and save the rule after naming it. Based on the rule, checks are performed on the columns of the selected staging table to filter out information you do not require. To specify DQ rules, follow these steps: **a.** Click Add + next to the Enable DQ check option. **b.** Under Master DQ, select one or multiple Primary Key options. All columns of the selected staging table are listed for you to select. **c.** Under DQ Rules, select a column from the Available Columns list. This list contains all columns of the selected staging table. **d.** Enter a rule name for the selected column of the staging table and specify the following checks for this rule: • Range Check DQ Rules. Specify the following range checks: – Is Range Check Required: Select Yes or No. If you select No, jump to the length check rule. If you select Yes, provide a value in the Minimum Value field. – Is Provided Minimum Value Inclusive: Select Yes or No. – Maximum Value: Provide a value in the Maximum Value field. – Is Provided Maximum Value Inclusive: Select Yes or No. • Length Check DQ Rules: Specify Is Length Check Required: Select Yes or No. If you select No, jump to the LOV check rule. If you select Yes, provide a value each in the Minimum Length and Maximum Length fields. • LOV Check DQ Rules: Specify is LOV Check Required: Select Yes or No. If you select No, jump to the Null Check DQ rule. If you select Yes, provide the LOV values in the LOV Values field. • Null Check DQ Rules: Specify the following Null check DQ rules: – Is NULL Check Required: Select Yes or No. If you select No, jump to the Is Null Value Allowed rule. If you select Yes, provide the null default value in the Null Default Values field. – Is NULL Value Allowed: Select Yes or No. If you select No, provide the null default value in the Null Default Values field. **e.** Click Save to save your DQ rule. **f.** Repeat these steps to define DQ rules for all the table columns based on your requirement.

4. Click **Save** to save the changes. The dataset is created and is visible on the canvas. It is also available for use in the Dataset pane.
5. To reuse a dataset you have created, click the **Dataset** icon on the upper-left corner to view the Dataset pane. Click **Expand** to open the list to display the available datasets, including those you have created. Click the dataset name you want and drag it into the canvas of the Pipeline Designer.

> **Note:**
>
> You can perform certain common tasks in all the widgets, such as Edit, Delete, filter, etc. For more information, see Common Tasks section.

9.10.1.2 Creating Persist

The Persist widget enables you to write data to database tables so that it can be used in other pipelines.

It is used to map columns of the source table to a destination table. The Persist widget helps you to map attributes from the input datasets to the target table, which will be stored.
To create a persist, follow these steps:

1. Navigate to the **Pipeline Designer** page.
2. **Drag and drop** the Persist widget from the widgets pane to the designer pane.
3. Hover on the Persist widget and click **Edit**. A dialog box is displayed.
4. Provide the details as described in the following table.

Table 9-18 Fields and Description - Creating Persist

Field	Description
Save As	Enter the name for the Persist widget.
Source Datasets	Displays the list of datasets that are linked to the persist widget.
Target Table	Select the target table to which you want to map the columns in the source dataset tables.

Table 9-18 (Cont.) Fields and Description - Creating Persist

Field	Description
Type	Select the type of mapping that you want to implement for the columns in the target table. The following options are available: • **Full Load**: This option enables you to truncate the existing data in the target table and load it with new data from the source datasets. • **SCD**: This option represents a slowly changing dimension. This option is used to map data from source datasets to the target table with both current and historical data stored in the target table. You can select the following options: – **Surrogate**: Values of this type are typically generated by incremental keys. For example, Sequence IDs. – **Unique**: Use this type for values which are unique across the dataset. For example, Customer Identifiers. – **Type 2**: Use this type for values that may be changed or added to. For example, Customer Names. Values of this type compare both current and historical data to provide the latest record as active. Historical values will be marked inactive. – **Direct**: Use this type for values that should consider only the current data for this record. For example, Data Origin. – **Incremental**: This option is used to map data from the source dataset to a target table incrementally. Incremental mapping adds new entries in addition to the existing data. • **Merge**: This option is used to map data from the source dataset to the target table such that both current and historical data are stored, and incremental data is also stored. • **Delta**: This option allows to identify the delta records of the source by comparing with the available records in the target.
Join	Available only if you have connected multiple datasets. For more information on Joining datasets, see Creating Join section.
Hints	Hints provide a mechanism to direct the optimizer to choose a certain query execution plan based on the specific criteria. Select the Type of SQL Operation from the drop-down list and provide a hint in the Hints field.

5. In the Map pane, follow these steps:

   a. Select the **Source Dataset** from the drop-down list on the left-hand side.

      The columns in the table that are associated with the selected source dataset are listed on the left-hand side.

      > **Note:**
      >
      > The source dataset table is referred to as Source Entity, and the columns in the Source Entity are referred to as Source Column.

   b. Select the **Target Table** on the right-hand side. The columns in the target table are listed on the right-hand side.

> **Note:**
>
> The target dataset table is referred to as the Target Entity. The columns in the Target Entity are referred to as the Target Column

- To Automap, click the **Link** icon. Source and target columns are auto-mapped based on Column Names and Data Types.
- To map source and target columns manually, select a source column and target column, and then click **Expand**.

  > **Note:**
  >
  > You must select columns of the same data type.

  The source column is mapped to the target column. The mapping details are displayed in the table on the right-hand side.
- To add a condition to the target column, click **Add +** and use the **Expression Builder** to create the condition. The result is displayed in the target column on the right pane.
- You can also import source and target columns from an Excel sheet. Click **Choose File** and select the Excel sheet.
- You can also export the mapped source and target columns to Excel using **Export**.

6. Click **Save** to save the changes. The persist is created.

   You can perform certain common tasks in all the widgets, such as Edit, Delete, filter, etc. For more information, see the Common Tasks section.

9.10.1.3 Creating Join

The Join widget enables you to combine or group multiple tables using various join operators.

To create a join, follow these steps:

1. Navigate to the **Pipeline Designer** page.
2. **Drag and drop** the Join widget from the widgets pane to the designer pane.
3. Hover on the Join widget and click **Edit**. A dialog box is displayed.
4. Enter the name in the **Name** field.
5. In the **Output** pane, follow these steps:
   a. Select the required tables from the drop-down lists on the left-hand side and right-hand side that you want to join.
   b. Select the join operators to join the two tables. For more information, see the Join Operators section.
6. Add a Join condition to the join table to save the widget.
7. To add a condition, click **Add +** on the right (Add Group and then Add Condition) and specify rules for the condition. You can add multiple groups and multiple conditions under each Group.

8. Click **Save** to save the changes. The join widget is created.

   You can perform certain common tasks in all the widgets, such as Edit, Delete, Filter, etc. For more information, see the Common Tasks section.

9.10.1.3.1 Join Operators

This topic describes different types of joins available in the data pipelines.

The following types of join operators are available:

- **Inner Join**: The Inner Join selects all rows from both participating tables as long as there is a match between the columns.
- **Left Join**: The Left Join returns all rows from the left table, with the matching rows in the right table.
- **Right Join**: The Right Join returns all rows from the right table, with the matching rows in the left table.
- **Full Join**: The Full Join combines the results of both the left and right outer joins and returns all rows from the tables on both sides.

9.10.1.4 Creating Filters

The Filter widget defines criteria that filter the data in the pipeline to use a subset of the available data records.

The data matching the filter conditions are obtained by applying a filter, which can be used to search and analyze behaviors of interest.
To create a filter, follow these steps:

1. Navigate to the **Pipeline Designer** page.
2. Drag and drop the **Filter** widget from the widgets pane to the designer pane.
3. Hover on the Filter widget and click **Edit**. The Filter pane is displayed.
4. Enter the name for the filter in the **Name** field.
5. Navigate to the **Filter** pane. The **Output** pane is displayed.
6. Configure the filter. For more information, see the Configuring Filters section.

9.10.1.5 Creating External Service

External Service refers to an existing set of services that the Customer can use to derive the risk of certain business entities, configure data movement for case management, create events, etc.

A business entity refers to parameters such as Customer, account, transaction, etc.
To create an external service, follow these steps:

1. Navigate to the **Pipeline Designer** page.
2. Drag and drop the **External Service** widget from the widgets pane to the designer pane.
3. Hover over the External Service widget and click **Edit**. The External Service pane is displayed.
4. Select the external Service from the **Name** drop-down list.

   Based on the external Service selected, the following details are auto-populated:

   - The description for the external Service is auto-displayed in the Description field.

- The corresponding details of the selected external Service are displayed in a table. The details include parameter names and parameter values associated with the External Service.

5. Click **Save** to save the changes. The external Service is created.

   You can perform certain common tasks in all the widgets, such as Edit, Delete, Filter, etc. For more information, see the Common Tasks section.

9.10.2 Creating a New Data Pipeline

You can create a pipeline and then configure the pipeline.
To create a pipeline, follow these steps:

1. Navigate to the **Pipeline Designer** page.
2. Click **Add** in the upper-right corner. The New Pipeline dialog box is displayed.
3. Provide the details as described in the following table.

Table 9-19 Fields and Description - Create a Pipeline

Field	Description
Name	Enter the name for the pipeline.
Description	Enter the description for the pipeline.
Add Search Tags	Enter the keywords for the pipeline. These keywords can be used as search tags while searching for a pipeline. Search tags are also used to group pipelines of the same type. These search tags appear as filters on the Pipeline page.
Type	Select the type of pipeline as either Scenario, Scoring, Data Loading, or Data.

9.10.3 Common Tasks

This topic describes common task to be performed in the data pipelines.

Configuring Filters

Users can configure a filter by defining various filter conditions. For more information on defining conditions, see Defining Filter Conditions.

To configure a filter, follow these steps:

1. Navigate to the **Output** pane.
2. Click **Add** corresponding to the Output pane to open the filter group. The filter group opens where you can add filter conditions.
3. Click **Add** in the Output pane to create a filter condition. The filter condition is displayed.
4. Define the filter condition. For more information, see Defining Filter Conditions.
5. Click **Include** to include the filter in the widget. The filter is included and is functional in the widget.
6. Click **Save** to save the changes. The filter is created. The entire expression that is formed as part of filter creation is displayed in the Output pane.

Defining Filter Conditions

You can define the filter conditions using one of the following:

- Expression Builder
- Tables
- Text

Defining Filter Condition Using Expression Builder

You can form filter conditions using all the operators given in the Expression Builder. The Expression Builder is used to define free flow text filter conditions.

To define a filter condition using the Expression Builder, follow these steps:

1. Click **Exp**. The Expression Builder dialog box is displayed.
2. Select the required Dataset, Attribute, and Runtime Parameters and operators. The resulting condition is displayed in the Condition field.
3. Click **Save** to save the changes.

Defining Filter Condition Using Tables

You can define filter conditions using the various columns of tables. The columns of the two tables are compared with each other using the required operators.

To define filter conditions using tables, follow these steps:

1. Click **Tables**.
2. Select the dataset or risk indicator and column on the left-hand and right-hand sides, and then select the operator.

Defining Filter Condition Using Text

You can define filter conditions using text. A particular column in a table is compared with the input text using the required operators.

To define filter conditions using text, follow these steps:

1. Click **Text**.
2. Select a dataset and column on the left-hand side, operator, and then enter the text in the field on the right-hand side.
3. Click **Save** to save the changes.

9.10.3.1 Creating Runtime Parameter

A Runtime parameter is a variable whose value can be defined and then called from within that same pipeline.

When you define a runtime parameter, you enter the default value to use. You can specify another value to override the default when you create or edit a job that includes a pipeline with runtime parameters.
To create a runtime parameter, follow these steps:

1. Configure a filter. For more information, see Configuring Filters.
2. Define the filter condition using Expression Builder.
3. Click **Add**. The New Runtime Parameter dialog box is displayed.

4. Provide the details as described in the following table.

Table 9-20 Fields and Description - New Runtime Parameter

Field	Description
Name	Enter the name for the runtime parameter.
Datatype	Enter the datatype for the runtime parameter.
Description	Enter the description for the runtime parameter.
Default Value	Enter the default values for the runtime parameter.

5. Click **OK**. The runtime parameter is created.

9.10.3.2 Editing Widget

To modify a widget, follow these steps:

1. Navigate to the **Pipeline Designer** page.
2. Select the widget that you want to modify.
3. Hover on the widget and click **Edit** icon. A dialog box is displayed.
4. Modify the required details.
5. Click **Save** to save the changes. The widget details are modified.

9.10.3.3 Deleting Widget

To delete a widget, follow these steps:

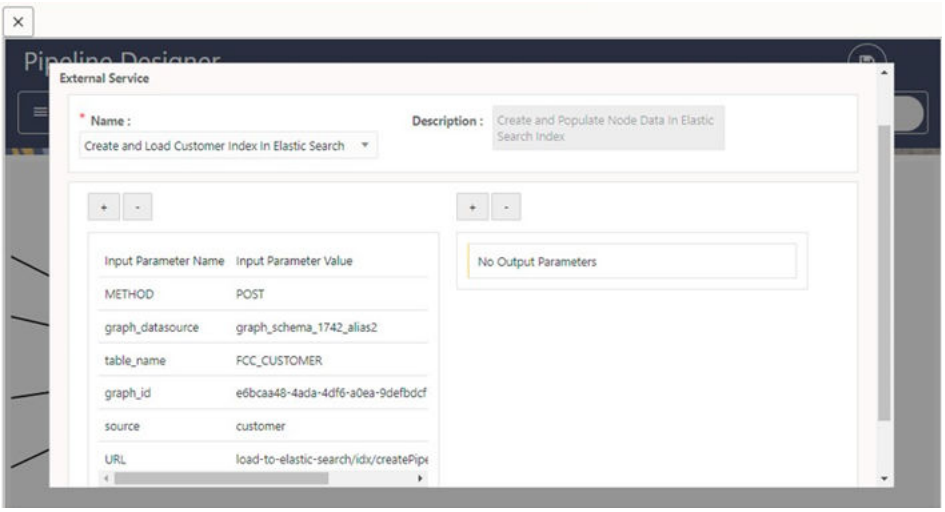
1. Navigate to the **Pipeline Designer** page.
2. Select the widget that you want to delete.
3. Hover on the widget and click **Delete** icon. The Pipeline Delete dialog box is displayed.
4. Click **Confirm**. The widget is deleted.

9.10.4 External Service Widget

This widget is used to add an external service.

It performs actions on the data, such as loading or moving the data or performing a virus scan. The Node information should be loaded into OpenSearch if similar edges are found in the graph between two nodes.

Figure 9-45 External Service



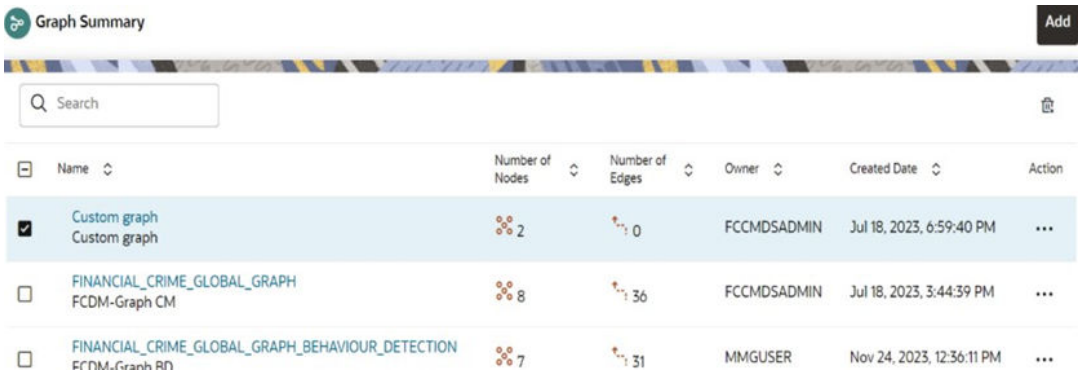
9.10.5 Creating an Index

An index is used to map node attributes and load these data into OpenSearch Indexes.

To create an index, follow these steps:

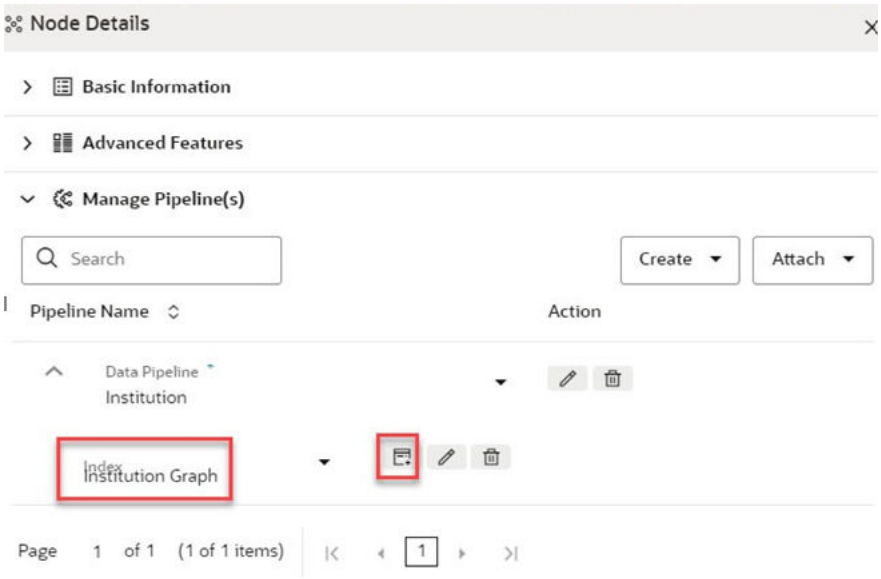
1. Click **Launch Workspace** next to corresponding Workspace to Launch Workspace to display the Dashboard window with application configuration and model creation menu.
2. On **Modeling** menu, click **Graphs**. The Graph Summary page is displayed.

Figure 9-46 Graph Summary



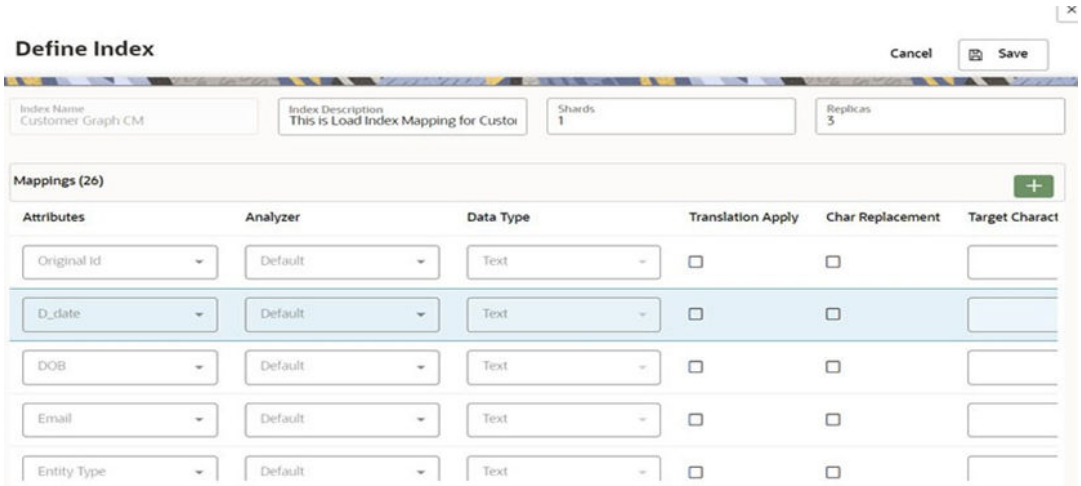
3. Click **Graph**. The sample graph page is displayed.
4. Hover over the graph pipeline and click **Node** on the graph.
5. On the Node Details page, from the **Attach** drop-down list, select **Data Pipeline**. The following page is displayed.

Figure 9-47 Node Details



6. Click **Show Attached Index** drop-down list to view the index. By default, an index is attached to the data pipeline.
7. If you need to create a new index then click **Create** to create the index. The Define Index window is displayed.

Figure 9-48 Define Index



8. Enter the following fields:
   a. **Index Name**: Enter the unique index name across all the graphs. The field is limited to 30 characters.
   b. **Index Description**: Enter the description of the index. The field is limited to 255 characters.
   c. **Shards**: Enter the required integer value. Enter the value as 1. The maximum value for shards is 1024.
   d. **Replicas**: Enter the required integer value. Enter the value as 3. The maximum value for replicas is 1024.

9. Click **Add** icon to add mapping to the index. The Mappings section is displayed.
10. Provide the details as described in the following table.

Table 9-21 Fields and Description - Mappings

Field	Description
Attributes	Select the Attributes from the drop-down. It lists all the attributes which are created in the node modeler attribute. **Note**: To create an index, you should select the attribute which is defined as Key Attributes in the Advanced Features group of the Node Details page.
Analyzer	It should be set as default for key attributes. For other attributes, you can select from the drop-down. The available options are: • DEFAULT • ADDRESS • ORGANIZATION • NAME • NAMESTOP **Note**: If you want to seed a new value in the Analyzer Type, then enter the value as FCC_IDX_JSON_CONTENT in the metadata table.
Data Type	It should be set as default for key attributes. For other attributes, you can select from the drop-down. The available options are: • DATE • NUMBER • TEXT **Note**: If you want to seed a new value in the Data Type, then enter the value as FCC_IDX_JSON_CONTENT in the metadata table. Data Type is set to default as "Text" and non-editable when Analyzer type is selected as "Default".
Translation Apply	Enable this option for language translation, if required.
Char Replacement	Enable this option for character replacement in the OpenSearch while matching, if required.
Target Characters	This filed is applicable only, if char Replacement field is enabled. Enter the target characters which have to be modified.
Replace With	This filed is applicable only, if char Replacement field is enabled. Enter the required characters to be replaced which is defined in the target characters.
Delete	Click Delete to delete the selected mapping, if required. **Note**: If the attribute is deleted in the Node Modeler Attributes Section of the Node Details page, then same attribute must be deleted in the Mappings section of the Define Index page.

11. Click **Save**. The index is created with mapped node attributes to load data into OpenSearch Indexes.
    - If you want to edit the index, click **Edit** icon. The Define Index page is displayed.
    - Edit the required fields and click **Save**. The index is saved.

> **Note:**
>
> You cannot edit the **Index Name** field.

- If you want to delete the index, click **Delete** icon. The index will be removed from the respective data pipeline.

A
Appendix

Topics:

A.1 Common Screen Elements in the Notebook

This topic describes common screen elements in the notebook that can be used to perform quick actions when preparing and executing the notebooks.

Table A-1 Elements in the Notebook

Icon Name	Action/Description
Modify Notebook 	Modifies the details of a notebook, such as a name, description, and (or) tags.
Hide Code 	Hides or shows the Code Section in all the paragraphs in a notebook.
Hide Result 	Hides or shows the Results Section in all the paragraphs in a notebook.

Table A-1 (Cont.) Elements in the Notebook

Icon Name	Action/Description
Re-initialize session 	It is a Compliance Studio add-on feature. Re-initializing the session internally first "invalidates the session," followed by the recreation of a new session and setting/initializing specific variable values. These variables are utilized in ASC and other FCC use cases. **Note**: • It is recommended to use the "Re-initialize session" instead of "Invalidate session" for ASC and FCC use cases in the Out-of-the-box notebooks. • By default, the Start Widget gets executed automatically on clicking the Re-initialize icon. All other paragraphs need to be executed manually.
Read-Only 	Sets the notebook to read-only mode. **Note**: The notebook is protected from edit, clear result, delete, share, reset session, and run paragraphs in Read-only mode.
Write 	Set the notebook to write mode.
Run Paragraphs 	Executes all the paragraphs in a notebook in sequential order.
Invalidate Session 	It is a base Data Studio feature that kills any existing session when a notebook/paragraph is executed subsequently and then, a new session is created on-demand.
Delete Notebook 	To delete the selected notebook.
Clear Result 	Clears results for all the paragraphs in a notebook. **Note**: This action clears all the results. You must rerun the paragraphs to view the results.

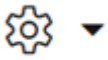
Table A-1 (Cont.) Elements in the Notebook

Icon Name	Action/Description
Clear Paragraph Dependencies 	Remove all defined paragraph dependencies.
Open as an iframe 	Opens a notebook in an Iframe. This allows a notebook to be embedded inside another webpage.
Share Notebook 	Share the notebook with another user, user group, or role.
Clone Notebook 	Creates a copy of a notebook. All paragraphs in the current notebook are replicated in the new notebook. The cloned notebook is created with the default name, Copy of <Current Notebook Name>.
Export Notebook 	Export the notebook to your computer as a DNSB file.
Print Notebook 	To print a notebook in the PDF format and save it in your local machine.
Default Template 	To apply the overall look and feel of the notebook using the default template.

Table A-1 (Cont.) Elements in the Notebook

Icon Name	Action/Description
Layout 	Sets the preferred layout. The available layouts are Zeppelin and Jupyter.
Default Template 	Applies the overall look and feel of the notebook using the default template.
Show Panel 	Shows or hides the Paragraph Settings Bar Commands, Results Toolbar, and Settings Dialog for a selected paragraph in a panel to the notebook's right.
Attach Credentials 	You can use this option to attach credentials (wallet and password) to the notebook to enable secure data access.
Versioning 	You can use this option to create versions for your notebook, which helps you analyze the changes based on the version control.
Run Paragraph 	To execute the code or query in a paragraph.
Enter Dependency Mode 	To add or remove dependent paragraphs. Paragraphs with dependent paragraphs are executed in the dependency order.

Table A-1 (Cont.) Elements in the Notebook

Icon Name	Action/Description
Comments 	To add comments to a paragraph.
Expand 	To expand a paragraph and view the paragraph in full-screen mode.
Show/Hide Line Numbers 	To show or hide line numbers in the code in a paragraph. **Note**:This button is applicable only to the code section.
Visibility 	To manage the visibility settings in a paragraph. It controls how a paragraph may be viewed by the author and other users who have access to the notebook.
Settings 	Users can perform the following actions: • Resize the width of a paragraph • Change the order of placement of the paragraphs by moving them up or down • Run the all above or all below paragraph from the selected paragraph. • Clear the paragraph result • Open the notebook in Embedded window • Disable the run action • Clone the paragraph • Delete a paragraph

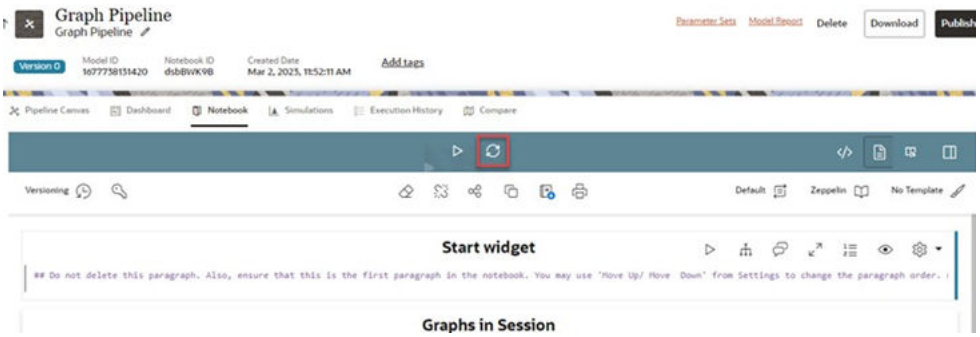
A.2 Managing Resource Utilization

Users can invalidate the session of that notebook and free up the system resources, if any of the Notebooks are no longer needed.

To invalidate the session, follow these steps:

1. Navigate to the **Workspace Summary** page and launch the workspace.
2. On **Modeling** menu, click **Pipelines**.
3. Open the Model in **Pipeline Designer** and click **Notebook** tab to view the notebook.

Figure A-1 Invalidate Session



4. Click **Invalidate Session** icon to invalidate the session.

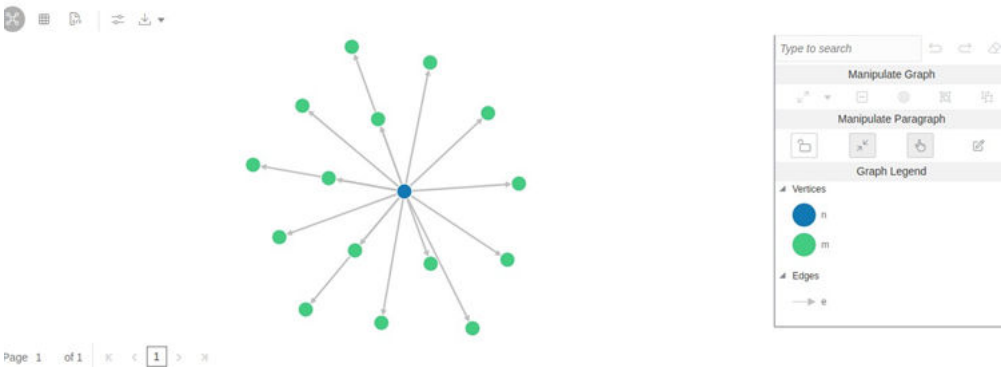
A.3 Using Data Visualizations

This topic describes data visualizations that are available in the Notebook Server.

Graph Visualization

The Notebook Server allows visualizing data in the form of a Graph Visualization. All actions of the graph visualization are collected inside a panel.

Figure A-2 Graph Visualization

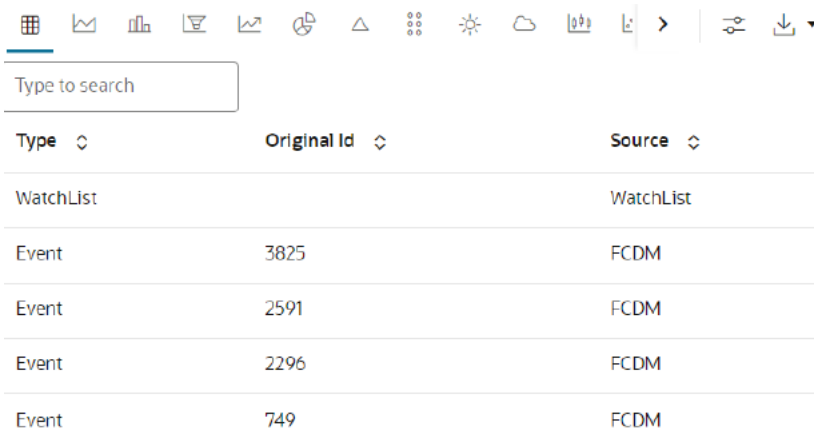


The panel can be re-sized and includes two states, default and minimized panel. These panels enable you to manipulate paragraphs and graphs and provide options to customize your graph.

Table

The Notebook Server allows you to visualize your data in the form of a Table Diagram. The table can be sorted by column in ascending or descending order. Additionally, the table can be filtered for a specific search term. Rows that do not contain this term are hidden from view, and the remaining rows highlight the location of the search term within the row.

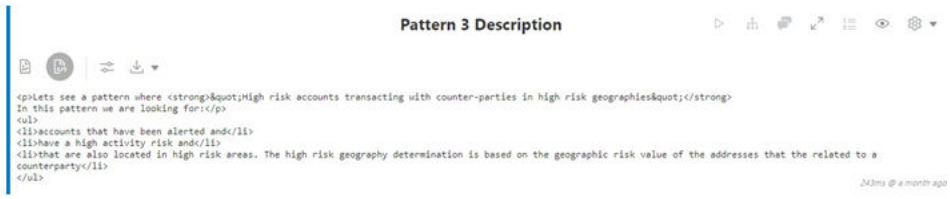
Figure A-3 Table



HTML or Markdown

The Notebook Server allows you to visualize your data in the form of HTML or Markdown text.

Figure A-4 HTML Text



Text

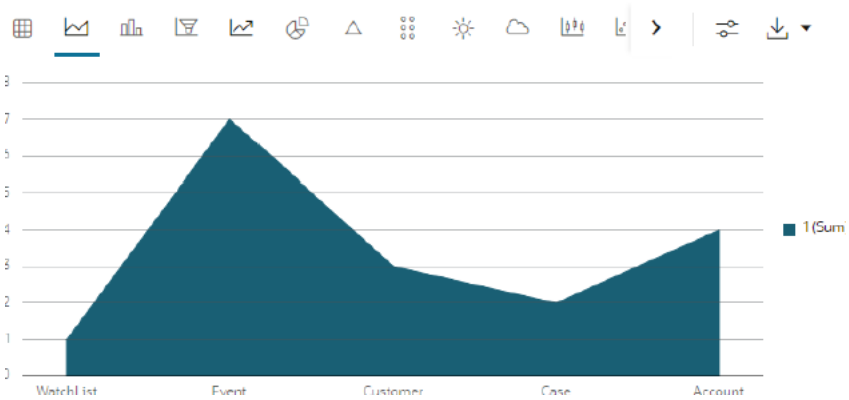
The Notebook Server allows you to visualize your data in structured and presented text for you to customize the data.

Figure A-5 Text

Area Chart

The Notebook Server allows you to visualize your data in the form of an Area Chart.

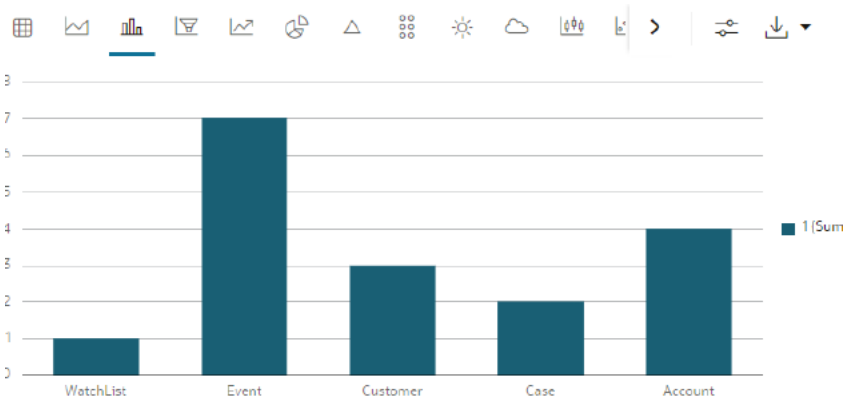
Figure A-6 Area Chart



Bar Chart

The Notebook Server allows you to visualize your data in the form of a Bar Chart.

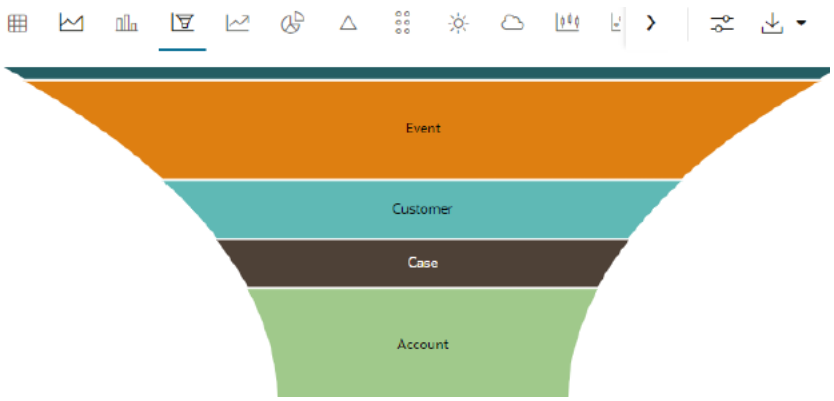
Figure A-7 Bar Chart



Funnel Chart

The Notebook Server allows you to visualize your data in the form of a Funnel Chart.

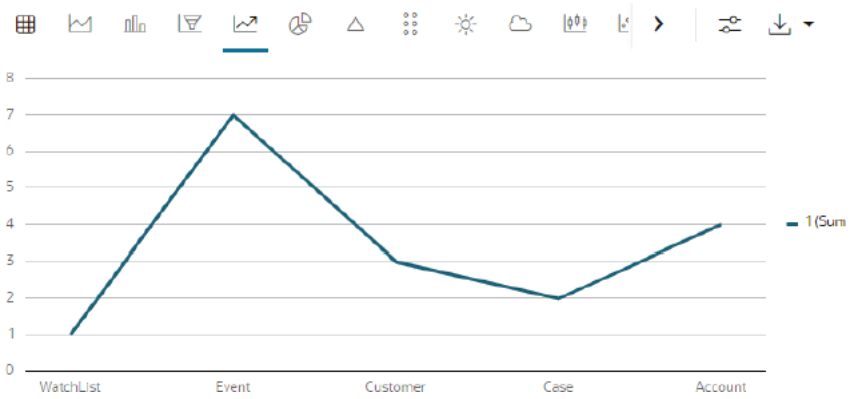
Figure A-8 Funnel Chart



Line Chart

The Notebook Server allows you to visualize your data in the form of a Line Chart.

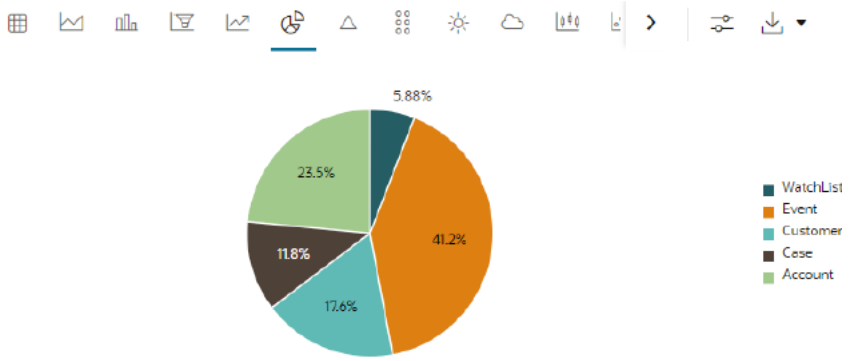
Figure A-9 Line Chart



Pie Chart

The Notebook Server allows you to visualize your data in the form of a Pie Chart.

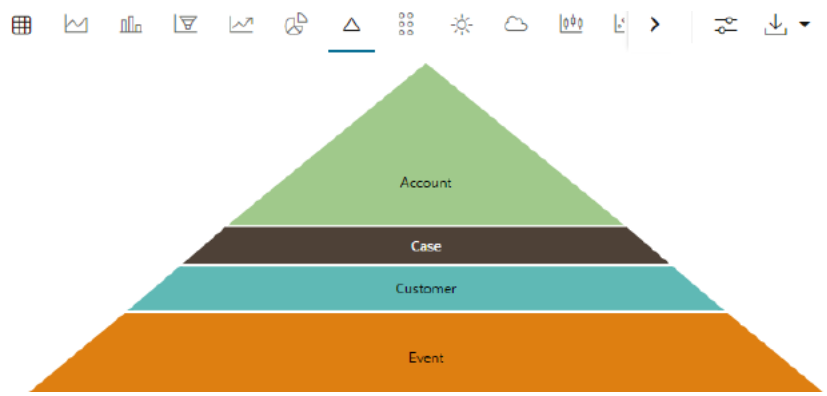
Figure A-10 Pie Chart



Pyramid Chart

The Notebook Server allows you to visualize your data in the form of a Pyramid Chart.

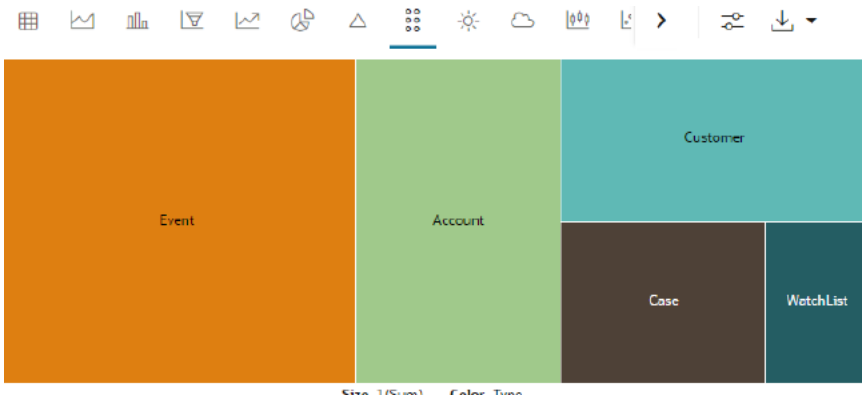
Figure A-11 Pyramid Chart



TreeMap Diagram

The Notebook Server allows you to visualize your data in the form of a Tree Map Diagram.

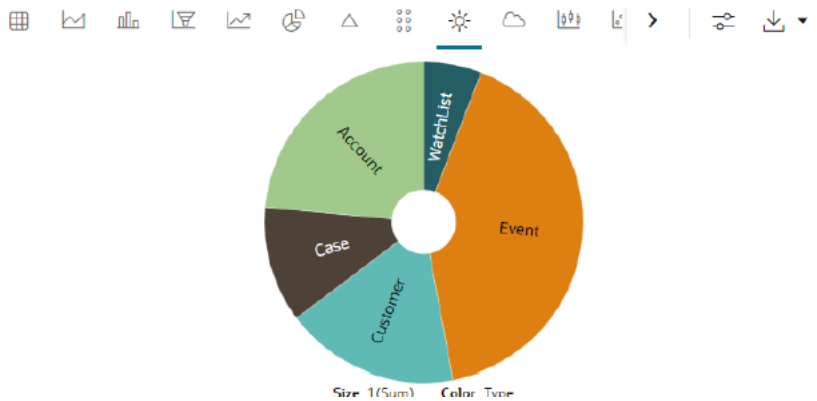
Figure A-12 Tree Map Diagram



Sunburst Diagram

The Notebook Server allows you to visualize your data in the form of a Sunburst Diagram.

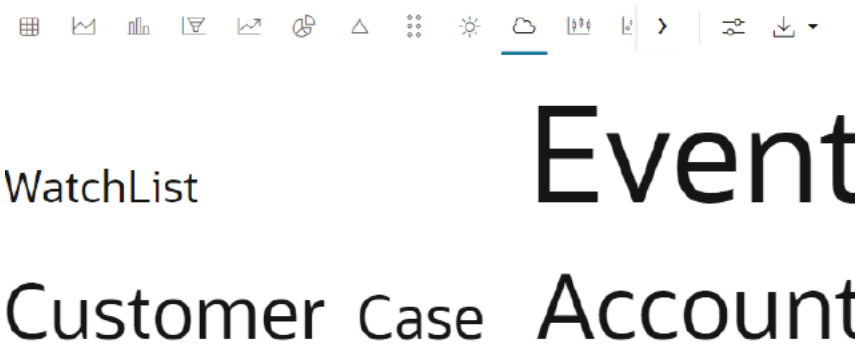
Figure A-13 Sunburst Diagram



Tag Cloud

The Notebook Server allows you to visualize your data in the form of tags. The tag cloud operation is used to identify the spots where there are more flags.

Figure A-14 Tag Cloud



Box Plot

The Notebook Server allows you to visualize your data in the form of a Box Plot.

Figure A-15 Box Plot

Scatter Plot

The Notebook Server allows you to visualize your data in the form of a Scatter Plot.

Figure A-16 Scatter Plot

Map Visualizer

The Notebook Server allows you to visualize your data on top of a Map.

Setting up the Visualizations

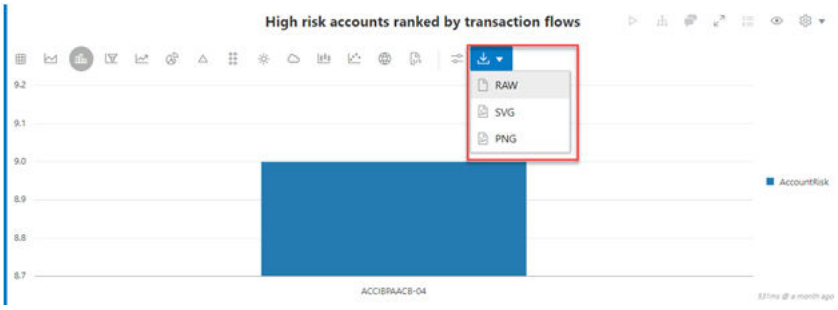
The visualization settings offer several ways of configuration, which are split into three tabs: General, Visualization, and Text. To set the configuration for your table view, click **Settings**.

These tabs in the Settings window allow you to customize the properties of the text, graphs, and visuals and view the visuals in various orientations. For example, you can view the Bar Chart to change the colors, bar type, and 360-degree view of the visuals. The Settings window is displayed and contains the following category:

- General
- Visualization
- Text

Download As

The displayed visuals can be downloaded using the **Download As** option. The available download formats are RAW, SVG, and PNG.



A.4 Dynamic Forms

This topic provide details about how you can specify the dynamic parameters for the existing fields.

Textbox

The textbox offers a simple user input that accepts any characters as follows:

- The name of the user input can be defined using `${name}`.
- A customized label can be set using `${name(Label)}`.
- A default value can be assigned using the equals operator `${name(label)=Default Value}or ${name=Default Value}`.
  For example, `${name(Label)=Default Value}`.

Select

A select field allows you to choose from a given selection as follows:

- The name, label, and default value can be set like the Textbox.
- Options follow the "`,(comma)`" after the default assignment `name=default,default|other1|other2` etc.
- The label of an option can be added in the brackets following the option `other1 (Other Option 1)`.
  For example: `${name(Label)=default,default(Default Option)|other1(Other Option 1)|other2}`.

Slider

A slider allows you to choose from a given range as follows:

- The first part of the slider dynamic form is the keyword `slider`, followed by brackets containing the values for minimum, maximum, and step size: `slider(<min>,<max>,<step size>)`.
- Then, after a: (colon), the name, label, and default value can be set similar to the Textbox.
  For example: `${slider(0,30,1.5):name(Label)=3}`.

Checkbox

A checkbox allows you to choose from a given selection as follows:

- The first part of the checkbox dynamic form is the keyword `checkbox`, followed by brackets containing the values join parameter of the options: `slider(<join parameter>)`.
- The join parameter is used between all selections and only shows up if at least two elements are selected.
- Then, after a : (colon), the name, label, default value, and options can be set similar to the Dynamic Forms.
  For example: `${checkbox(or):name(Label)=default,default(Default Option)|other1(Other Option 1)|other2}`.

Date Picker

A date-picker allows you to choose from a given date as follows:

- The first part of the checkbox dynamic form is the keyword `date`, followed by brackets containing the output format: date `(day)` or `date(yyyy-MM-dd`.
- The name, label, and default value can be set similar to the Textbox.
  For example:`${date(yyyy-MM-dd):name(Label)=2019-02-27}`.

Time Picker

A time picker allows you to choose from a given time as follows:

- The first part of the checkbox dynamic form is the keyword `time`, followed by brackets containing the output format: `time(HH:mm)` or `time(HH)`.
- The name, label, and default value can be set similar to the Textbox.
  For example: `${time(HH:mm):name(Label)=T13:30}`.

DateTime Picker

A datetime picker allows you to choose from a given date and time as follows:

- The first part of the checkbox dynamic form is the keyword `dateTime`, followed by brackets containing the output format, that is, `dateTime (YYYY-MM-dd hh:mm)`.
- The name, label, and default value can be set similar to the Textbox.
  For example: `${dateTime(yyyy-MM-dd hh:mm):name(Label)=2019-02-28 11:30}`.

Setting Values through a Notebook URL

In the web interface, dynamic form values can also be set by passing a name-value pair as the parameter of a notebook URL.

For example, to set the value of the dynamic form named `myDynamicForm` to `FCCStudio` , the parameter `&form_myDynamicForm=FCCStudio` must be added to the URL. If there are multiple dynamic forms with the same name in the same notebook, all of them will be set.

> **Note:**
>
> It is recommended not to pass sensitive information via the URL. If you have sensitive information, you can use the REST API to change the values of a dynamic form.

A.5 Quantifind Risk Report

Quantifind integration for real-time risk reports in Notebook Server enables financial institutions to discover revenue drivers and risk signals, including fraud and money.

> **Note:**
>
> These reports are only available for Notebook Server integration with ECM.

The results are displayed as a card, which displays the risk status of the identified node details. Based on the risk, you can perform the required measures for the risk analysis.

A.6 Using Off loaded Subgraph

This topic describes instruction to load graphs into the in-memory graph server.

There are two options for loading graphs into the in-memory graph server (PGX).

- **In-memory**: It collects all the data from the nodes and edge tables and it will be transformed into a single global graph in the Oracle Graph Analytics Engine (PGX).
- **Offloaded**: It collects all the data from the nodes and edges and keeps it ready to be loaded in Oracle Graph Analytics Engine (PGX) so that users can load subgraphs while doing graph analytics.

The following are the advantages of using subgraph:

- This feature provides a much more flexible way to declare the scope of the data to be loaded into PGX server.
- For global graphs with a high volume of records, the PGX server may need to be very large to hold the full global graph. Subgraph loading allows for many smaller graphs to be loaded which requires a smaller PGX server size.

Creating and defining Graph Pipelines

- For information on how to create a graph, see the Graphs section.
- For information on how to define the graph, see the Updating Graph Definition section.

A.6.1 Loading an Initial Subgraph

This topic is an example of loading a subgraph for a given use case and this can be replicated depending on the users use case for the subgraph.

You can load the subgraph by executing PGQL query in the notebook.
To load the subgraph, follow these steps:

1. Click **Launch Workspace** next to the corresponding Workspace to Launch Workspace to display the Dashboard window with application configuration and model creation menu.
2. Navigate to `<OFS_COMPLIANCE_STUDIO_INSTALLATION_PATH>/deployed/mmg-home/ mmg-studio/server/builtin/notebooks` directory and download the following notebooks:

   `Sub Graph_Dynamic Graph Loading.dsnb`

   `Sub Graph_Dynamic graph loading from OOB graph.dsnb`
3. On **Modeling** menu, click **Pipelines**.
4. Create an Objective to import the notebook.

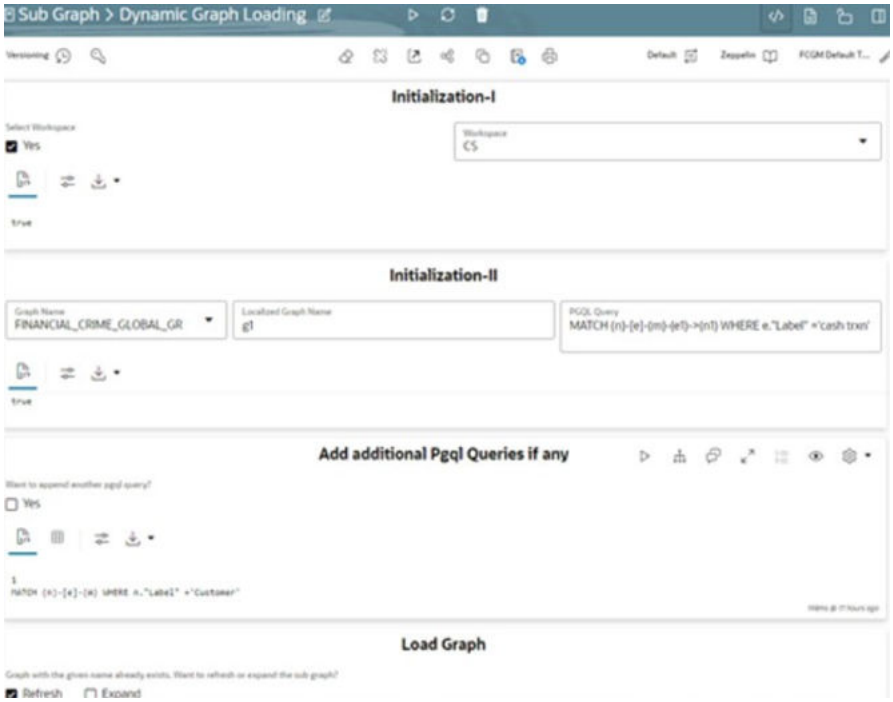
A.6.1.1 Dynamic Graph Loading

This notebook loads graphs from the property graphs created by graph pipeline execution.

To execute the PGQL query, follow these steps:

1. Click **Dynamic Graph Loading** on the Sub Graph notebooks to execute the PGQL query.

Figure A-17 Dynamic Graph Loading



2. In the Initialization-I paragraph, select **Yes** in the **Select Workspace** check box to list the available workspaces where graph pipelines are executed.
3. In the Initialization-I paragraph, select **Workspace** from the drop-down list to get the list of graphs executed in that workspace.
4. In the Initialization-II paragraph, select **Graph Name** from the drop-down list that lists the property graphs executed from graph pipelines.
5. In the Initialization-II paragraph, enter the **Localized Graph Name**.

   Localized graph name is the name with which the subgraph will be created. This will be the name for all references of the subgraph in the session.

   The graph will be loaded with the name entered in this textbox.
6. In the Initialization-II paragraph, enter the **PGQL Query** in the textbox.

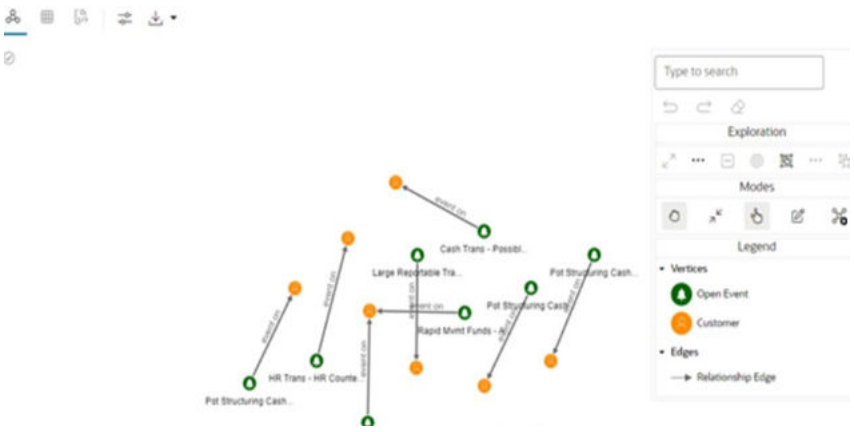
   The graph will be loaded from the entered pgql query. For information on sample queries, see the Dynamic Graph Loading from OOB Graph section.
7. In Add additional Pgql Queries if Any paragraph, select **Yes** in the **Want to append another pqgl query?** check box to load graph from more than one pgql query.

   A textbox is displayed to enter the PGQL query if the option **Yes** is selected. Rerun the paragraph until the number of pgql queries has to be appended.
8. In the Load Graph paragraph, the graph is loaded based on the PGQL query execution. If the graph with the same name already exists in the current session, a check box is displayed with the following options:
   - **Refresh**: This option will drop the graph and load it with the new graph.
   - **Expand**: This option will expand the existing graph with pgql queries added from the above paragraphs.

> **Note:**
>
> If any field is modified, then you need to **re-execute** the paragraph.

The following figure shows the sample graph based on the executed PGQL query.

Figure A-18 Sample Graph

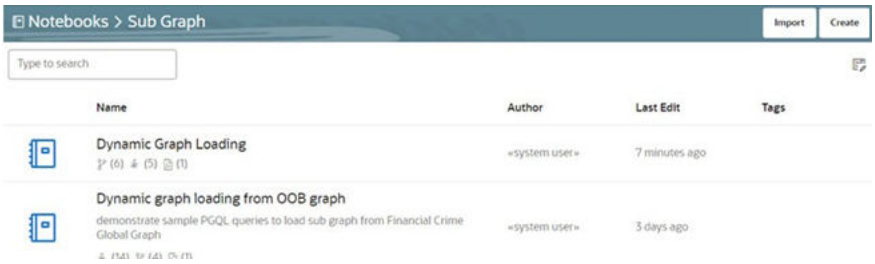


A.6.1.2 Dynamic Graph Loading from OOB Graph

This notebook provides the sample PGQL queries for dynamically loading graphs from property graphs (PG VIEWS) in the database.

> **Note:**
>
> **This notebook is created for sample purpose only**.

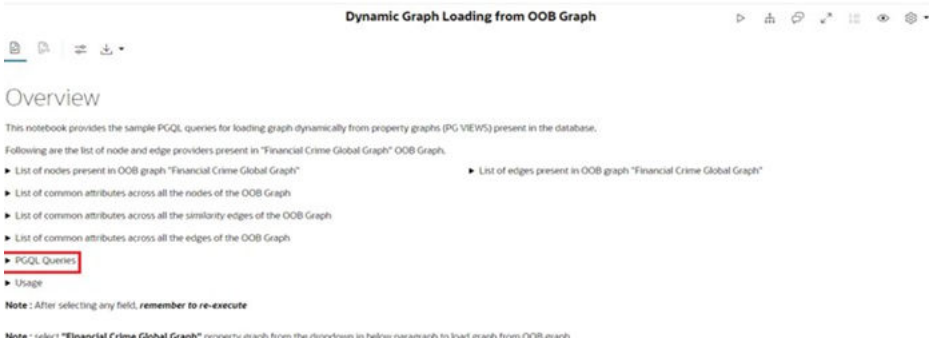
Figure A-19 Sub Graph



To access the sample PGQL queries for loading subgraph, follow these steps:

1. Click **Dynamic Graph Loading from OOB Graph**. The notebook page is displayed.

Figure A-20 Dynamic Graph Loading from OOB Graph



2. Click **PGQL Queries** group to view the list of sample queries to be executed for loading the subgraph as described in the following table.

Table A-2 Sample PGQL Query

PGQL Query	Description
MATCH (p)-[e]-(m)	This will load the complete property graph.
MATCH (p:"label(p)")- [e:"label(e)"]-(m:"label(m)")	This will load two node providers and the corresponding edge provider to connect two nodes. **Note**: Replace label (p) and label (m) with the node provider names and label (e) with the edge provider name.
MATCH (p:"label(p)")- [e:"label(e)"]-(m:"label(m)") WHERE id(p)="label(p)(Id)"	This will load the graph corresponding to the Id. Multiple conditions can be included similarly in the WHERE clause, separated by " AND "OR" condition. **Note**: Replace label (p) and label (m) with the node provider names and label (e) with the edge provider name. Replace Id with the Id property value
"MATCH (p:"label(p)")-()-()- (m:"label(m)") WHERE id(p) = 'label(p)(Id)"	This will only load p and m vertices, but not the edge or vertices linking them. **Note**: Replace label (p) and label (m) with the node provider names and replace Id with the Id property value.
"MATCH ()"	This is not allowed as the query must have at least one vertex. **Note**: This query is not allowed for execution.
"MATCH ()-[e]-()"	This is not allowed for loading an edge because it should load both the source and destination vertex. **Note**: This query is not allowed for execution.
"MATCH (v)-[e]-()"	This is not allowed for loading an edge because it should load both the source and destination vertex. **Note**: This query is not allowed for execution.

A.7 APIs for Model Pipelines

All the listed APIs/Functions can be utilized by users in Model Pipelines.

Table A-3 Python Paragraph Scripting

API Type	Functionality	Prerequisite	Notebook Execution Script	Notebook Script Input	Minimum Supported Version
Set User	Set the user for a session	-	%python from mmg.constants import * user = "SAUSER" set_user(user)	user - username to be set	8.1.1
Get User	Get the user that has been set	Set User	%python from mmg.constants import * get_user()	-	8.1.1
Attach Workspace	Attach the workspace for a session	-	%python from mmg.workspace import attach_workspace workspace = "CS" attach_workspace(workspace)	workspace- workspace to be attached	8.1.1
Get Workspace	Get the workspace that has been attached	Attach Workspace	%python from mmg.workspace import get_workspace get_workspace()	-	8.1.1
List Workspaces	Lists all workspaces available to a user	Set User	%python from mmg.workspace import list_workspaces list_workspaces()	-	8.1.1
Check AIF	Check if AIF is set or not	Set User	%python from mmg.workspace import check_aif check_aif()	-	8.1.1

Table A-3 (Cont.) Python Paragraph Scripting

API Type	Functionality	Prerequisite	Notebook Execution Script	Notebook Script Input	Minimum Supported Version
List All Data Sources	List all data sources available in a workspace for a set user. The default workspace chosen is the one that is attached.	Set User, Attach Workspace	%python from mmg.workspace import list_datasources list_datasources(order=None)	order - None ensures that all the data sources will be listed for the attached workspace	8.1.1
List All Data Sources of Workspace	List all Data Sources of a specific Workspace that is not the attached workspace for a set user.	Set User	%python from mmg.workspace import list_datasources workspace="HELLO NEW" list_datasources(workspace,order=None)	workspace - Name of workspace for which data sources should be listed order - None ensures that ALL data sources will be listed for that workspace	8.1.1
Fetch First Order Data Source	This method will fetch the data source related to attached workspace with default order 1.	Set User, Attach Workspace	%python from mmg.workspace import list_datasources list_datasources()	order - 1 by default, if not specified	8.1.1
Fetch First Order Data Source of Workspace	This method will fetch the data source with default order 1 related to specified workspace.	Set User	%python from mmg.workspace import list_datasources workspace="CS" list_datasources(workspace)	workspace - Name of workspace for which data source should be listed order- 1 by default if not specified	8.1.1
Fetch Nth Order Data Source	This method will fetch the data source related to attached workspace with given order	Set User, Attach Workspace	%python from mmg.workspace import list_datasources list_datasources(order=2)	order - Order of data source to be fetched	8.1.1

Table A-3 (Cont.) Python Paragraph Scripting

API Type	Functionality	Prerequisite	Notebook Execution Script	Notebook Script Input	Minimum Supported Version
Fetch Nth Order Data Source of Workspace	This method will fetch the data source with given order related to specified workspace.	Set User	%python from mmg.workspace import list_datasources workspace="CS" list_datasources(workspace,order=2)	workspace - Name of workspace for which data source should be listed order = Order of data source to be fetched	8.1.1
Get Connection Object	Returns Oracle/ SQL Alchemy Connection Object to Data Source of order 1 in attached workspace.	Set User, Attach Workspace	%python import cx_Oracle from mmg.workspace import get_connection conn=get_connection() if not isinstance(conn, cx_Oracle.connect): print("Not Connection Object") print(conn)	-	8.1.2.0
Get Connection Object To Specific Data Source	Returns Oracle/ SQL Alchemy Connection Object to specified Data Source	Set User	%python import cx_Oracle from mmg.workspace import get_connection datasource="DS001" conn=get_connection(datasource) if not isinstance(conn, cx_Oracle.connect): print("Not Conn Object") print(conn)	datasource - Name of datasource for which the connection object has to be returned	8.1.2.0

Table A-3 (Cont.) Python Paragraph Scripting

API Type	Functionality	Prerequisite	Notebook Execution Script	Notebook Script Input	Minimum Supported Version
Get Connection Object To Data Source Using Order	Returns Oracle/ SQL Alchemy Connection Object to Data Source of specified order in attached workspace.	Set User, Attach Workspace	%python import cx_Oracle from mmg.workspace import get_connection datasource=2 conn=get_connection(datasource) if not isinstance(conn, cx_Oracle.connect): print("Not Conn Object") print(conn)	datasource - Order number of datasource in attached workspace for which the connection object has to be returned	8.1.2.0
Test Connection Object	This script can be used to test the connection object to data source	Get Connection Object OR Get Connection Object To Specific Data Source OR Get Connection Object To Data Source Using Order	%python query="SELECT table_name FROM user_tables" cur = conn.cursor() cur.execute(query) print(cur.fetchall()) if cur: cur.close()	-	8.1.2.0

Table A-4 PYTHON LINEPARSER APIs (for MMG version 8.1.0)

API Type	Functionality	Notebook Execution Script	Note
Attach Workspace	Attach the workspace for a session	%python mmg.attachWorkspace(workspace) #eg-mmg.attachWorkspace("CS")	workspace - workspace to be attached (String) Sets 'mmg' python variable which contains the information about the attached workspace. Output- Print a boolean in one line on the status of workspace attachment and another line of boolean on status of AIF enabled and attached.

Table A-4 (Cont.) PYTHON LINEPARSER APIs (for MMG version 8.1.0)

API Type	Functionality	Notebook Execution Script	Note
List Workspaces	Lists all workspaces available to a user	%python mmg.listWorkspaces()	Sets 'ofs_wsList' python variable which is the list of all available workspaces. Output- Print the list of workspaces.
Check AIF	Check if AIF is set or not	%python print(aif__isAttached)	Output- Boolean on the status of AIF.
Get Connection Object	Returns cx_Oracle Connection Object	%python mmg.getConnection(schema_name) #eg- mmg.getConnection("OFSAA")	schema_name = name of the schema for which connection is required (String). Sets 'conn' python variable which is the cx_Oracle connection object to be used for firing queries.
Publish	Fetch and sets the Notebook JSON dump	%python mmg.publish()	Sets 'emf_para' python variable which is the JSON dump for the notebook component. Output- Print the set 'emf_para' value.
Register	Register the Notebook	%python mmg.register(register) #eg- mmg.register({"modelname":"model1",\ # "modeldescription":"Model description"}) #eg- mmg.register(emf_para) {After doing mmg.publish()}	register - JSON Stringify object for RegisterNotebookBean Object. Sets 'publishedNotebookId' python variable which is the studio notebook id for the published version. Output- Prints the published notebook id.
Load Flat File	-	%python mmg.loadFlatFile() #eg- mmg.loadFlatFile() Incomplete	-
Profile Data	-	%python mmg.profileData(code,title) #eg- mmg.profileData(ABC,XYZ)	-

Table A-5 MMG Library Python APIs

Index	API Type	Functionality	Prerequisite	Notebook Execution Script
1	Set User	Set the user for a session	-	`%python from mmg.constants import * user = "SAUSER" set_user(user)`
2	Get User	Get the user that has been set	Set User	`%python from mmg.constants import * get_user()`
3	Attach Workspace	Attach the workspace for a session	-	`%python from mmg.workspace import attach_workspace workspace = "CS" attach_workspace(workspace)`
4	Get Workspace	Get the workspace that has been attached	Attach Workspace	`%python from mmg.workspace import get_workspace get_workspace()`
5	Get Base URL	Get MMG Studio Service Base URL	-	`%python from mmg.constants import * get_mmg_studio_service_url()`
6	List Workspaces	Lists all workspaces available to a user	Set User	`%python from mmg.workspace import list_workspaces list_workspaces()`
7	Check AIF	Check if AIF is set or not	Set User	`%python from mmg.workspace import check_aif check_aif()`
8	List All Data Sources	List all data sources available in a workspace for a set user. The default workspace chosen is the one that is attached.	Set User, Attach Workspace	`%python from mmg.workspace import list_datasources list_datasources(order=None)`

Table A-5 (Cont.) MMG Library Python APIs

Index	API Type	Functionality	Prerequisite	Notebook Execution Script
9	List All Data Sources of Workspace	List All Data Sources of a specific Workspace that is not the attached workspace for a set user	Set User	`%python from mmg.workspace import list_datasources workspace="HELLO NEW" list_datasources(workspace,order=None)`
10	Fetch First Order Data Source	This method will fetch the data source related to attached workspace with default order 1	Set User, Attach Workspace	`%python from mmg.workspace import list_datasources list_datasources()`
11	Fetch First Order Data Source of Workspace	This method will fetch the data source with default order 1 related to specified workspace	Set User	`%python from mmg.workspace import list_datasources workspace="CS" list_datasources(workspace)`
12	Fetch Nth Order Data Source	This method will fetch the data source related to attached workspace with given order	Set User, Attach Workspace	`%python from mmg.workspace import list_datasources list_datasources(order=2)`
13	Fetch Nth Order Data Source of Workspace	This method will fetch the data source with given order related to specified workspace	Set User	`%python from mmg.workspace import list_datasources workspace="CS" list_datasources(workspace,order=2)`

Table A-5 (Cont.) MMG Library Python APIs

Index	API Type	Functionality	Prerequisite	Notebook Execution Script
14	Get Connection Object	Returns Oracle/ SQL Alchemy Connection Object to Data Source of order 1 in attached workspace	Set User, Attach Workspace	`%python import cx_Oracle from mmg.workspace import get_connection conn=get_connection() if not isinstance(conn,cx_Oracle.connect): print("Not Conn Object") print(conn)`
15	Get Connection Object To Specific Data Source	Returns Oracle/ SQL Alchemy Connection Object to specified Data Source	Set User	`%python import cx_Oracle from mmg.workspace import get_connection datasource="DS" conn=get_connection(datasource, conn_type=None) if not isinstance(conn,cx_Oracle.connect): print("Not Conn Object") print(conn)`
16	Get Connection Object To Data Source Using Order	Returns Oracle/ SQL Alchemy Connection Object to Data Source of specified order in attached workspace	Set User, Attach Workspace	`%python import cx_Oracle from mmg.workspace import get_connection datasource=2 conn=get_connection(datasource, conn_type=None) if not isinstance(conn,cx_Oracle.connect): print("Not Conn Object") print(conn)`

Table A-5 (Cont.) MMG Library Python APIs

Index	API Type	Functionality	Prerequisite	Notebook Execution Script
17	Get Connection Object	Returns Oracle/ SQL Alchemy Connection Object to Data Source of order 1 in attached workspace	Set User, Attach Workspace	`%python import oracledb from mmg.workspace import get_conn conn=get_conn() if not isinstance(conn ,oracledb.Conne ction): print("Not Conn Object") print(conn)`
18	Get Connection Object To Specific Data Source	Returns Oracle/ SQL Alchemy Connection Object to specified Data Source	Set User	`%python import oracledb from mmg.workspace import get_conn datasource="DS" conn=get_conn(d atasource, conn_type=None) if not isinstance(conn ,oracledb.Conne ction): print("Not Conn Object") print(conn)`
19	Get Connection Object To Data Source Using Order	Returns Oracle/ SQL Alchemy Connection Object to Data Source of specified order in attached workspace	Set User, Attach Workspace	`%python import oracledb from mmg.workspace import get_conn datasource=2 conn=get_conn(d atasource, conn_type=None) if not isinstance(conn ,oracledb.Conne ction): print("Not Conn Object") print(conn)`
20	Get Objective Hierarchy	Returns the complete objective hierarchy including pipeline name	-	`%python print(objective Id)`

Table A-5 (Cont.) MMG Library Python APIs

Index	API Type	Functionality	Prerequisite	Notebook Execution Script
21	Fetch FTPSHARE Path	Get complete ftpshare path on the server	-	`%python from mmg.datasets import fetch_ftpshare fetch_ftpshare()`
22	Get Pipeline Name	Get the Pipeline Name to which the current notebook is attached	Attach Workspace	`%python from mmg.constants import get_pipeline_name get_pipeline_name()`
23	Get Pipeline ID	Get ID of Pipeline to which the current notebook is attached	Attach Workspace	`%python from mmg.constants import get_pipeline_id get_pipeline_id()`
24	Get Pipeline Version	Get version of Pipeline to which the current notebook is attached	Attach Workspace	`%python from mmg.constants import get_pipeline_version get_pipeline_version()`
25	Get MISDATE	Return MISDATE from pipeline execution if set otherwise takes default value provided	Attach Workspace	`%python misdate=ds.date_picker(name='$FICMISDATE$',format='yyyy-MM-dd', default_value='2000-01-01', label='FICMISDATE') print(misdate)`
26	Get BATCHRUNID	Returns BATCHRUNID if set from pipeline, otherwise takes default value	-	`%python batchrunid=ds.textbox(name='$BATCHRUNID$',default_value='Batch_auto',label='BATCHRUNID') print(batchrunid)`

Table A-5 (Cont.) MMG Library Python APIs

Index	API Type	Functionality	Prerequisite	Notebook Execution Script
27	Get TASKID	Returns TASKID if set from pipeline, otherwise takes default value	-	`%python taskid=ds.textbox(name='$TASKID$',default_value='task1',label='TASKID') print(taskid)`

Miscellaneous Helper Python Scripts

Table A-6 Miscellaneous Helper Python Scripts

Index	Name	Description	Prerequisite	Script
1	Create Connection Object Using Wallet Alias	This script can be used to test the connection to any Oracle database using wallet alias. Make sure to close connection object after use.	-	`%python wallet_alias = "WALLET_ALIAS_NAME" import oracledb oracledb.init_oracle_client() conn = oracledb.connect(dsn=wallet_alias`
2	Test Connection Object	This script can be used to test the connection object to data source	Get Connection Object OR Get Connection Object To Specific Data Source OR Get Connection Object To Data Source Using Order	`%python query="SELECT table_name FROM user_tables" cur = conn.cursor() cur.execute(query) print(cur.fetchall()) if cur: cur.close()`
3	Close Connection Object	-	Connection Object should be created using any of the above methods	`%python query="SELECT table_name FROM user_tables" cur = conn.cursor() cur.execute(query) print(cur.fetchall()) if cur: cur.close()`
4	Close Cursor	-	Cursor Object should be created using any of the above methods	`%python if cur: cur.close()`

Table A-6 (Cont.) Miscellaneous Helper Python Scripts

Index	Name	Description	Prerequisite	Script
5	Check environment variables	For example: Check wallet alias value of MMG_LIB_WALLET_ALIAS	-	`%python import os print(os.environ['MMG_LIB_WALLET_ALIAS'])`

Table A-6 (Cont.) Miscellaneous Helper Python Scripts

Index	Name	Description	Prerequisite	Script
6	Save Dataframe to a File	Save dataframe to a file stored on server	-	`%python from mmg.datasets import fetch_ftpshare, save_dataframe from mmg.workspace import get_workspace from mmg.constants import get_pipeline_version misdate=ds.date_picker(name='$FICMISDATE$', format='yyyy-MM-dd', default_value='2000-01-01', label='FICMISDATE') batchrunid=ds.textbox(name='$BATCHRUNID$', default_value='Batch_auto', label='BATCHRUNID') taskid=ds.textbox(name='$TASKID$', default_value='task1', label='TASKID') filename="df.csv" filepath=fetch_ftpshare() + "/mmg/" + get_workspace() + "/pipelines/" + objectiveId + "/" + get_pipeline_version() + "/output/" + misdate + "/" + batchrunid + "/" + taskid + "/" + filename`

Table A-6 (Cont.) Miscellaneous Helper Python Scripts

Index	Name	Description	Prerequisite	Script
				`res=save_dataframe(df,filepath)` `message=res["messages"] if res["status"]=="ERROR": raise Exception(message) else: print(message)`

Python Linepraser APIs (for MMG version 8.1.0)

Table A-7 Python Linepraser APIs (for MMG version 8.1.0)

Index	API Type	Functionality	Notebook Execution Script	Note
1	Attach Workspace	Attach the workspace for a session	`%python mmg.attachWorkspace(workspace) #eg- mmg.attachWorkspace("CS")`	workspace = workspace to be attached (String) Sets 'mmg' python variable which contains the information about the attached workspace Output- Print a boolean in one line on the status of workspace attachment and another line of boolean on status of AIFenabled and attached
2	List Workspaces	Lists all workspaces available to a user	`%python mmg.listWorkspaces()`	Sets 'ofs_wsList' python variable which is the list of all available workspaces Output- Print the list of workspaces
3	Check AIF	Check if AIF is set or not	`%python print(aif__isAttached)`	Output- Boolean on the status of AIF

Table A-7 (Cont.) Python Linepraser APIs (for MMG version 8.1.0)

Index	API Type	Functionality	Notebook Execution Script	Note
4	Get Connection Object	Returns cx_Oracle Connection Object	`%python mmg.getConnection(schema_name) #eg- mmg.getConnection("OFSAA")`	schema_name = Name of the schema for which connection is required (String) Sets 'conn' python variable which is the cx_Oracle connection object to be used for firing queries
5	Publish	Fetch and sets the Notebook JSON dump	`%python mmg.publish()`	Sets 'emf_para' python variable which is the JSON dump for the notebook component Output- Print the set 'emf_para' value
6	Register	Register the Notebook	`%python mmg.register(register) #eg- mmg.register({"modelname":"model1",\ # "modeldescription":"Model description"}) #eg- mmg.register(emf_para) {After doing mmg.publish()}`	register = JSON Stringify object for RegisterNotebookBean Object Sets 'publishedNotebookId' python variable which is the studio notebook id for the published version. Output- Prints the published notebook ID
7	Load Flat File	`%python mmg.loadFlatFile() #eg- mmg.loadFlatFile() Incomplete`	-	-
8	Profile Data	-	`%python mmg.profileData(code,title) #eg- mmg.profileData(ABC,XYZ)`	code = title =

A.8 Public APIs for Scheduler Operations

Public APIs for Scheduler Operations

The following APIs are exposed for the Scheduler Operations.

Note:

1. All the below APIs are POST requests.
2. All the APIs accept below values in request header:
   - ofs_tenant_id - Optional
   - ofs_service_id - Optional
   - ofs_workspace_id - Respective workspace id where the batch needs to be created
   - ofs_remote_user – MMG login user The ofs_workspace_id and ofs_remote_user are mandatory fields.

Base URL: `http(s)://<MMG_Service_HostName>:<MMG-Service_Port>/<CONETXTNAME>/`

Table A-8 APIs for Scheduler Operations

Functionality	API	Sample Request JSON	Sample Response JSON	Comments

Execute Immediate	rest-api/v1/external/trigger	{
		"batchName": "batch1",
		"batchType": "rest",
		"includedTasks" :"task1,task2",
		"excludedTasks" :"task1,task2",
		"heldTasks":"task1,task2"
		}
			{
			"severity": "info",
			"summary": "Object triggered successfully with Run Id:
			batch1_2022-08-17_1660721567845_1",
			"batchRunId": "batch1_2022-08-17_1660721567845_1",
			"details": "Object triggered successfully.",
			"status": "success",
			"statusCode": "0"
			}
				• The includedTasks/ excludedTasks/ heldTasks are options keys.
				• The batchType will be “group” for batch group execution.

Table A-8 (Cont.) APIs for Scheduler Operations

Functionality	API	Sample Request JSON	Sample Response JSON	Comments
Execution Status	rest-api/v1/external/status	{ "batchRunId": "batch1_2022-08-17_1660721567845_1", "tasks": ["task1","task2"] }	{ "severity": "info", "batchRunId": "batch1_2022-08-17_1660721567845_1", "taskStatusList": [{ "taskCode": "task1", "taskStatus": "SUCCESSFUL", "statusCode": "0" }, { "taskCode": "task2", "taskStatus": "EXCLUDED", "statusCode": "4" }], "batchStatusCode": "0", "batchList": [], "batchStatus": "SUCCESSFUL", "status": "success", "statusCode": "0" }	• "tasks" is options key. If not passed, then response will contain status all the tasks inside a batch.
Rerun	rest-api/v1/external/rerun	{ "batchName": "batchgroup1", "batchRunId": "batchgroup1_2022-08-17_1660720814942_1" }	{ "severity": "info", "summary": "Object triggered successfully for rerun with Run Id: batchgroup1_2022-08-17_1660730049819_1", "batchRunId": "batchgroup1_2022-08-17_1660730049819_1", "details": "Object triggered successfully.", "status": "success", "statusCode": "0" }	

Table A-8 (Cont.) APIs for Scheduler Operations

Functionality	API	Sample Request JSON	Sample Response JSON	Comments

Restart	rest-api/v1/external/restart	{
		"batchName": "MMG_R1",
		"batchRunId": "MMG_R1_2022-07-15_1657867378160_1"
		}
			{
			"severity": "info",
			"summary": "Object triggered successfully for restart with Run Id: MMG_R1_2022-07-15_1657867378160_1",
			"batchRunId": "MMG_R1_2022-07-15_1657867378160_1",
			"details": "Object triggered successfully.",
			"status": "success",
			"statusCode": "0"
			}	

Interrupt	rest-api/v1/external/interrupt	{
		"batchName": "B2001",
		"batchRunId": "B2001_2022-05-22_1653222717896_1"
		}
			{
			"summary": "Execution interrupted successfully for Run Id: B2001_2022-05-30_1653233511394_1",
			"severity": "info",
			"batchRunId": "B2001_2022-05-30_1653233511394_1",
			"details": "Execution interrupted successfully.",
			"statusCode": "0",
			"status": "success"
			}	

Table A-9 Failed Response Body

Failed Sample Response JSON	Description

Batch Status Failed	{
	"severity": "info",
	"batchRunId": "BT-BI-EXCHG_RATES_EOD_2022-05-27_1653662703599_1",
	"batchStatusCode": "-1",
	"batchList": [],
	"batchStatus": "FAILED",
	"statusCode": "0",
	"status": "success"
	}

Table A-9 (Cont.) Failed Response Body

Failed Sample Response JSON	Description

Object Not found	{
	"severity": "error",
	"summary": "Object does not exist.",
	"details": "Object does not exist.",
	"error": {
	"errorCode": "OBJECT_NOT_EXIST",
	"errorMsg": "Object does not exist."
	},
	"statusCode": "-3",
	"status": "failed"
	}

Table A-10 Status Codes in the Response Body

Status Code	Description
0	Success
1	Not Started
2	Ongoing
4	Excluded
5	Held
-1	Failure
-2	Interrupted
-3	Object does not exist
-4	Invalid arguments passed in request/not enough parameters in the Request body
-5	Invalid request headers/request headers missing
-6	No executable job is present
-7	Job is already interrupted