Oracle Financial Services Compliance Studio

User Guide

Release 8.1.2.3.0

January 2023

F48799-01

ORACLE®
Financial Services

ORACLE®

OFS Compliance Studio User Guide

This software or hardware and documentation may provide access to or information about content, products, and services from third parties. Oracle Corporation and its affiliates are not responsible for and expressly disclaim all warranties of any kind with respect to third-party content, products, and services unless otherwise set forth in an applicable agreement between you and Oracle. Oracle Corporation and its affiliates will not be responsible for any loss, costs, or damages incurred due to your access to or use of third-party content, products, or services, except as set forth in an applicable agreement between you and Oracle.

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.

Oracle customers that have purchased support have access to electronic support through My Oracle Support. For information, visit http://www.oracle.com/pls/topic/lookup?ctx=acc&id=info or visit http://www.oracle.com/pls/topic/lookup?ctx=acc&id=trs if you are hearing impaired.

Document Control

Table 1 lists the document control of this guide:

Table 1: Document Control

Version Number	Revision Date	Change Log
8.1.2.3.0	January 2023	Updated User-group field in the Configure Basic Details. Added a new Using Scenario Conversion Utility section.
8.1.2.1.0	December 2022	Updated a note (4.e) in the Using Graph Definition section.
8.1.2.1.0	September 2022	OJET Upgrade (all UI elements are updated according to UI in the entire document). Added the following sections: • Managing Graph Pipelines • Managing Datasets • Using Scheduler Service • Compliance Studio Global Graph • Mapping Graph Datasource • Using Graph Definition • Managing Data Pipeline • Accessing the Model Pipelines • Import a Workspace Model Data into a New Model • Cloning Ruleset (Match) Updated the following sections: • Using the OFS Compliance Studio Application • Using Audit Trail • ML4AML Model Templates Added new interpreter-based paragraphs in the About Paragraphs section. Added new list of nodes, edges, and pre-seeded indexes in the Pre-configured Graph FCGM Model section. Removed scenarios such as HRG - AC Focus, RMF - AC Focus, RMF - CUS Focus in the Using Example Scenario Notebooks section. Added new method **Generate Sequence** in the Creating Data Survival Rules section. Removed a note in the Creating Data Survival Rules section. Updated steps in the Settings Window section. Added **Re-initialize session** icon in the Common Screen Elements in a Notebook section.

Table 1: Document Control

Version Number	Revision Date	Change Log
8.1.2.0.0	March 2022	Updated the following sections: • About Paragraphs • ML4AML Model Templates • Entity Resolution • Using Ruleset (Match/Merge) • Data Survival Rules • Using Manual Decisioning Add the Using Merge and Split Global Entities section.
8.1.1.1.0	November 2021	Updated for the Data Survival Rules section.
8.1.1.0.0	October 2021	Created the user guide for the release v8.1.1.0.0 version.

Table of Contents

1   Preface

This preface provides information for the Oracle Financial Services Compliance Studio application (OFS Compliance Studio) User Guide.

Topics:

- About this Guide
- Audience
- Additional Resources
- Conventions Used
- Acronyms Used

1.1   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.

1.2   Audience

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

1.3   Additional Resources

This section identifies additional resources for the OFS Compliance Studio Application Pack.

You can access the following OFS Compliance Studio documents from the Oracle Help Center (OHC).

- OFS Compliance Studio Matching Guide
- OFS Compliance Studio Administration Guide
- OFS Compliance Studio Release Notes
- OFS Compliance Studio Installation Guide
- OFS Compliance Studio Architecture Guide

For more information about how Oracle Financial Services solves real business problems, see our website at www.oracle.com/financialservices.

1.4   Conventions Used

Table 2 lists the conventions used in this guide.

Table 2: Conventions Used in this Guide

Convention	Meaning
Italics	- Names of books, chapters, and sections as references - Emphasis
Bold	- The object of an action (menu names, field names, options, button names) in a step-by-step procedure - Commands typed at a prompt - User input
`Monospace`	- Directories and subdirectories - File names and extensions - Process names - Code sample, including keywords and variables within the text and as separate paragraphs, and user-defined program elements within the text
Hyperlink	Hyperlink type indicates the links to external websites and internal document links to sections.

1.5 Acronyms Used

Table 3 lists the acronyms used in this guide.

Table 3: Acronyms Used in this Guide

Conventions	Description
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
PGX	Parallel Graph Analytics / Graph Analytics Engine

Table 3: Acronyms Used in this Guide

Conventions	Description
FCGM	Financial Crime Graph Model

2   About OFS Compliance Studio

Topics:

- Introduction
- Capabilities offered by Compliance Studio
- Architecture Overview
- Process Flow

2.1   Introduction

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.

2.2   Capabilities offered by Compliance Studio

- Purpose-Built for Fighting Crime
  - Fully defined and sourced Financial Crime Graph Model supporting detection and investigation
  - Provided Accelerators for finding the needles in the haystack.
  - What if Analysis for existing Scenarios
  - Integration with ECM and Investigation Hub 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.x releases of Oracle Financial Compliance Management Anti Money Laundering (AML), Enterprise Case Management, and Fraud applications.
- Entity Resolution for AML
  - 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
  - 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 etc.
- Analytics of Choice
  - Choose from our proprietary models or bring your own
  - Fully embedded Graph Analytics Engine and Financial Crime Model
  - Supports non-ML
  - Embedded with a highly scalable in-memory Graph Analytics Engine (PGX)
  - Industry's most intuitive Graph Query Language to gain rapid insights

- 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
- 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 build typology detection models, detect anomalies, and risk score customers or events
  - Event Scoring for false positive prediction and disposition
  - Dashboards and KPIs to measure the impact of ML and generate insights for business users
  - Ongoing Monitoring of Model Performance and Concept Drift
  - Automated Scenario Calibration and Scenario Conversion Utility for Oracle AML Scenarios
- 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.

2.3 Architecture Overview

See the OFS Compliance Studio Architecture Guide for more details on Compliance Studio architecture.

2.4 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: OFS Compliance Studio Process Flow
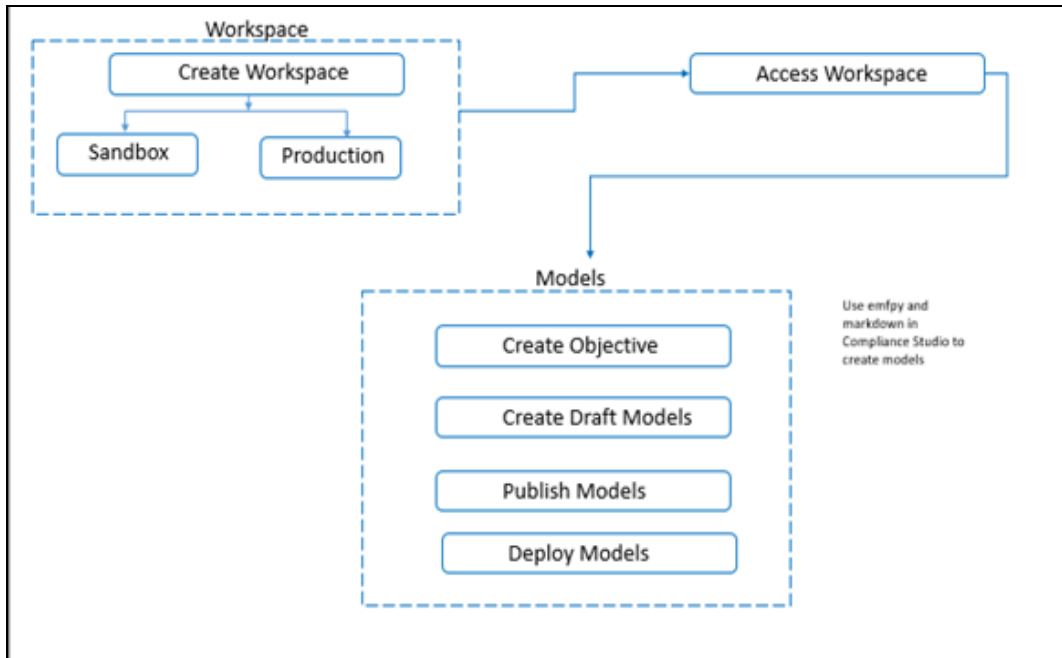
3   Getting Started

This section provides 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.

Topics:

- Accessing the OFS Compliance Studio Application
- Using the Application UI
- Mapping User Groups
- Using the OFS Compliance Studio Application

3.1   Accessing the OFS Compliance Studio Application

After the application is installed and configured, you can access the OFS Compliance Studio application.

To access OFS Compliance Studio, follow these steps:

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

Figure 2: OFS Compliance Studio Application Login Page
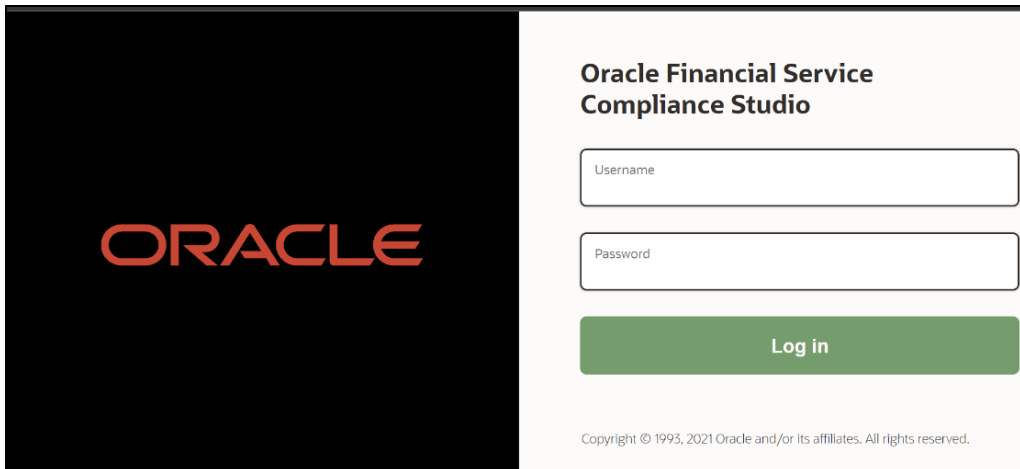
NOTE	- The login page only appears in the case of FCCMRealm.
	- In the case of SAMLRealm, the user will see the login page of their Identity Provider.

2. Enter your **User ID** and **Password**.

3. Click **Log in**. The Workspace Summary page is displayed.

Figure 3: Compliance Studio Application Selection
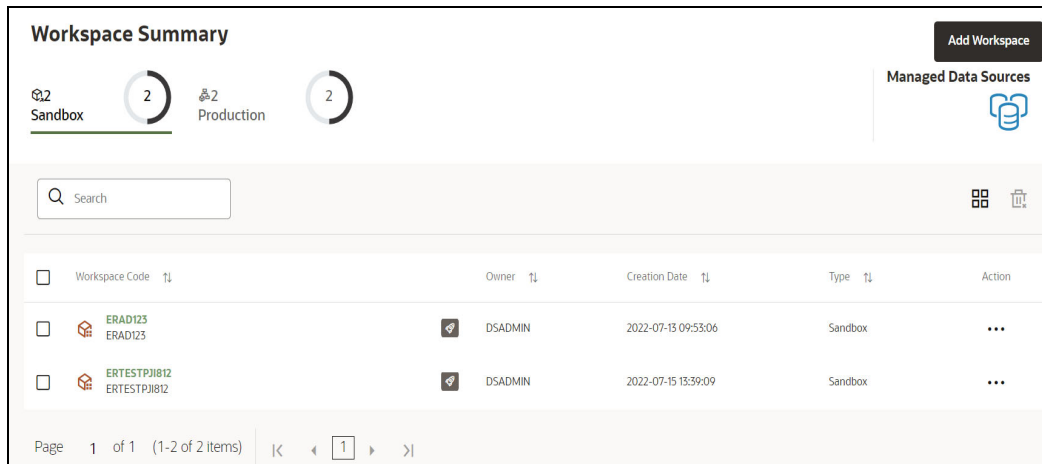
3.2 Using the Application UI

This section describes the user interface of the OFS Compliance Studio application. These are the common features that are found across the modules of OFS Compliance Studio. It describes the organization of the user interface and provides step-by-step instructions for navigating through the application.

Topics:

- Home Page Components

3.2.1 Home Page Components

The Home Page contains the following sections:

- Header

- Home icon

- Application Navigation
- Workspace Summary

Figure 4: Home Page
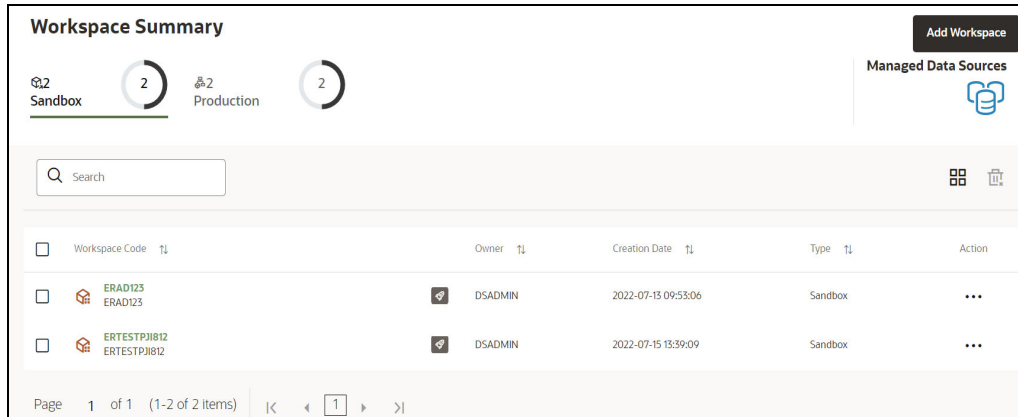
3.2.1.1   Home Page icon

- Click this icon to refresh. This icon is also used to navigate the Workspace Summary page from Dashboard.

3.2.1.2   Application Navigation

- Click **Application Navigation** icon to navigate to Dashboard, Dataset, Advance Options, and other sections of OFS Compliance Studio application.

3.2.1.3   Header

The following user-interface elements are displayed for OFS Compliance Studio:

Figure 5: Header
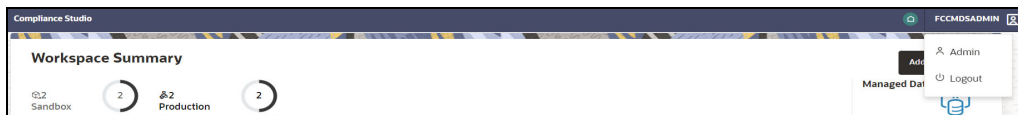
- **User Name:** Displays the logged-in user name. Click this to select from the following options in the drop-down list:

- **Admin:** Navigate to the Identity Management page. For more information, see OFS Admin Console User Guide.

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

3.2.1.4   Workspace Summary

Workspace Summary shows the list of Workspaces. Select the checkbox next to the Workspace that you want to View, Edit, Copy, or Delete. Click **Add** to create a new Workspace. 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 Type.

3.3 Mapping User Groups

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 the **Admin**. The Identity Management window is displayed.
	▪ For more information on permissions, see OFS Admin Console User Guide.

3.4 Using the OFS Compliance Studio Application

The Compliance Studio Application has the following key capabilities and concepts described in greater detail in this document.

3.4.1 Workspace Management

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 Using Workspaces section for information on how to use Workspace Management in the application.

3.4.2 Model Management

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

3.4.3 Model Actions

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

For more information, see the Using Model Actions section.

3.4.4 Graphs (Graph Pipeline)

Compliance Studio allows users to harness the power of Graph Analytics using our in-built in-memory Oracle Graph Analytics Engine (PGX). A pre-configured financial crime graph is provided which loads data from our Financial Crime Data Model and ICIJ through configurable pipelines.

The Graph Pipeline functionality within Compliance Studio allows users to easily define graphs, attach underlying data, and match pipelines to populate data in the graph.

For more information, see the Managing Graph Pipelines section.

3.4.5 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 Managing Datasets section.

3.4.6 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 Using Scheduler Service section.

3.4.7 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 Using Audit Trail section.

3.4.8 Data Studio Options

The Data Studio Options window allows you to view the list of options that are associated in the OFS CS. For more information, see the Accessing Notebook Server Options section.

3.4.9 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.

3.4.10 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.

3.4.11 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.

3.4.12 ML4AML Model Templates

The prepackaged Model templates allow 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)
- Typology Scenarios
    - Shell Account detection

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

4   Using Workspaces

Workspace management allows you to create and manage workspaces in the OFS Compliance Studio application.

Workspace 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
- Managed Data Sources

Topics:

- Accessing the Workspace Summary Page
- Accessing the Compliance Studio Workspace
- Creating new Workspaces
- Launching a Workspace
- Managing Workspaces
- Managing Data Sources

4.1   Accessing the Workspace Summary Page

The Workspace Summary page gives access to the various workspace functions such as create and delete.

NOTE	Users Groups must be mapped to the Workspace Admin role to access the Workspace Summary page.

To access the Workspace Summary page, follow these steps:

After login to **OFS Compliance Studio** application, the **Workspace Summary** page is displayed. The page displays workspace records in a table or card view. You can launch a workspace using the launch option or double clicking on it.

Figure 6: Workspace Summary Page
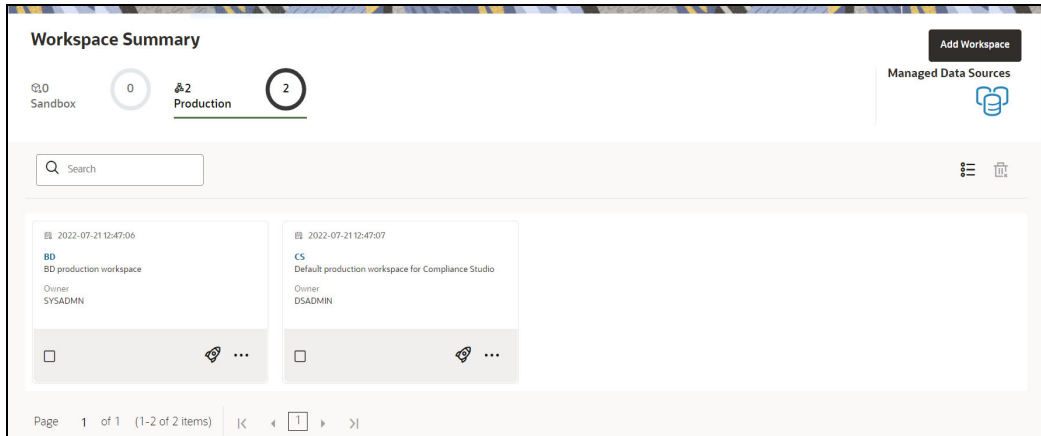
The Table 4 provides descriptions for the fields and icons on the **Workspace Summary** page.

Table 4: Workspace Summary
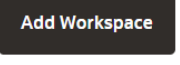
Field or Icon	Description
Search	The field to search for Workspace. Enter specific terms in the field for which you want to search, and press Enter on the keyboard to display the results. The search is available for all the fields.
Card and Grid View	Click the Card View [icon] or Grid View [icon] on the top-right corner of this pane to view the workspace as a block or a table, respectively.
Delete All	Select the Workspaces and click **Delete** [icon] to delete all the Workspaces.
Workspace Code	The code of the Workspace.
Description	The description for the Workspace
Creation Date	The date on which the Workspace was created.
Data Source Type	Shows the type of the Data Source.
Type	Shows the type of Workspace as Sandbox or Production.
Attachment	The Production Workspace shows the attached Sandbox in Card view.
Add Workspace	Click **Add Workspace** and select Workspace to create a new Workspace. See the Creating new Workspaces section for more information.
Launch Workspace	Click **Launch Workspace** [icon] next to corresponding workspace to display the **Workspace Dashboard** with application configuration and model creation menu. See the Launching a Workspace section for more information.

Table 4: Workspace Summary

Field or Icon	Description
Populate Workspace	Click Action ⋯ next to corresponding Workspace and select **Populate Workspace** 🔷 to populate the workspace with dataset data. See the Populating the Workspace section for more information.
Edit Workspace	Click Action ⋯ next to corresponding Workspace and select **Edit Workspace** ✏️ to edit the Workspace. See the Editing the Workspace section for more information.
Delete Workspace	Click Action ⋯ next to corresponding Workspace and select **Delete Workspace** 🗑️ to delete the Workspace. See the Populate Workspace section for more information.

4.2 Accessing the Compliance Studio Workspace

The out-of-the-box Compliance Studio Workspace allows you to perform Matching, Entity Resolution, ETL, etc. For more information, see the Using the Compliance Studio Workspace section.

4.3 Creating new Workspaces

The Workspace creation requires entry of the source of dataset, validation, and deployment. Besides, the OFS Compliance Studio 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. For information on User Groups, Mappings, see the **User Access and Permission Management** section in the OFS Compliance Studio Administration and Configuration Guide.

To create a Workspace, follow these steps:

1. Navigate to **Workspace Management Summary** page to display the **Workspace Summary** page. The page displays workspace records in a table.

2. Select ➕ to create a Workspace in the Workspace **Summary** window.

Figure 7: Workspace Creation Window
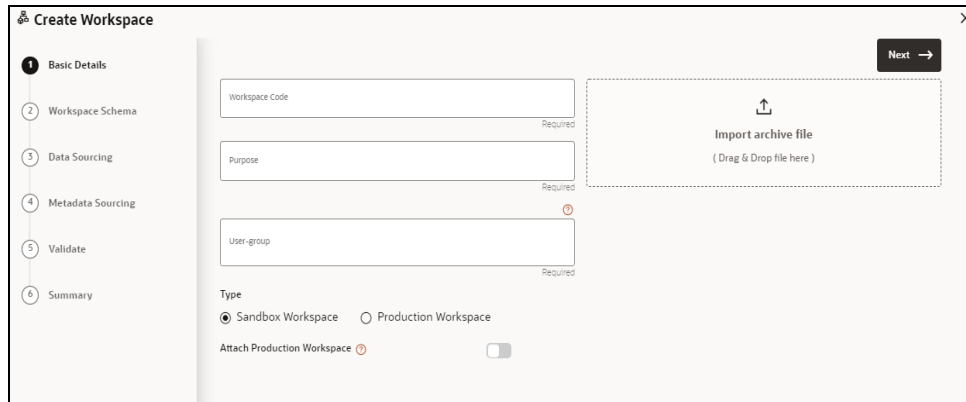
The window displays a progress indicator at the left that indicates the active window where you are entering details. Click **Previous** to go back a step and click **Next** to go to the next step.

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

- Configuring Basic Details
- Configuring Workspace Schema
- Configuring Data Sourcing
- Configuring Metadata Sourcing
- Validating Workspace
- Displaying Summary

4.3.1 Configuring Basic Details

Enter basic configuration details in this window.

Figure 8: Configure Basic Details
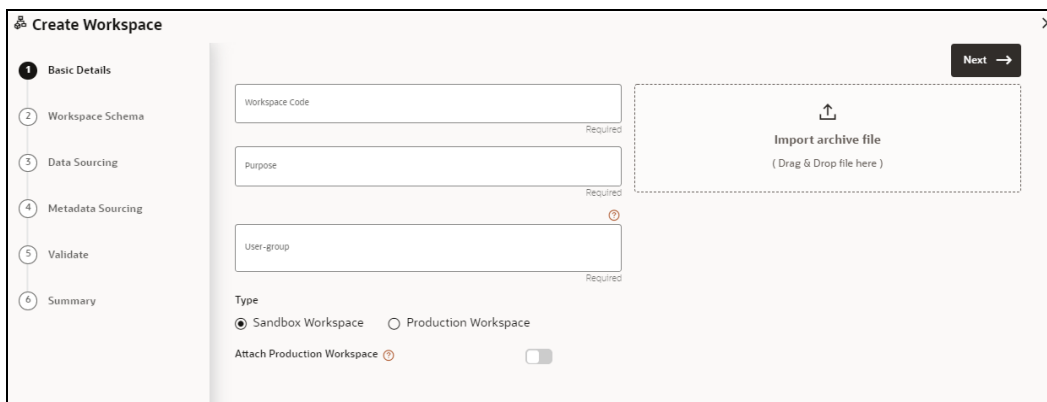
To configure the basic details for the workspace, follow these steps:

1. Enter the required details in the **Basic Details** pane as shown in the Table 5.

Table 5: Configure Basic Details

Field	Description
Workspace Code	Enter the code of the workspace. This field is limited to 20 characters.
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. ● Modeling Approver ● Modeling Reviewer ● Modeling User (Mandatory)
Type	Select the type of Workspace as Sandbox or Production. UGDOMMAP function should be mapped to the user performing sandbox creation, else create operation will fail.
Attached Production Schema	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 OFSAA Schema, 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.
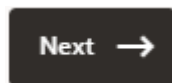
2. Click [Next →] to go to the next step.

4.3.2   Configuring Workspace Schema

Select the schema operation and enter connection details.

Figure 9: Workspace Schema

To configure the Workspace schema, follow these steps:

1. Select the required option from the **Meta Schema** and **Data Schema** drop-down list.

2. You can upload a new data model with the details you provide in the Data Sourcing drop-down list. You can select an existing Schema or add a new Schema.

    ▪ **Meta Schema**: Every workspace has a schema to store the meta information such as batch ID and so on. This is used for saving all the definitions and metadata related to models.

    ▪ **Data Schema**: This is the actual data used for model building. You can use multiple data schemas for one Meta Source.

    To add a new Schema, follow these steps:

1. Click Add   .

    The **Create Data Source** window is displayed.

Figure 10: Create Data Source with Oracle Database

Figure 11: Create Data Source with Hive Database

2. Enter the required details as shown in the Table 6.

Table 6: Create Data Source

Field	Description
Database Source Name	Enter the connection URL to the database for the data schema.
Description	Enter the description of database connection.
Type	Enter the type of the database connection.
Database Type	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 Create Data Source with Oracle Database), then following additional fields are displayed to enter details: ● Table Owner: Enter the Oracle Database name. If you select Database Type as Hive (see Create Data Source with Hive Database), then following additional fields are displayed to enter details: ● User Name: User Name / Principal is used for Kerberos authentication. ● Table Owner: Enter the Hive schema. ● JDBC Connection String: Enter the JDBC Connection String. ● 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. ● Realm File Name: Enter the Name of the configuration file present in conf directory. **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 OFS Compliance Studio Installation Guide.
Wallet Alias	Enter the Wallet Alias. This value should be same as configured using **Oracle Wallet**. For more information, OFS CS Administration and Configuration Guide.

3. Click **Create.**

   OR

   Click **Cancel** to skip the changes.

> **NOTE**   Click **Test Connection** to check the connection.
>
> A success message is displayed.
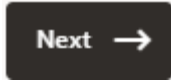
4. Click [Next →] to go to the next step.

4.3.3 Configuring 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 Hive 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.

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 and Hive data sources.

Enter the details in this window.

NOTE	This step does not apply to the prepopulated schema and OFSAA Infodom options in Configuring Workspace Schema.

Figure 12: Data Sourcing - External Data Source
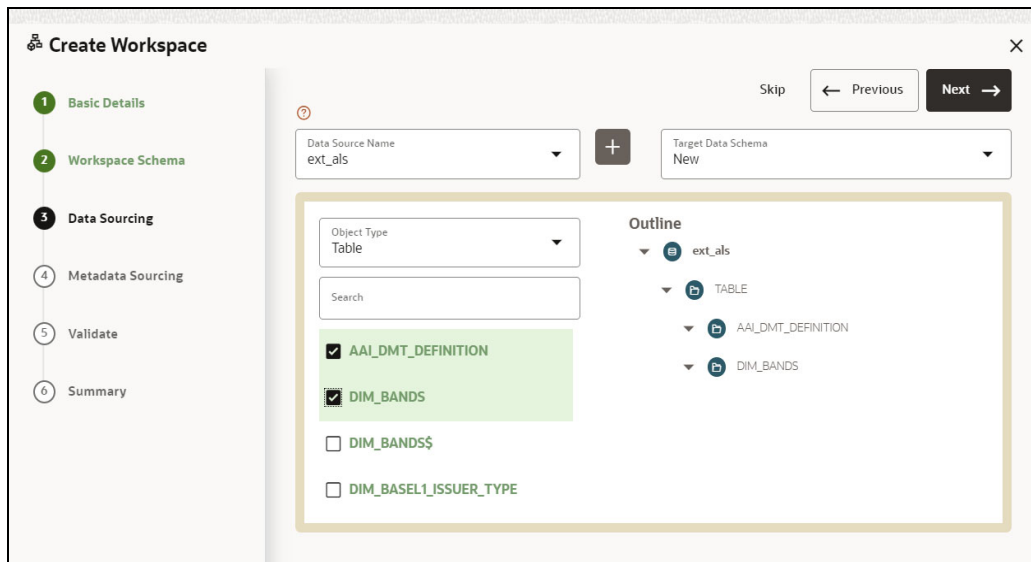
To configure Data Sourcing, follow these steps:

1. You can select a Data Source from **Data Source Name** drop-down list or create a new Data Source.

   To create a new Data Source, see the Configuring Data Sourcing section.

2. Select the **Target Data Schema**.

   You can select multiple Data Sources for a Target Data Schema.

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.

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:

- If Table A has columns C1, C2 and C3

- If Table B has columns C1, C2 and C4

    Then the data from the first table will be used.

    During the data population, only columns C1 and C2 will be used and those will be marked in Green color.

3. 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

4. Click   Next →   to go to the next step.

    OR

    Click **Skip** to skip the step.

4.3.4   Configuring 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.

> **NOTE**   This step is optional.

Figure 13: Configure Metadata Sourcing
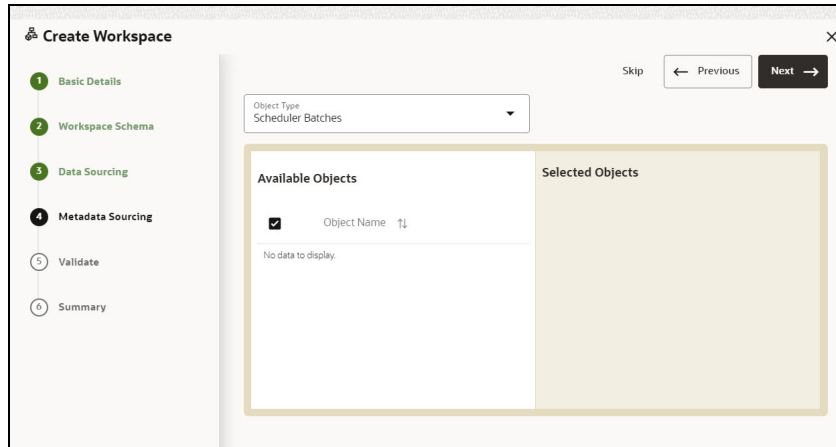
To configure Metadata Sourcing, follow these steps:

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

    The **Available Objects** are displayed.

2. Select the objects that will be displayed under **Selected Objects** group.

3. Click Next to go to the next step.

    OR

    Click **Skip** to skip the step.

4.3.5   Validating Workspace

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

Figure 14: Validate Workspace
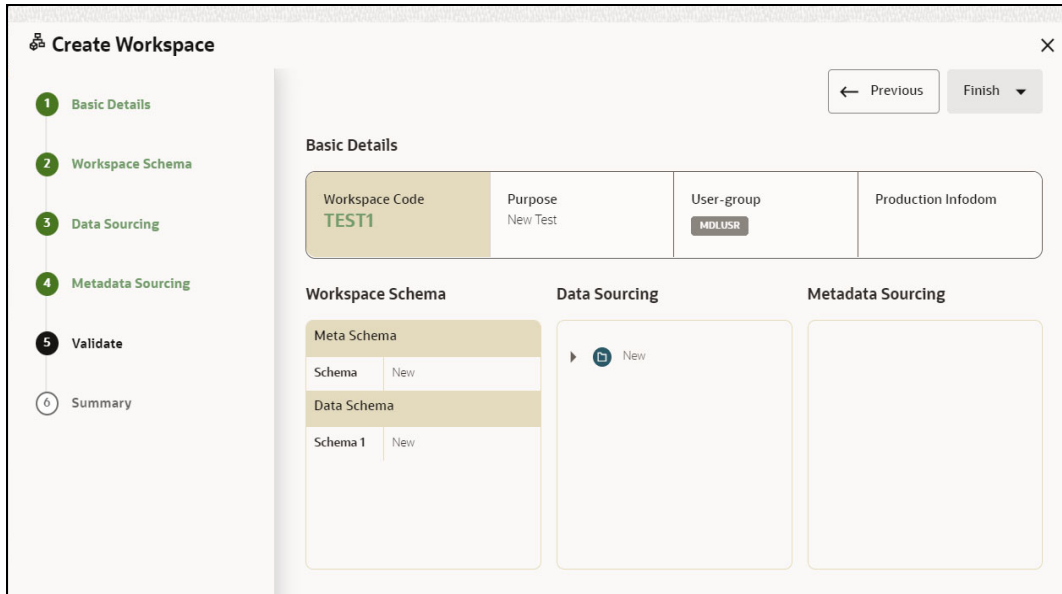
To validate the Workspace and deploy, follow these steps:

1. Review the details in the **Validate** pane.

2. Click **Finish** to creation of the Workspace using **Physicalize Sandbox** option or **Download Configurable Archive**.

Figure 15: Finish Menu
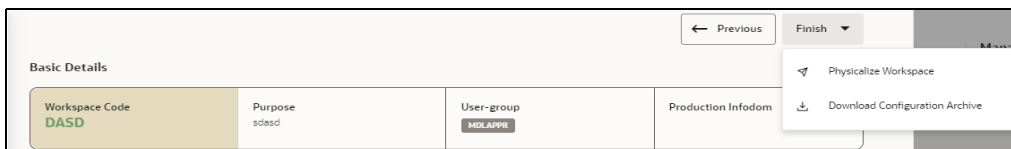
- When you click **Download Configurable Archive**, it exports the metadata information of the workspace in .zip format which can be used later using the Import option.

- When you click **Physicalize Sandbox**, it creates actual workspace.

4.3.6   Displaying Summary

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

Figure 16: Workspace Creation Summary
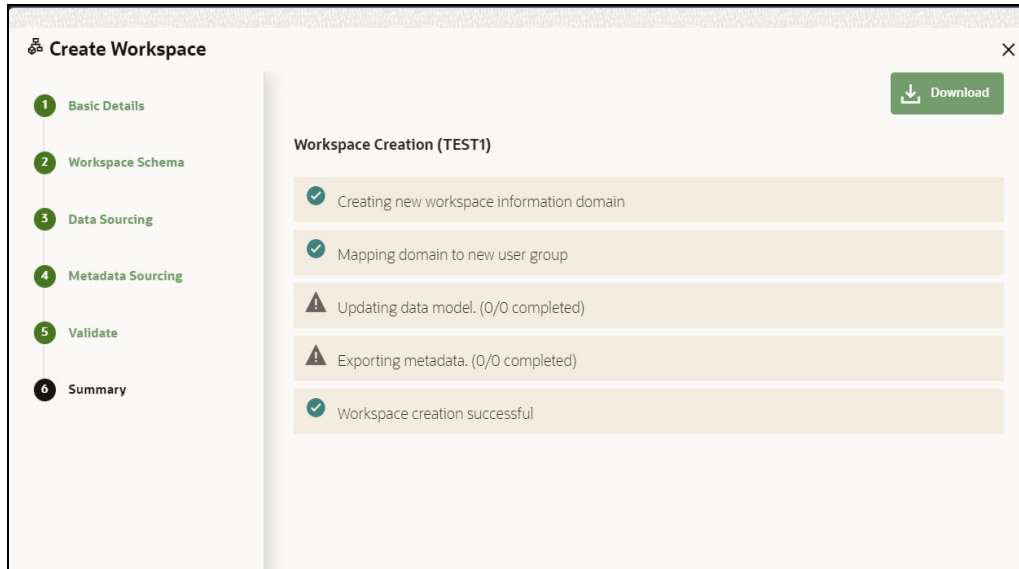
- Click **⬇ Download** to download the deployment report.

4.4 Managing Workspaces

4.4.1 Launching a Workspace

The launching of the Workspace displays the **Workspace** window. The workspace displays a menu for Models and an application configuration and model creation submenu as shown in the following list:

- Dashboard
- Dataset
- Model Pipelines
- Model Actions
- Graphs
- Scheduler Service
- Audit Trail
- Data Studio Options
- Ruleset Details
- Manual Decisioning
- Merge and Split Global Entities

To launch a workspace, perform the following steps:

1. Click **Launch** icon 🚀 next to corresponding Workspace.

   The Workspace Dashboard is displayed.

Figure 17: Workspace Dashboard
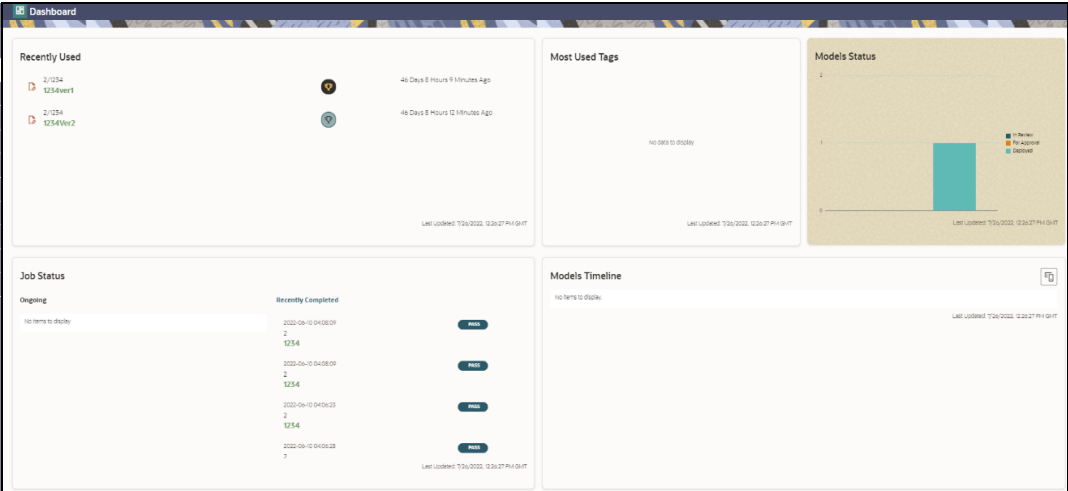
4.4.2   Viewing the Workspace

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 Workspace** ✎ to view the workspace.

    The Basic Details of the Workspace is displayed.

Figure 18: View Workspace Window

3. Navigate to **Workspace Schema**, **Data Sourcing**, and **Metadata Sourcing** tabs to view the respective details.

4.4.3 Populating the 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 Workspace** to populate the Workspace with data from a dataset data in the **Populate Workspace** window.

Figure 19: Populate Workspace Window

The Table 7 provides descriptions for the fields in the **Populate Workspace** window.

Table 7: Populate Workspace

Field	Description
Workspace Code	The code of the Workspace.
Purpose	The description for the Workspace.
Creation Date	The date 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.
Data Filter - Global	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.
Data Filter - Table level	Provide the data filters individually on the tables here. **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. Global Filters will not be 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.
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:
	● 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.
	● Unlimited
	  Here, all the errors will be ignored during the data population.

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.

> **NOTE**   You may require approval from an approver to populate the
> workspace.

- 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.

4. Navigate to Scheduler Service and select Define Task.

Figure 20: Task Parameters

5. Enter the following parameters for workspace population.

- **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. |

4.4.4 Editing the Workspace

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 Workspace** 🖉 to edit the workspace.

Figure 21: Edit Workspace Window
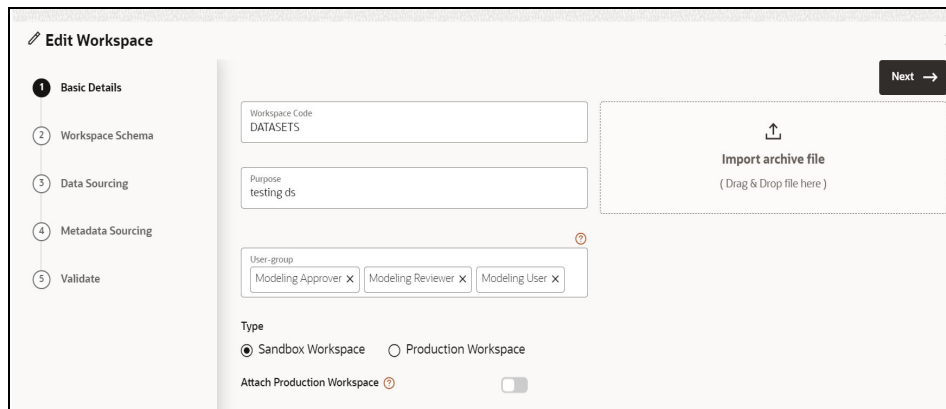
| NOTE | You can modify the Workspace Type from Sandbox to Production or vice versa. |

The Table 8 provides descriptions for the fields in the **Edit Workspace** window.

Table 8: Edit Workspace

Field	Description
Workspace Code	Enter the code of the workspace. **NOTE:** You cannot edit the Workspace Code.
Purpose	The description for the workspace.

Table 8: Edit Workspace

Field	Description
User-group	Click on this field to display a list of User-group values. Select the required value. • Modeling Approver • Modeling Reviewer • Modeling User
Type	Select the type of Workspace as Sandbox or Production. UGDOMMAP function should be mapped to the user performing sandbox creation, else create operation will fail. **NOTE**: You can modify the Workspace Type from Sandbox to Production or vice versa.
Attached Production Workspace	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 OFSAA Schema, the model is promoted to the environment.
Import archive file	You cannot import the archive file while editing the Workspace.

3. Click **Save**.

4.4.5   Deleting the Workspace

To delete a Workspace, follow these steps:
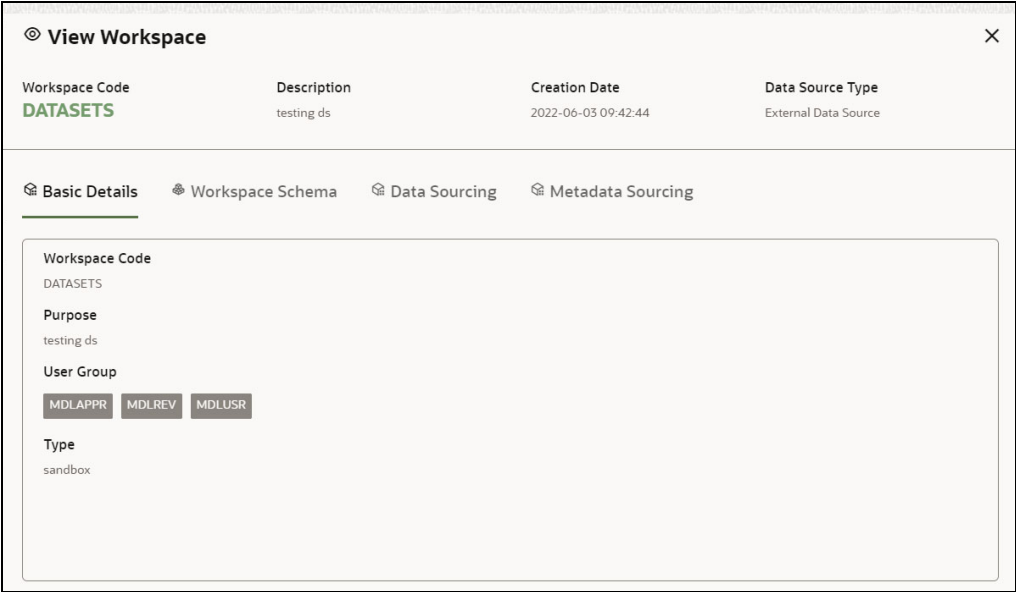
1. Navigate to **Workspace Management** to display the **Workspace Summary** page. The page displays workspace records in a table.

2. Click Action ⋯ next to corresponding Workspace and select **Delete Workspace** 🗑 to delete the workspace.

Figure 22: Delete Workspace Window
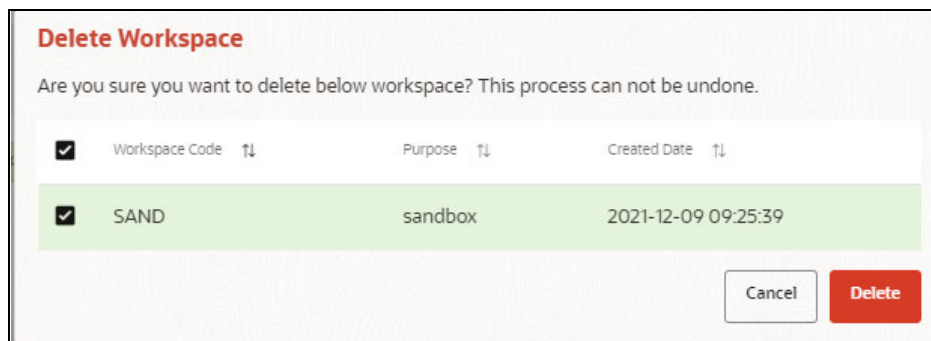
The Table 9 provides descriptions for the fields in the **Delete Workspace** window.

Table 9: Delete Workspace

Field	Description
Workspace Name	The name of the workspace.
Purpose	The description for the workspace.
Creation 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	You may require approval from an approver to delete a workspace. You must de-link the Sandbox workspaces from Production Workspace before deleting the Production Workspace.

4.5   Managing Data Sources

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

to navigate to **Data Source Summary** window.

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

Figure 23: Data Source Summary
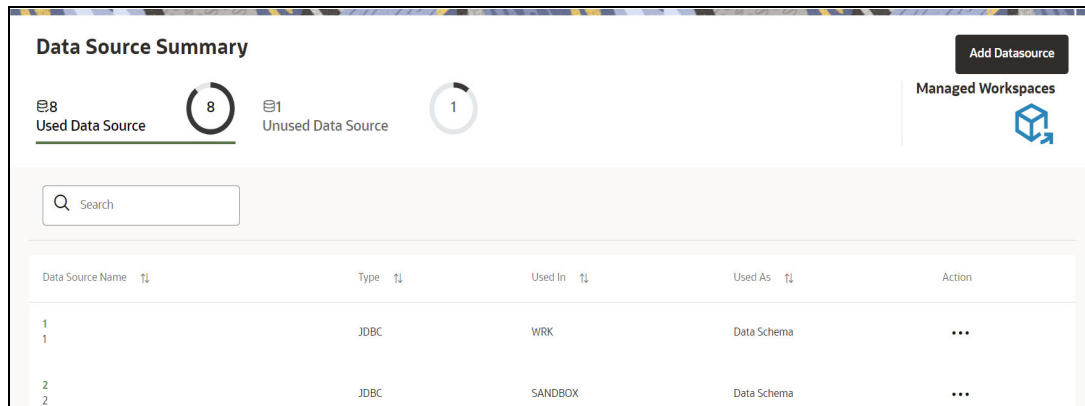
The Data Source Summary is divided into two sections:

- **Used Data Sources:**

  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 Sources:**

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

Figure 24: View Data Source

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

1. Navigate to **Workspace Summary** window.

2. Click Managed Data Sources to navigate to **Data Source Summary** window and click

    Add Datasource icon .

   The Create Data Source window is displayed.

   For more information, see the Configuring Data Sourcing section.

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

1. Navigate to **Data Source Summary** window.

2. Click **Add Datasource** icon Add Datasource .

For more information, see the Configuring Data Sourcing section.

5   Managing Datasets

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).

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.

This is similar to T2Ts, where you can select data from Data Source, such as file, table or other 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.

5.1   Accessing the Dataset Summary Page

The Dataset Summary page gives access to the various Dataset functions such as create, view, and delete.
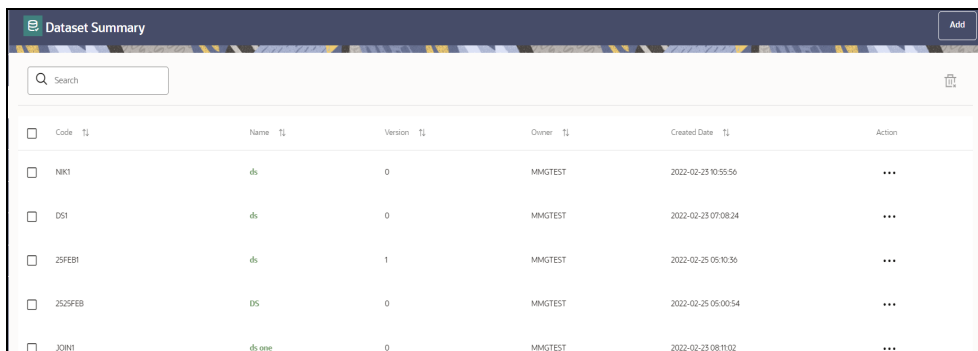
> **NOTE**   Users Groups must be mapped to the SBADMIN role to access the Sandbox Summary page.

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. Click **Dataset** to display the **Dataset Management** window.

   This window displays the dataset records in a table.

   Figure 25: Dataset Summary Page

The Table 10 provides descriptions for the fields and icons on the **Dataset Summary** page.

Table 10: Dataset Summary Details

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 **Add**	Click Add to create a new Dataset. See the Create a Dataset section for more information.
Delete 🗑	Click Delete to delete multiple Datasets.
Action •••	Click the three dots to perform View/Edit/Delete functions on selected dataset.

5.2   Create a Dataset

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

5.2.1   Creating a Dataset

To create a Dataset, follow these steps:

1. Navigate to Dataset Summary page to display the **Dataset Summary** page.

   This page displays the dataset records in a table.

2. Click **Add** to create a Dataset in the Dataset **Summary** window.

Figure 26: Dataset Creation Window

The following steps show the various phases from Dataset creation:

- Dataset Transformation
- Source Selection
- Dataset Creation
- Dataset Transformation
- Summary

5.2.2 Dataset Transformation

Enter basic configuration details in this window.

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

1. Enter the required details in the **Dataset Transformation** window as shown in the Table 11.

Table 11: Dataset Transformation Details

Field	Description
Code	Enter the 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. Click **Advanced Settings** to select the Dataframe Library.

The Advanced Settings window is displayed.

Figure 27: Advanced Settings

3. Select the Dataframe Library for the Dataset as Pandas (Default) or Modin and click **Close**.

4. Click Next [Next] to go to the next step.

5.2.3 Source Selection

This section allows you to define the source of data. Here, you can choose the data structures from an existing datasources.

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

1. Enter the required details in the **Source Selection Details** pane as shown in the following table.

Figure 28: Source Selection
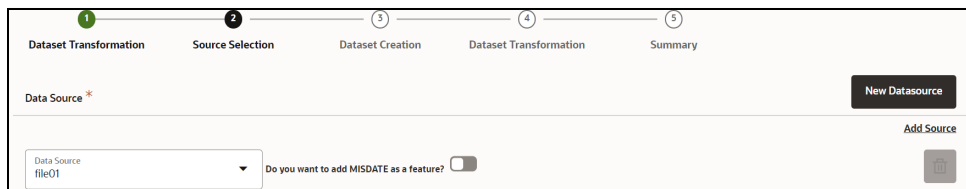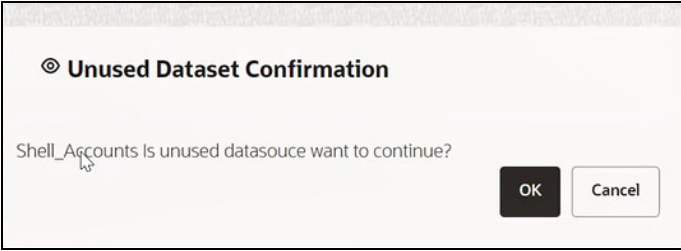
NOTE	The MISDATE feature is displayed only when the Data Source field is selected as a **File Type**.
	If enabled, you need to add the Physical and Feature name for the Data Source.

Figure 29: New Data Source
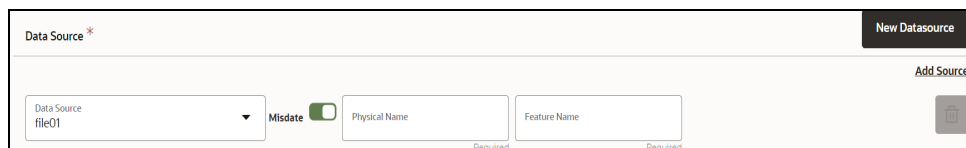
The Table 12 lists the data source details:

Table 12: Data Source Details

Field	Description
Data Source	The Data Source 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. A confirmation message is displayed to tag an unused datasource with workspace. **Figure 30: Confirmation Pop-up** Here, either use the unused or create a new datasource using New Datasource . When you create a new datasource, it becomes as unused. For more information, see Creating Data Source. You can add more Data Sources using Add More **Add Source** link. Use Delete to delete added Data Source.

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

5.2.4   Dataset Creation

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 DatasetCreation, follow these steps:

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**
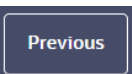
pane. You can use [icon] to select or use [icon] to de-select the data in the **Selected Entities** pane.

Figure 31: Dataset Creation
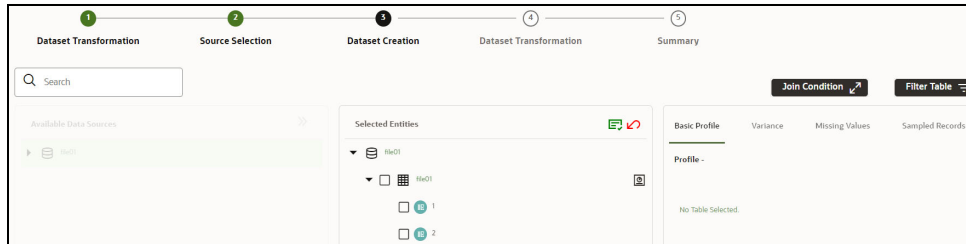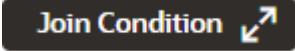
2. For more information on Joins, see Joining Tables section.

    For more information on Filtering, see Filtering Tables section.

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

5.2.4.1   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 1 table

Click Join Condition [Join Condition] to navigate to Join Condition window.

NOTE: As of now, only INNER JOIN is present.

Figure 32: Join Tables
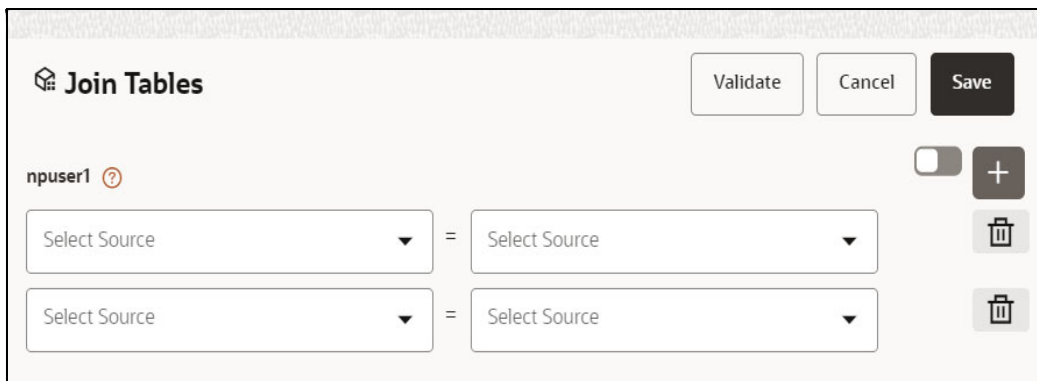
You can join Tables either using drop-down or using the Script.

5.2.4.1.1   Using drop-down

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

5.2.4.1.2   Using Script

1.   To use script, click [toggle] .

The Join Tables window is displayed to add scripts:

Figure 33: Join Tables

You must provide the script in below format:

Figure 34: Script

2.   Click **Validate** to check the Join conditions.

3.   Click **Save**.

5.2.4.2   Filtering Tables

You can also use the filters in dataframe creation from entity.
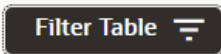
1.   Click Filter Table [Filter Table] to navigate to Filter Tables window.

Figure 35: Filter Tables
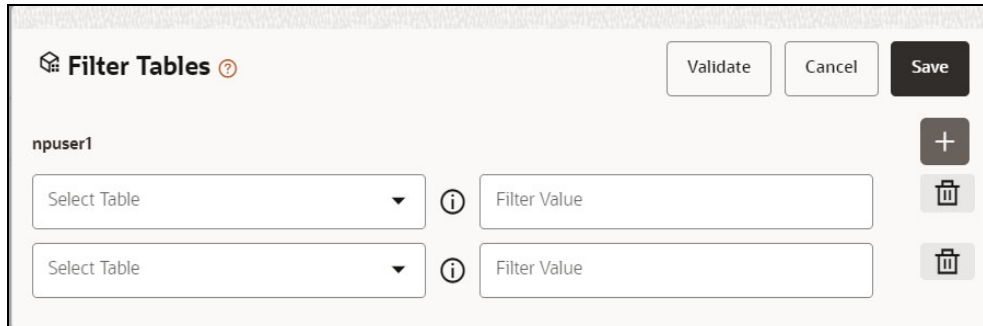
There are certain rules to be followed when adding filters

Figure 36: Rules while Adding Filters

I. Filtervalues will not be applied in tables from datasources joined using joinscript.

II. Only one filtervalue entry is allowed for one table.If multiple entries are present, the last filtervalue entry will be used for that table.

III. Example : Filter Value : *Column1*=1234 AND *Column2*='US'

SELECT * FROM *tablename* WHERE *Column1*=1234 AND *Column2*='US'.

2. Multiple filters can be added by using the + symbol.

3. Click **Validate** to check the Filter conditions.

4. Click **Save**.

5.2.4.3   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, and you must fix these missing values before using the selected table. You can perform three types of data transformations to correct PCA, Scaling, and Missing Values.

Figure 37: Basic Profile Tab
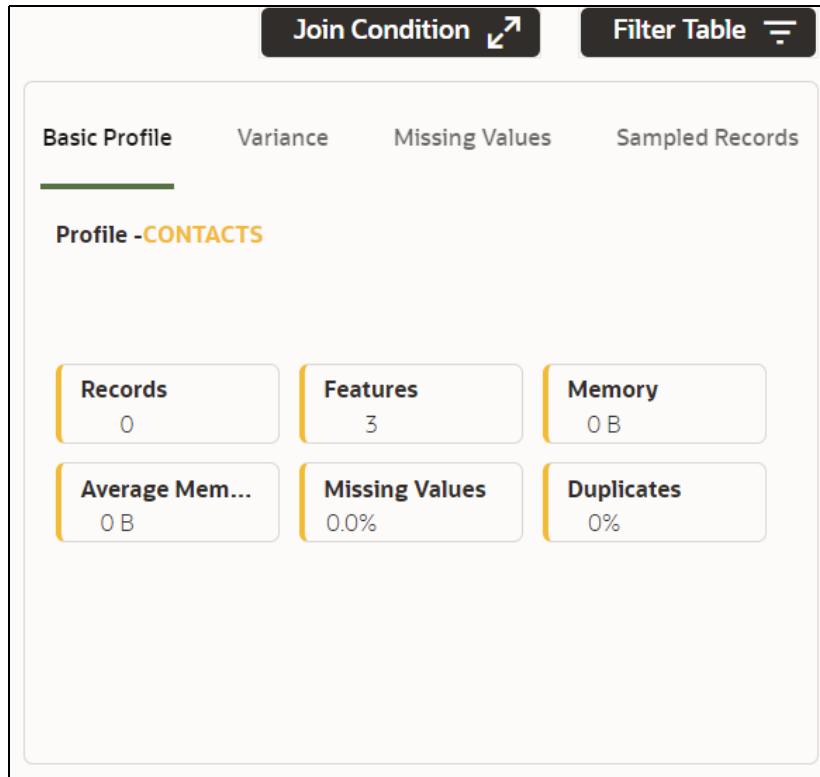
5.2.5   Dataset Transformation

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.

Figure 38: Dataset Transformation
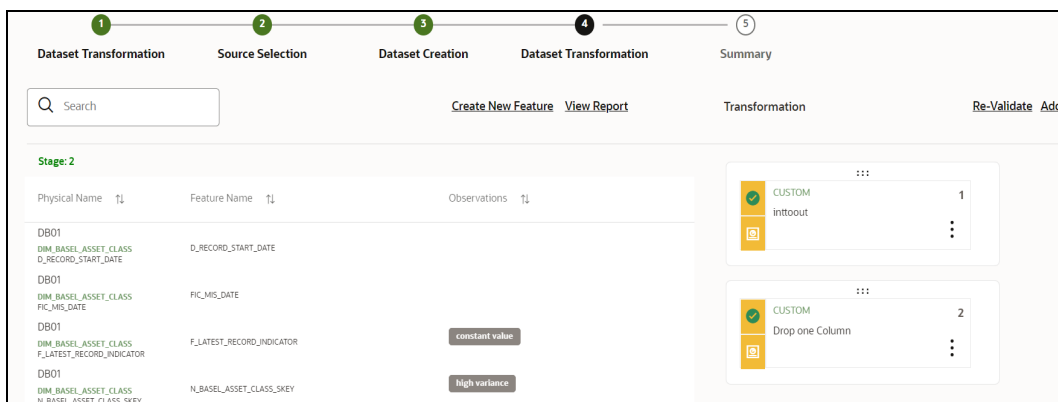
This window shows the following columns:

- **Physical Name:** Shows the Physical Name of column.

- **Feature Name:** Shows the Physical Name of column. You can edit the Logical Name.

- **Observations:** The Observations field shows the value of column. For example, if any column is missing any value, you can easily identity missing value columns. Use Transformation button to fix these values. In addition, if all workspace/Infodom has one value, then Observation column can have Constant value and you can ignore this column to include.

- **Create Feature:** To create a new feature/column by doing operations on the existing features.

- **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 ⋮ icon and select **Re-Validate** option.

- **Add:** 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 following actions

- Re-ordering of transformations

- Insertion of transformation
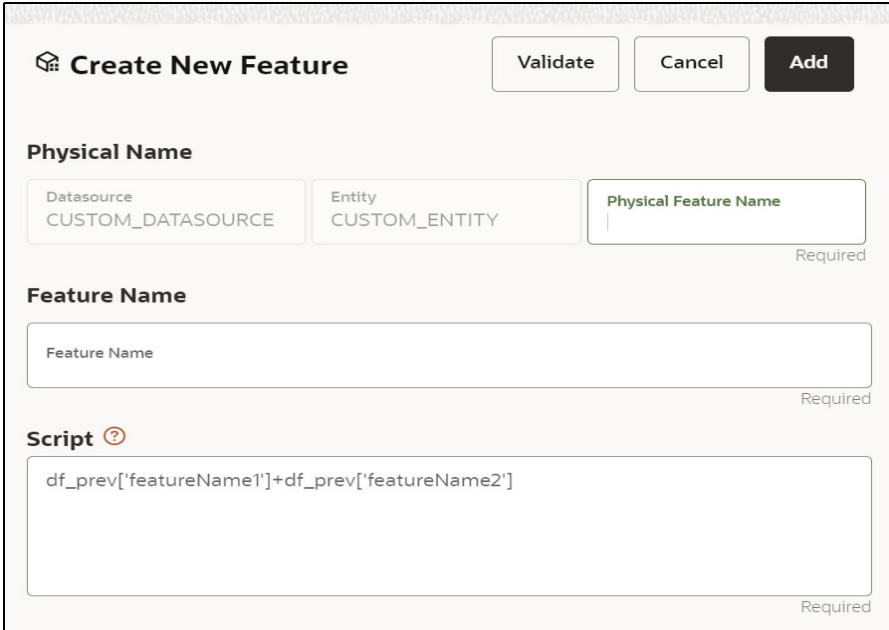
5.2.5.1   Create New Feature

The Create New Feature window allows you to create a new feature. This new feature is used in defining the Transformation. 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 39:   Create New Feature

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.

5.2.5.2   Add Data Transformation

To add the data transformation, follow these steps:

1. Click **Add**.

The Transform window is displayed.

Figure 40: Transform Window
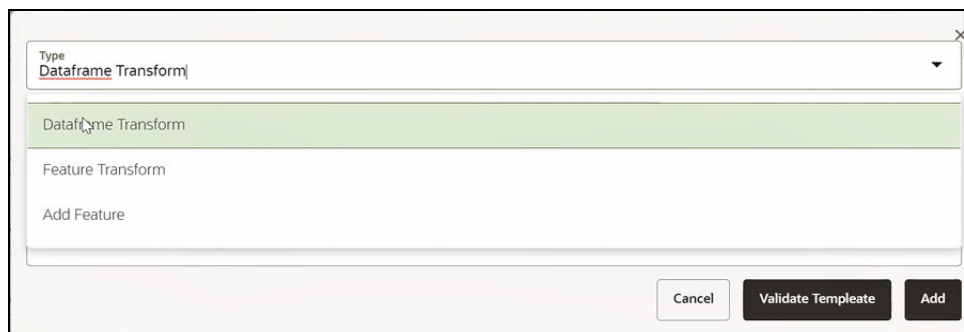
2. Select the Transformation Type.

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 41: Transformation

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 column, then perform following steps to add the transform.

1. Click Add on Dataset Transformation window.

2. Select Type as Feature Transform.

3. Select Transform as Impute Missing Value.

4. Select the Physical Name from Physical Name drop-down list.

5. 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 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)

5.2.5.3   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 42: Report
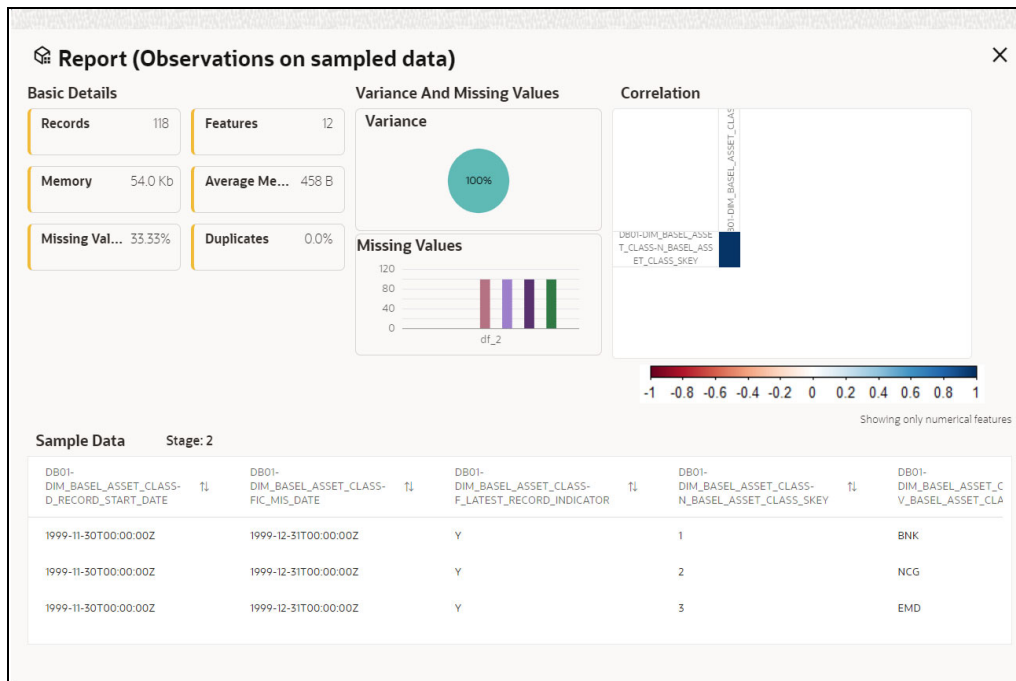
5.2.5.4   Re-ordering of Transformations

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.

5.2.6   Display Summary

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

Figure 43: Dataset Creation Summary
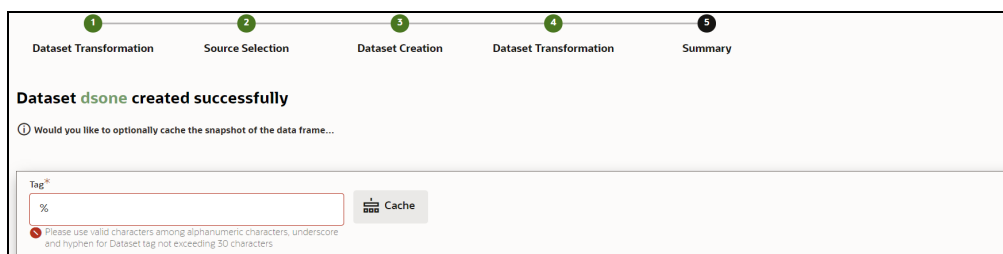
A confirmation message is displayed as "Dataset #DATASETNAME# is created successfully."

Enter the cache (tag) name having alphanumeric chac, underscore and hyphen not exceeding 30 characters and click on Cache.

The Table 13 provides information on the error and the troubleshooting procedure in case of dataset failure.

Table 13: Dataset Failure Details

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"` error while 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**

5.3   Caching the Data

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 data, and you want to use intermediate data for use. For example, if you have 1 million data 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.

> **NOTE**   Unset the https/http proxy details before starting the services or add the relevant entries in no proxy with mmg hosted server details.

5.4   View a Dataset

To edit a Dataset, follow these steps:

1.   Navigate to Dataset Summary Page.

2.   Click ⋯ next to corresponding Dataset and select View ◉ .

5.5   Edit a Dataset

To edit a Dataset, follow these steps:

1.   Navigate to Dataset Summary Page.

2. Click ••• next to corresponding Dataset and select Edit ✎ . You can edit the Dataset fields except Code and Dataset Name.

5.6 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 🗑 .

5.7 Dataset Transformation Backend Details

5.7.1 Creating Dataset Metadata

Metadata for creating a dataset can be defined from UI. A new link has been created for this in the LHS panel.

5.7.2 Caching Dataset

- 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 2 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.

5.7.3 Listing tags of Datasets Cached from UI

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

The Table 14 lists the tag list:

Table 14: Tag List

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

5.7.4 Deleting Cached Dataset

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

The Table 15 lists the delete dataset:

Table 15: Delete 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

5.7.5 Listing all Datasets with Metadata saved from UI

The Table 16 lists the datasets:

Table 16: All Datasets

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

5.7.6 Fetching Dataset in Notebook

Dataset whose metadata has been saved from UI can be fetched in 2 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.

The Table 17 lists the fetching dataset:

Table 17: Fetching Dataset

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

5.7.7 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 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.

The Table 18 lists the caching data frame:

Table 18: Cache User's Dataframe

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

5.7.8   Fetching Data Frame Cached from Notebook

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

The Table 19 lists the fetch dataset:

Table 19: Fetching Data Frame

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

5.7.9   List Tags of all Data frames Cached from Notebook

The Table 20 lists the tags of all Data frames Cached from Notebook:

Table 20: All Data Frames

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

5.7.10   Sample Custom Scripts

The Table 21 displays the sample scripts which the users can refer to create transformations.

Table 21: Sample Script

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

5.7.11   Transformations

The Table 22 lists the various transformations which can be done from the UI.

Table 22: Transformation

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
4.	Encode Cyclical Features	This function encodes a cyclical feature having hour, minute data etc 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
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   Managing Model Pipelines

Model Pipeline allows you to create and publish models based on the workspaces created from datasets in the database. The published models are then deployed in production to be consumed by users. Modelers create models by using the Notebooks in the Data Studio. The Notebooks are used to create training models and the iterations of comparison between various models lead to the elimination of undesired models and filters a few robust ones that can be considered for deployment in production. Modelers then use their better judgement to consolidate their choice and fix on one model - the champion model. The champion model is also called the scoring model or the actual model in this document.

Topics:

- Prerequisites
- Access the Workspace Dashboard Window
- Accessing the Model Pipelines
- Creating, Reviewing, Approving, and Deploying a Model
- Executing Models using Scheduler Service
- Import a Workspace Model Data into a New Model
- Using View Models
- Editing Models
- Deleting Objectives and Draft Models

6.1   Prerequisites

The prerequisites for model pipeline are as follows:

- To create a model, your user profile must be mapped to the Group.
- To create a model, a workspace must be deployed. See the Creating new Workspaces section for more information.
- To approve and deploy a model, your user profile must be mapped to the Modeling Administrator Group.

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

6.2   Access the Workspace Dashboard Window

The Workspace Dashboard window allows you to view the models of launched workspace. The Workspace Dashboard shows the following details of a launched workspace:

- Recently Used model
- Most Used Tags models
- Status of models
- Job Status of models
- Timeline of models

To access the Dashboard window, follow these steps:

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

    The page displays workspace records in a table.

2. Click ![rocket icon] next to corresponding Workspace to Launch Workspace, The **Workspace Dashboard** window is displayed with application configuration and model creation menu.

    Figure 44: Workspace Dashboard
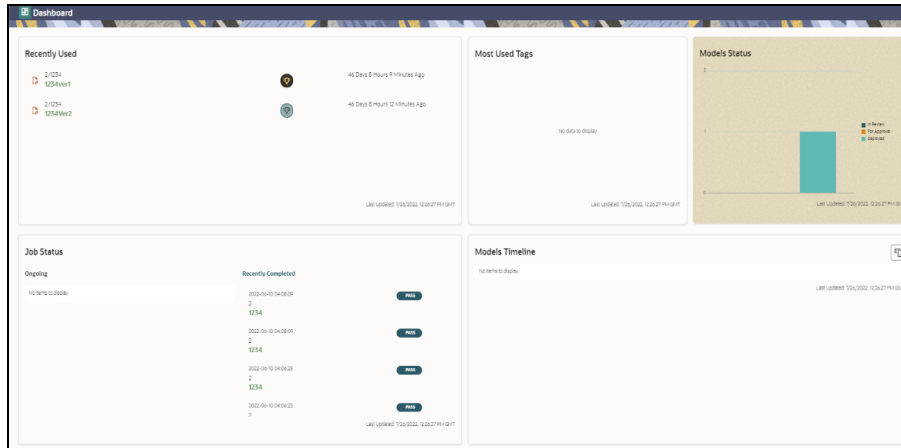
3. Click ![home icon] to navigate back to Workspace Summary window.

6.3   Accessing the Model Pipelines

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

To access the Model Pipeline window, follow these steps:

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

    The page displays workspace records in a table.

2. Click ![rocket icon] next to corresponding Workspace to Launch Workspace.

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

3. Click **Model Pipelines** ![model pipelines icon] to display the **Model Pipelines** page.

    The window displays folders that contain models and model records in a table. When you click on any draft for Objective Name, it shows the type of model, such as champion.

Figure 45: 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.

- **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 objective. You can search the models using the Filter by Version and/or Filter by Status.

  - In the Flat list of models, you can see the following details of published models: Model Name, ID, Description, Objective ID, Objective Name, Owner, Version, Deployment status.

Figure 46: Flat Option

- **Hierarchical option**: The Model Management window shows the hierarchical list of all the models (published and drafts) using the Hierarchical option.

4. Navigate to Model.

Figure 47: Draft Model

The Table 23 provides descriptions for the fields and icons on the **Model Pipeline** page.

Table 23: Model Pipeline

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 to navigate back to Workspace Summary page.
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 and Model	Displays the name of the Objective
Description	Displays the description for the Objective or Model.
Created User	Displays the user ID of the user who created the Model. This information does not display for an Objective.

Table 23: Model Pipeline

Field or Icon	Description
Tags	Displays the tags associated with the Objective or Model. #AML #Alerts This information does not display for an Objective.
Edit Model	Select ✏ to edit the models in Data Studio. See the Editing Models section for more information. This information does not display for an Objective.
Delete Model	Select 🗑 to delete the model from the **Confirmation** dialog box. Click **OK** to process or click **Cancel** to cancel. This information does not display for an Objective.
Create Objectives and Models	Select Add ➕ to display a list with the following options: ● Draft ● Objective To create Models, select Draft. To create Objectives, select Objective. See the Creating Objective (Folders) and Creating Draft Models sections for more information. However, if you want to know about the cycle of model creation, see the Creating, Reviewing, Approving, and Deploying a Model Section.

6.4   Creating, Reviewing, Approving, and Deploying a Model

Model creation and deployment undergoes a workflow of Model Governance where the following types of users in the system have privileges that restrict the activities, they can do in the model creation and deployment workflow.

The Table 24 provides descriptions for the fields and icons on the **Model Pipeline** page.

Table 24: Model Pipeline

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. ● Model Acceptance ● Make Champion – Local **NOTE:** User Groups must be mapped to the MDLDEPLOY role to access Requester functions.

Table 24: Model Pipeline

Field or Icon	Description
Reviewer	You can review models. However, to approve and publish, the model must be approved by a user with approver privileges. • Model Acceptance • Model Acceptance + Promote to Production • Promote to Production **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. • Promote to Production • Make Champion – Global **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:

- Creating Objective (Folders)
- Creating Draft Models
- Publishing Models (Scoring)
- Viewing Model Details
- Understanding Model Governance
- Request Model Acceptance
- Review Models and Move to Approve or Reject
- Approve Models and Promote to Production
- Deploying Models in Production and Make it a Global Champion

6.4.1 Creating 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 corresponding Workspace to Launch Workspace to display the **Dashboard** window with application configuration and model creation menu.

2. Click **Model Pipeline** to display the **Model Pipeline** window.

   The window displays folders that contain models and model records in a table.

Figure 48: Create Model Objectives
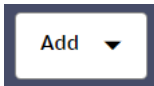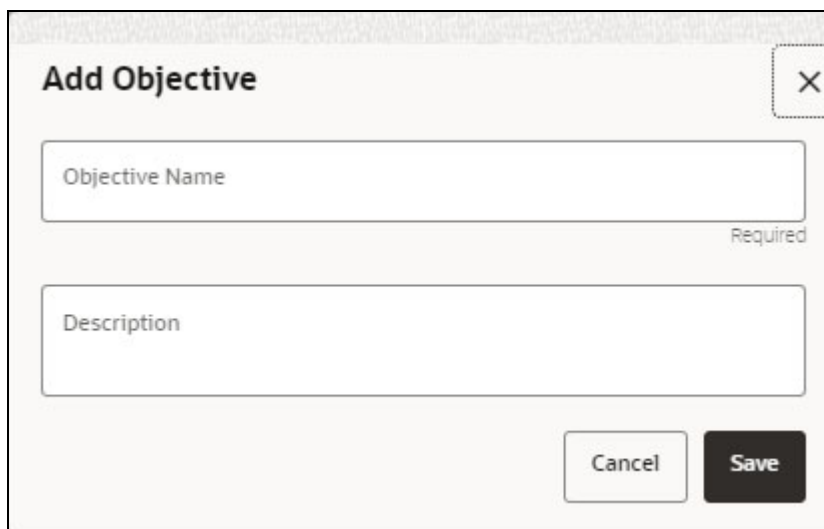
3. Click **Add** [Add ▼] and select **Objective** from the list to display the **Objective Details** dialog box.

Figure 49: Select Objective from Add
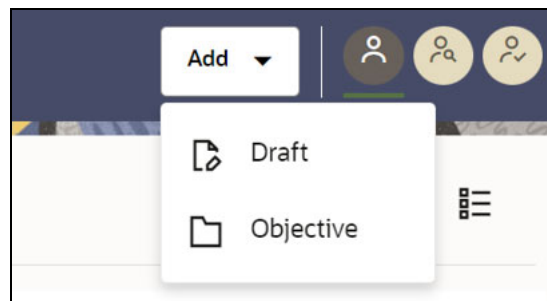
4. Enter details in Objective **Name** and **Description** fields in the Add **Objective** dialog box.

Figure 50: Objective Details Dialog box

5. Click **Save**.

6.4.2   Creating Draft Models

Create Models that are classified as draft models. These models will be reviewed before being sent for Scoring.

To create a draft Model, follow these steps:
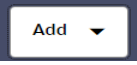
1. Click **Launch Workspace** ![rocket icon] next to corresponding Workspace to Launch Workspace to display the **Dashboard** window with application configuration and model creation menu.

2. Click **Model Pipeline** ![pipeline icon] to display the **Model Pipeline** window.

    The window displays folders that contain models and model records in a table.

3. Click **Add** ![Add button] and select **Draft** from the list to display the **Add Draft** dialog box.

    Figure 51: Create Model
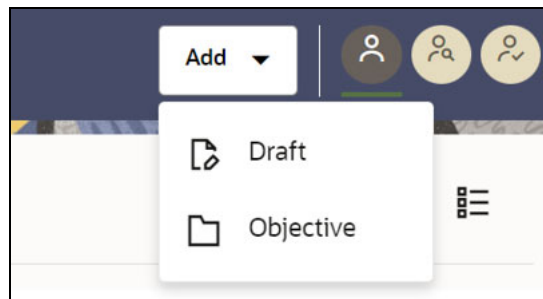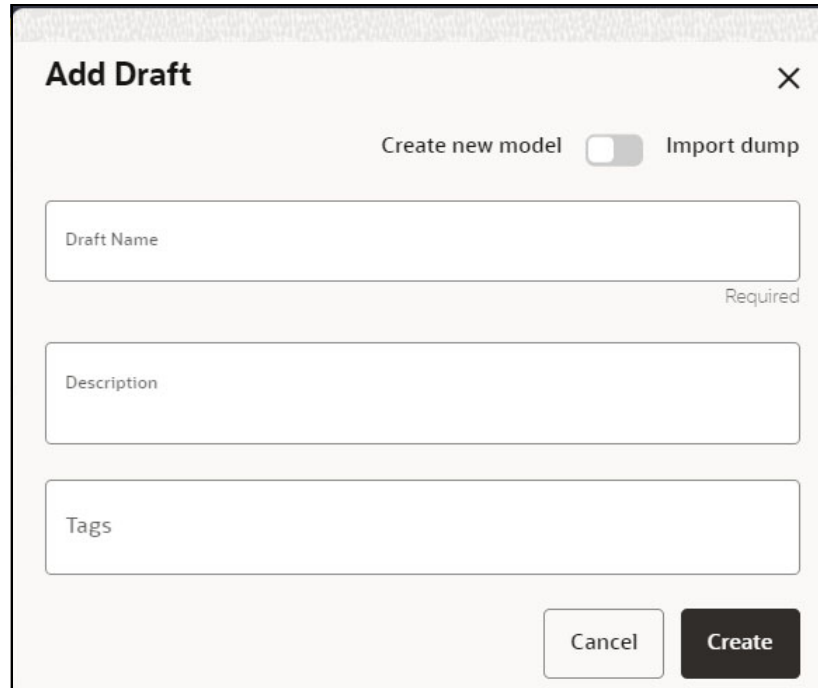
    Create New Model is the default setting in the **Model Details** dialog box. Drag the toggle button to select **Import Dump**. Use Create New Model to start from a blank Notebook in Compliance Studio. Import Dump lets you drag and drop an existing file with model data and modify it. To import a model data dump from another model, see the Import a Workspace Model Data into a New Model section. You can also create a draft under Objective (Folder) also. Click an Objective to open it.

    To create a new model, follow these steps:

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

Figure 52: Model Details - Create New Model

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

c.   Click Create.

6.4.2.1   Creating a Model 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.

The following types of paragraph creation are supported for model creation:

●   Add empty Paragraph: To create Model building scripts in Python.

Figure 53: Create Draft Models in Model Studio Notebooks
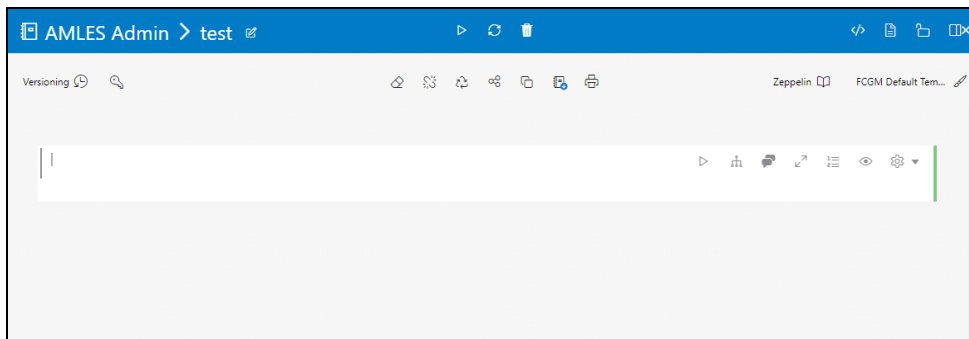
To create Paragraphs in the Model Studio Notebooks, follow these steps:

1.   Open Model Studio in OFS CS. Follow the instructions in the Creating Draft Models section to open it.

2. Click in Model Studio to add a Paragraph in the Notebook and add the scripting instruction to fetch connection objects in the following format. This creates a cx_oracle based connection object to the datadom of the infodom being passed:

```
conn = mmg.getConnection(<infodom name>)
```

> **NOTE** The fetch connection objects function works only with %Python interpreters and requires an Oracle client installed on the server that runs Model Studio.

Click to run the script and display the Notebook ID and the Model Objective ID.

3. Click in Model Studio to add a Paragraph in the Notebook and add in the following format to fetch current Notebook and Model Objective details:

- Notebook Id: **currentNotebookId**

- Model Objective: **objectiveId**

Click to run the script and display the Notebook ID and the Model Objective ID.

4. Click 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
print('threshold value is : ${[threshold]}')
```

Click to run the script and display.

After the Draft Models are created, publish the Notebooks to create Scoring Models. See the Publishing Models (Scoring) section for more information.

6.4.2.1.1 About Notebook

Notebooks are used by data scientists and data analysts to quickly explore the scenario and take actions based on the graphical insights. It offers several advantages over any local script or tool. When correctly set up by the organization, a notebook offers direct connections to all necessary data sources without additional effort from the user.

A notebook segments a computation in individual steps called paragraphs. These paragraphs contain an input and an output section. Each paragraph executes separately and modifies the global state of the notebook.

6.4.2.1.2 Advantages of Using the Notebooks

The following are the advantages of using notebooks:

- Notebooks increase productivity by facilitating incremental improvement. It is easy to use, rerun, and modify only the relevant paragraph.

- Users can develop a notebook and then iterate the notebooks based on their needs.

- Notebooks data are stored in the working memory.

- Notebooks offer advanced interactive visualizations. These range from simple line charts and bar graphs to maps and custom visualizations that are useful for the data scientists to consume the insights.

6.4.2.1.3 Managing your Notebooks

A notebook is a collection of paragraphs and acts as a container to hold one or more paragraphs. Each notebook is a collection of documentation and snippets of executable code. The notebook allows large scripts to be broken into a modular collection of executable code with tailored results. Different languages, such as Groovy, Scala, Python, and Oracle's own property graph query language (PGQL), can be combined into one notebook. Each notebook is mapped to the role of the logged-in user.

After creating a notebook, you can create Paragraphs in a notebook. To know more about the actions performed in a notebook, see the Common Screen Elements in a Notebook section.

6.4.2.1.4 About Paragraphs

A paragraph is a piece of code that can be executed to obtain the result. Paragraph offers a workbench to author code or query using interpreter-friendly scripting languages supported in Notebook Server.

Data scientists and analysts can use paragraphs to explore financial crime data by authoring queries and applying these queries on the financial crime data lake to fetch results in various formats. This enables them to discover and model financial crime patterns rapidly.

Paragraphs are like the pages of the Notebook. A notebook can be developed with multiple types of paragraphs and executed to visualize the results. The Notebook Server notebooks allow you to create as many paragraphs as possible by using a normal paragraph option or the paragraph with interpreters. In Notebook Server, you can add the following interpreter-based paragraphs:

- jdbc paragraph
- pgx-java paragraph
- pgx algorithm paragraph
- pgx-python
- pgql paragraph
- spark paragraph
- pyspark paragraph
- fcc-python
- fcc-python-ml4aml
- fcc-python-sane
- markdown

Table 25 lists the Ready-to-use interpreter in Compliance Studio:

Table 25: Ready-to-use interpreter

Interpreters	Description
ore Interpreter	The ore interpreter is a ready-to-use interpreter used to connect to BD schema. This Interpreter is used to write a notebook in R language (ORE: Oracle R Enterprise). Additional configuration must be required as a prerequisite or post-installation with a manual update in the interpreter setting. In the ore Interpreter, you can configure the database server's Oracle system identifier (SID) to configure different databases, configure to display the output number of rows, and set the database's hostname server to connect the ore Interpreter, schema details, etc.
fcc-python Interpreter fcc-python-ml4aml fcc-python-sane	The fcc-python interpreter is used to write Python code in a notebook to analyze data from different sources, machine learning, artificial intelligence, etc. In the fcc-python interpreter, you can configure the Python installed path, set the maximum number of results that must be displayed, change the Python version, add Python Packages, etc. The python interpreter uses a python virtual environment. Out-of-box Compliance Studio has three variants of python interpreters, fcc-python, fcc-python-ml4aml, and fcc-python-sane.
jdbc Interpreter	The jdbc interpreter is a ready-to-use interpreter used to connect to Studio schema. This Interpreter is used to connect and write SQL queries on any schema without any restriction. In the jdbc Interpreter, you can configure schema details, link Wallet Credentials to the jdbc Interpreter, etc.
md Interpreter	The md interpreter is used to configure the markdown parser type. This Interpreter is used to display text based on Markdown, which is a lightweight markup language. The connection does not apply to this Interpreter.
pgql Interpreter (part of PGX interpreter)	The pgql interpreter is a ready-to-use interpreter used to connect the configured PGX server. This Interpreter is used to perform queries on the graph in Compliance Studio. PGQL is a graph query language built on top of SQL, bringing graph pattern matching capabilities to existing SQL users and new users interested in graph technology but who do not have an SQL background.
pgx-python (part of PGX interpreter)	The pgx-python interpreter is a ready-to-use interpreter used to connect to the configured PGX server. It is a python-based interpreter with a PGX python client embedded to query the graph present in the PGX server. By default, this Interpreter points to ml4aml Python Virtual environment.
pgx-algorithm Interpreter (part of PGX interpreter)	The pgx-algorithm interpreter is a ready-to-use interpreter used to connect to the configured PGX server. This Interpreter is used to write an algorithm on the graph, and it is also used in the PGX interpreter.
pgx-java Interpreter (part of PGX interpreter)	The pgx-java interpreter is a ready-to-use interpreter used to connect to the configured PGX server. It is **Java11** based interpreter with a PGX client embedded in it to query on graphs present in the PGX server.

Table 25: Ready-to-use interpreter

pyspark Interpreter	The pyspark interpreter connects to the big data environment by default. Users must write code for connection either in the Initialization section or the notebook's paragraph.

This Interpreter is used to write the pyspark language to query and perform analytics on data present in big data. This requires additional configuration, which must be performed as a prerequisite or as post-installation with the manual change of interpreter settings.

In the pyspark Interpreter, you can configure the Python binary executable to use for PySpark in both driver and workers, set true to use Python, else set to false, etc. |
| spark Interpreter | The spark interpreter connects to the big data environment by default. Users must write for connection either in the Initialization section or the notebook's paragraph.

This Interpreter is used to perform analytics on data present in the big data clusters in the Scala language. This requires additional configuration, which must be performed as a prerequisite or as post-installation with the manual change of interpreter settings.

In the spark interpreter, you can configure the cluster manager to connect, print the Read–eval–print loop (REPL) output, the total number of cores to use, etc. |
| ore Interpreter | The ore Interpreter has been deprecated. This interpreter is not recommended since it will be removed in future versions of OFS Compliance Studio. The "R" Interpreter will be introduced instead of ore Interpreter. |

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

6.4.2.1.4.1 *Creating a Paragraph*

Paragraphs can be created using interpreters that are supported in Notebook Server.

To create a paragraph in a notebook, follow these steps:

1. Navigate to the Notebook page.

2. Hover above or below a paragraph and click ⊕ to add a paragraph to your Notebook.

   To create a paragraph based on the interpreter, you can choose the respective icons displayed for the interpreters as shown. You can hover over the icons to view the name of the interpreter-based paragraph.

   Figure 54: Interpreter Icons
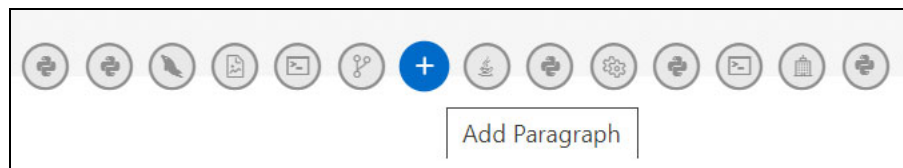
The added paragraph can be used to write queries based on your requirement to plot and visualize your data. For more information about data visualization, see Data Visualization.

6.4.2.1.4.2 *Writing a Sample paragraph*

After creating the paragraph using the available interpreter, you can write your query in the paragraph to execute and display the visuals or analysis as results.

Let us take an example to write a **PGQL Paragraph** using **PGQL Interpreter** and querying for the graph plots.

To create and write a PGQL query in the paragraph, follow these steps:

1. Hover above or below a paragraph and click ⚙ **Add PGQL Paragraph.**

2. Specify the title for your paragraph. In this case, **Transfer of Money Through External Entity.**

Figure 55: Transfer of Money Through External Entity

3. Specify the PGQL query that you want to execute from the Global Graph and click **Execute**.

Figure 56: Execute a Sample PGQL Query
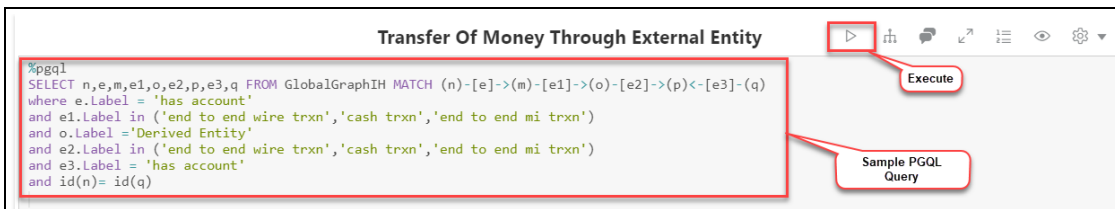
The graph results are displayed as shown.

Figure 57: Graph Results
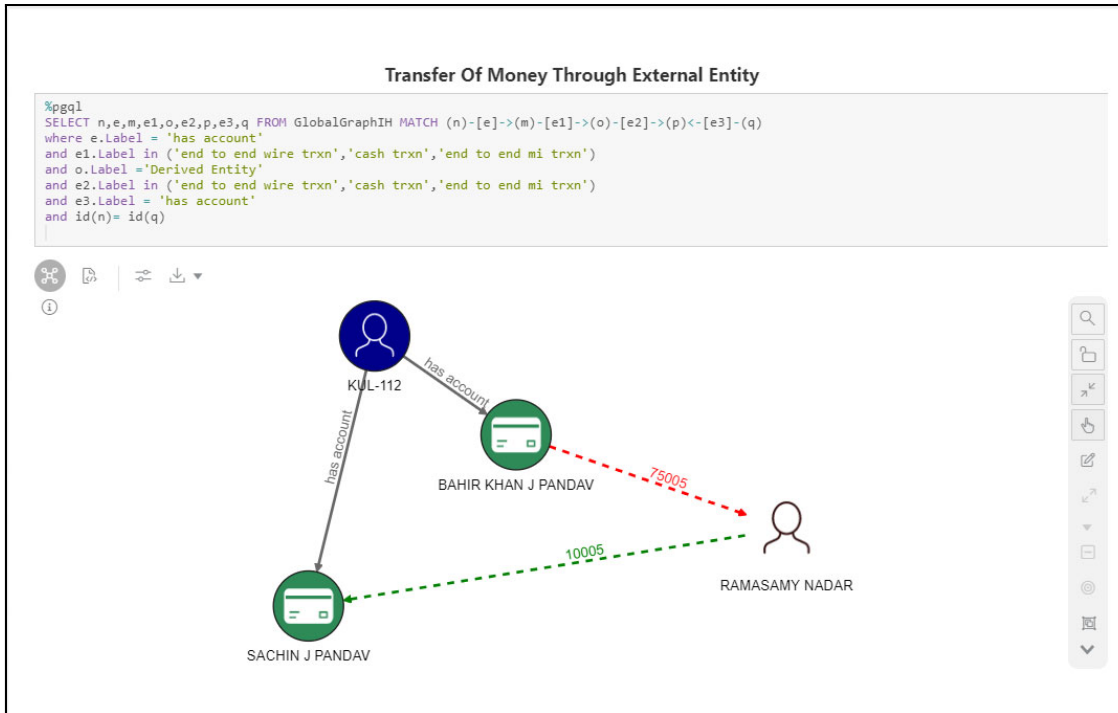
Similarly, you can help us, other interpreters, to create your paragraph, write your queries and execute them to visualize the data. For more information about data visualization, see Data Visualization.

6.4.2.1.4.3 *Paragraph Dependencies*

The Paragraph Dependencies help you to add dependencies between paragraphs. The dependents of a paragraph are automatically executed after the original paragraph itself, or any graph manipulation on the original paragraph is executed.

To create paragraph dependencies, follow these steps:

1. Click **Dependency** on the **Paragraph Settings Bar** of a paragraph. The **Dependency Mode** window is displayed.

2. Select the paragraphs to add as dependents.

   The order in which the paragraphs are selected appears as a number over the selected paragraphs. The number indicates the order in which the dependent paragraphs will be executed.

Figure 58: Financial Crime Graph Patterns screen
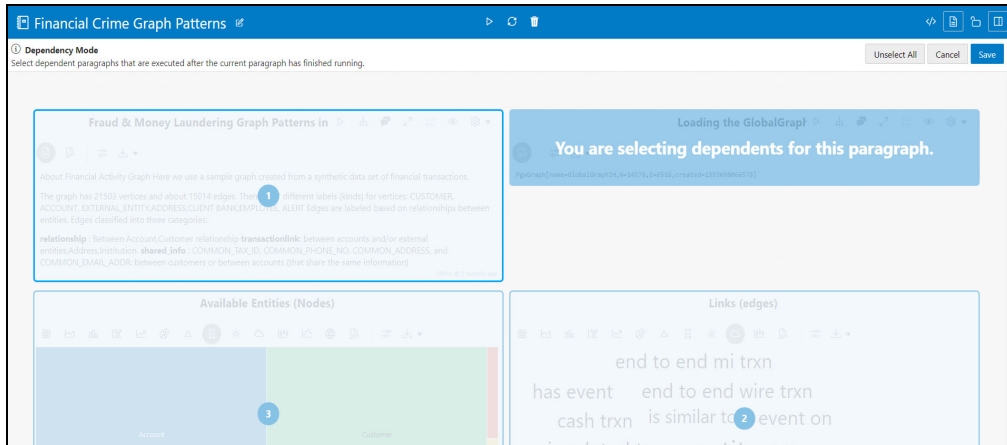
3. Click **Save**. The changes are saved. Every time a paragraph is executed or graph actions are applied, its dependent paragraphs will be executed automatically.

6.4.2.1.4.4 Running All Paragraphs

You can run all the paragraphs in the notebook by clicking ▷ **Run Paragraphs** in the notebook toolbar to execute the complete notebooks and view the data in a single go.

Figure 59: Run the Paragraphs button
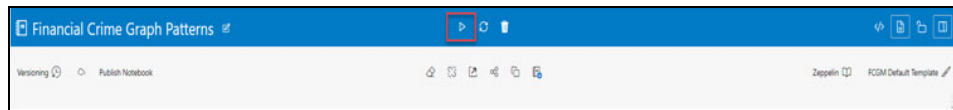
When you execute all the paragraphs at once, the paragraphs are executed in the top to bottom order. If a paragraph was deleted during the **Run Paragraphs** job execution, it is ignored and continues for the rest of the paragraphs.

6.4.2.1.5 Data Visualization

The notebooks provide you the flexibility to visualize and customize your data. See available data visualizations in the Using Data Visualizations.

Notebook Server supports multiple layouts to give the resemblance of familiar systems such as Project Jupyter and Apache Zeppelin.

6.4.2.1.6 Jupyter Layout

The Jupyter Layout is a single-column layout that is something that the formal documents look like. To get the Jupyter-Experience, paragraphs can hide certain aspects, such as Input, Selection, and Result. For example, this allows showing only HTML when a Markdown paragraph is executed.

Figure 60: Jupyter Layout
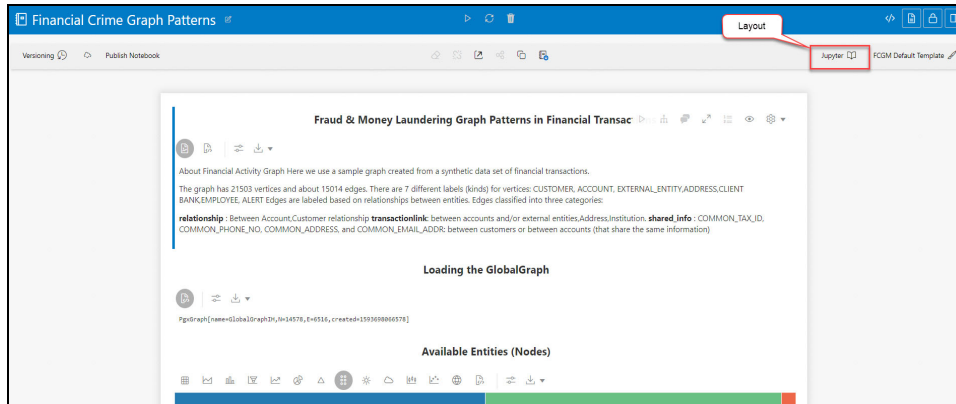
6.4.2.1.7   Apache Zeppelin Layout

The Zeppelin Layout allows you to resize paragraphs and have multiple paragraphs next to each other.

Figure 61: Zeppelin Layout
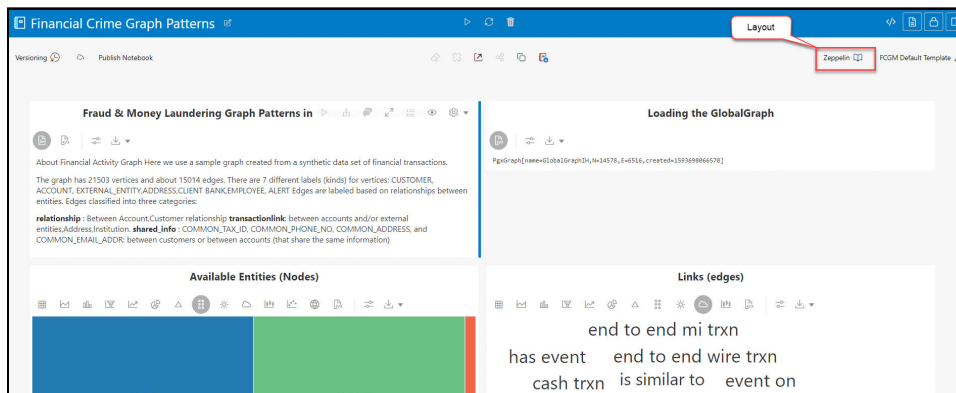
6.4.2.1.8   Comments

You can add comments on the paragraph to highlight changes or suggestions to the visuals. Comments can be added from the user interface, and they can delete their own comments.

You can click 💬 **Comments** in the paragraph toolbar. The Comments view shows all paragraph comments, such as:

- **Add Comment**

  Use a text area field labeled as New Comment on the bottom of comments.

- **Delete Comment**

  Use a bin icon on the right of comments (you can only delete your comments).

Figure 62: Interpolating nodes
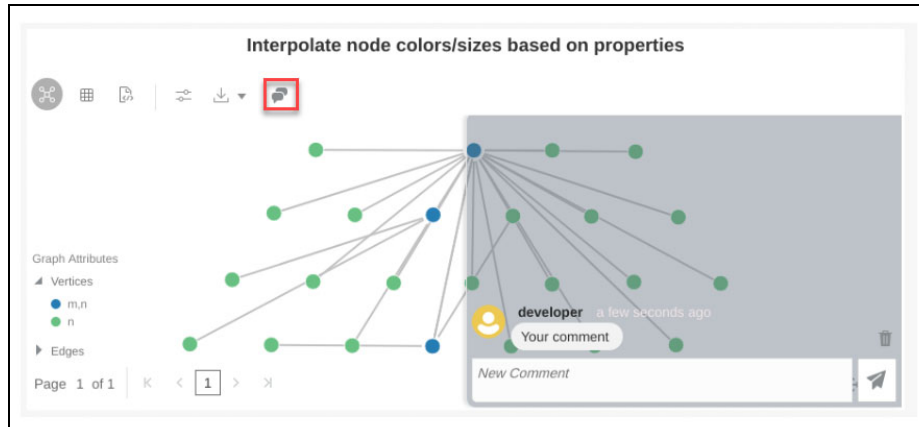
6.4.2.1.9 Dynamic Forms

Dynamic Forms in Notebook Server are the dynamic input fields that you can create for the paragraphs. When the paragraphs with the dynamic input fields are executed, the existing paragraph properties are replaced with these dynamic values, displaying the results. For more information, see Dynamic Forms.

6.4.2.2 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 OFS Compliance Studio, the attach workspace call will be made from mmg service and proper workspace will get attached.
	If the OFS Compliance Studio is opened outside of MMG application, 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 the using the Training and Experimentation link types. One flow can be break into 2-3 flows for execution purpose.

When a model is edited using the Pipeline Designer/Data Studio, and executed, 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:
	a. 1. P1 -> P1a -> P1b -> P2
	b. 2. P1 -> P3 -> P2

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

- Model Report
- Download
- Delete

6.4.2.2.1 Model Report

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

To view and download Report, follow these steps:

1. Open a Model in Pipeline Designer.

2. Click **Model Report**.

   The Model Report window is displayed.

   Figure 63: Model Report
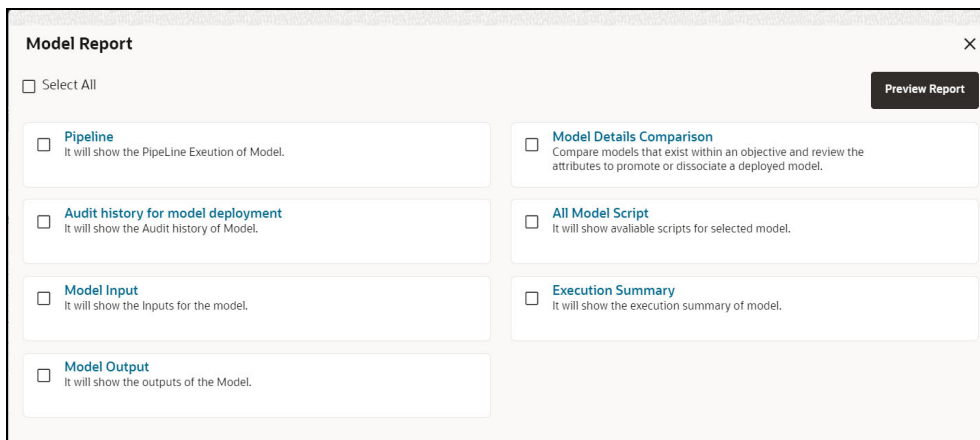
3. Select the Parameters to generate the report.

4. Click **Preview Report**. The report is displayed based on selected parameters. The following is a sample report.

Figure 64: Model Report
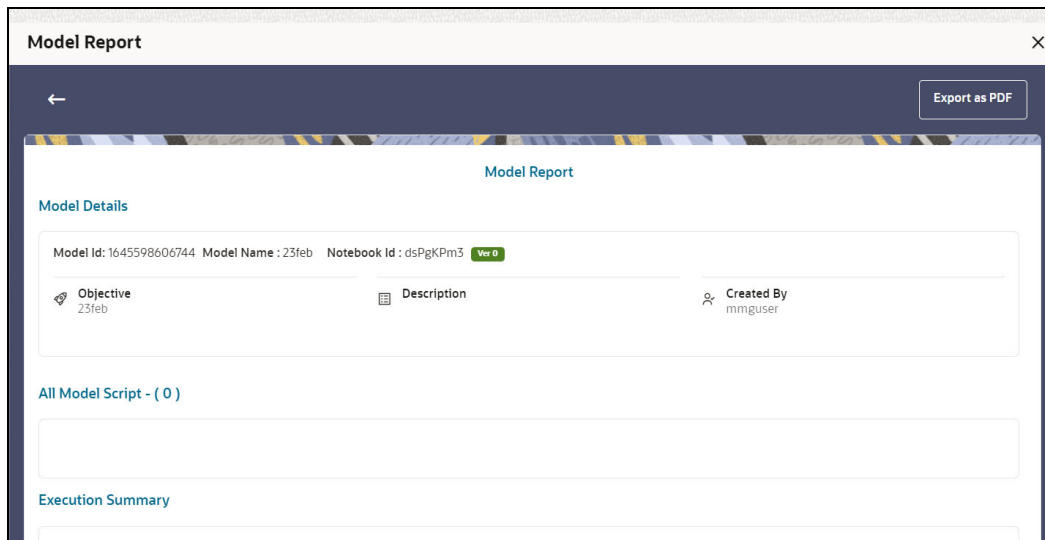
5.   Click **Export as PDF** to save the report as PDF in local system.

6.4.2.2.2   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.

6.4.2.2.3   Delete

This option is used to delete the Model.

6.4.2.2.4   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. Audit Trail window also captures the clone information.

To clone the model details, follow these steps:

1.   Open a Published Model in Pipeline Designer.

Figure 65: Clone Model
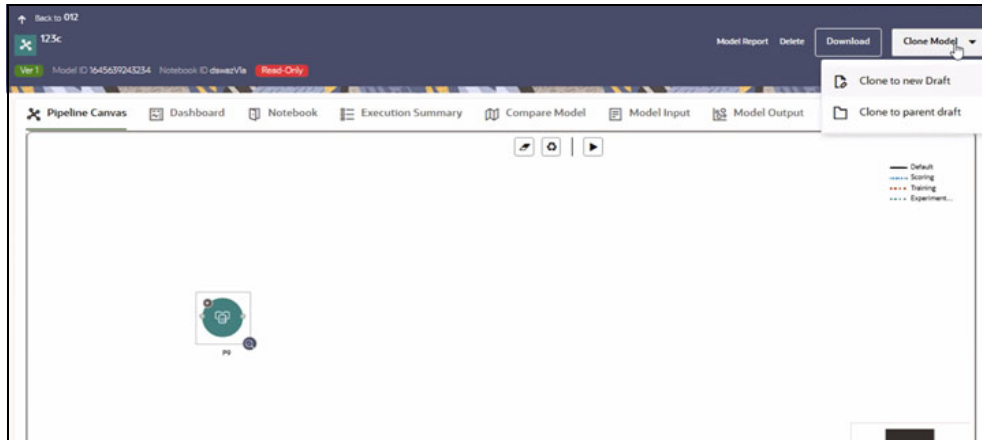
2. Select Clone to new Draft or Clone to Parent draft.

6.4.2.2.5 Pipeline Designer

The following sections are available on the Pipeline Designer window:

- Pipeline Canvas
- Execution Summary
- Notebook
- Dashboard
- Compare Model
- Model Input
- Model Output

| NOTE | 1. Published models have the 'Dashboard' as the default tab. |
| | 2. Drafts models have the 'Pipeline Canvas' as the default tab. |

6.4.2.2.6 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

6.4.2.2.6.1 *Creating a Pipeline*

To create a paragraph using pipeline, follow these steps:

1. Navigate to the **Pipeline Designer** page. For more information, see the Creating Draft Models section.

2. Drag and drop the Paragraph widget from the widgets pane in the upper-left corner of the designer pane.

    Drag and drop the 'START' widget from LHS widget list in the canvas screen. You can edit this Node except script. 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. Import of dsnb/ipynb will not have a start widget.

3. Hover the mouse over the Paragraph widget and click to add Basic Details.

    Figure 66: Pipeline Designer
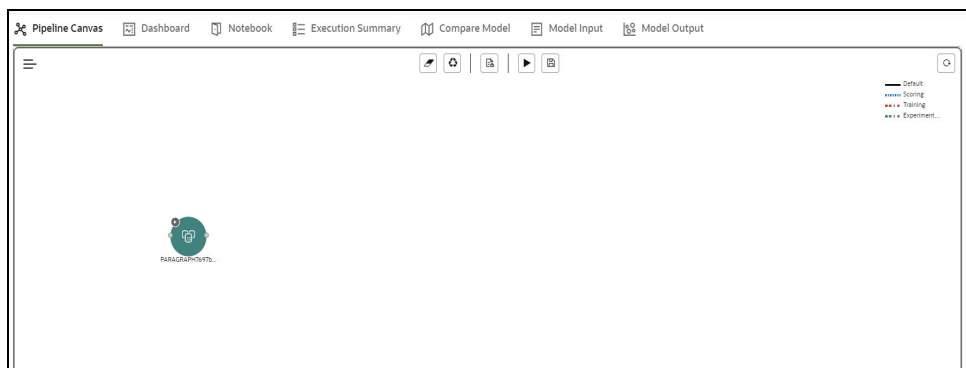
Figure 67: Basic Details for Paragraph

4. Provide details as described in the Table 26:

Table 26: Basic Details for Paragraph

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 paragraph details are displayed. Keep the Track Output to ON in case you want to execute the paragraph from Execution tab.
Script	Shows the script.

You can perform the following functions on Pipeline Canvas window:

■ Clear Execution Details: Clears the execution details.

■ Re-initialize session: Allows you open a new session.

■ Studio Notebook: Allows you to view the paragraphs in Studio application.

■ Publish: Allows publishing the Notebook.

■ Save: Allows saving the Notebook.

- Delete: Click to Delete 🗑 icon 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.

- Download: Use Download to download the current working version in opened in canvas

    ▪ Execute: Allows to execute the Notebook. For more information, see Executing a Pipeline

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 doesn't want to execute all the paragraphs and want to execute only Paragraph 1 and Paragraph 2. Paragraph 1 takes 15 mins time for execution and Paragraph 2 takes 10 mins 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 68: Confirmation Pop-up

Do you want to close the current session? (Note that session will be closed automatically after 10 minutes of inactivity!)

Yes   No
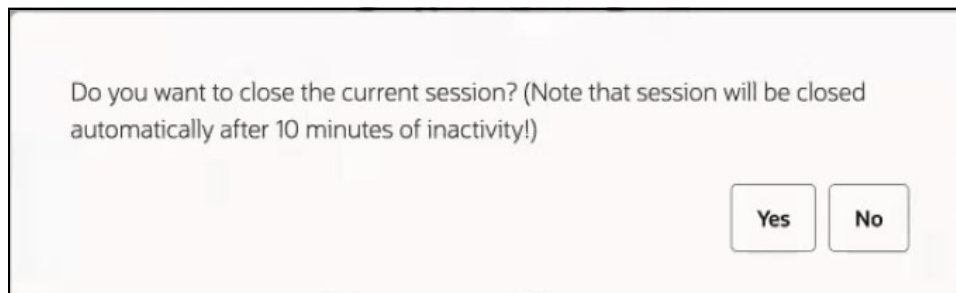
- If user clicks Yes, then execution thread will be closed after for this given session time.

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

User can configure these values in Application.xml file. For more information, see the OFS Compliance Studio Installation Guide.

6.4.2.2.6.2 *Viewing a Pipeline*

The Pipeline canvas window allows you to view the Pipeline using 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.

- Hide/Show Overview ⊙ : Shows or Hides the overview of widgets.

- Show Full Labels ▤ : Shows the full name of activity. By default, it shows only 14 characters of label name.

6.4.2.2.6.3 Creating Script

The Script Template window allows you to manage scripts. These scripts can be called for paragraph.

Figure 69: Script Template
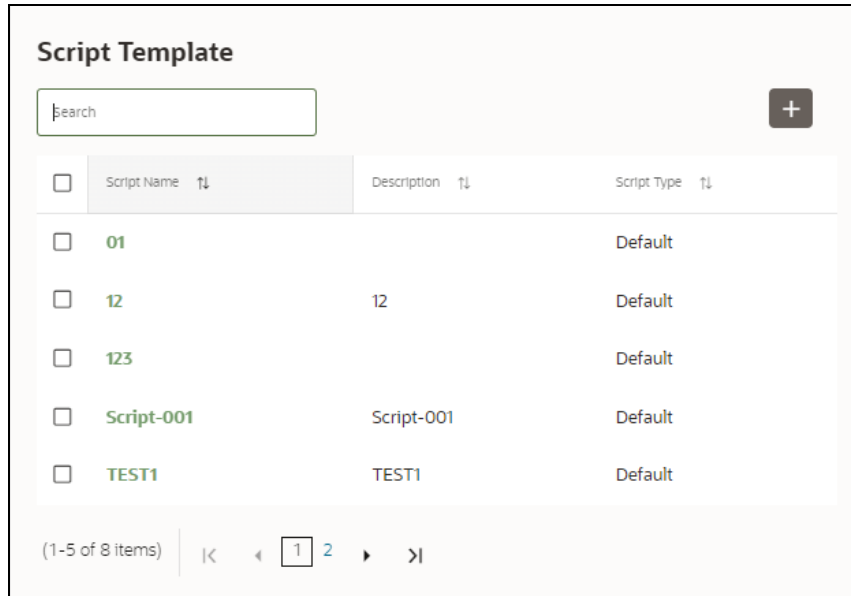
To add a script, follow these steps:

1. Click Tasks ☰ to navigate to Script Template.

2. Click Add ➕ in Script Template window.

3. Enter the details in the Table 27:

Table 27: Script Template Details

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

Figure 70: Script Details

You can select Parameter set and link type for execution in the canvas dashboard tab. This parameter set is available from already saved list. When you select this and there are no common keys, it is displayed along with the original parameters. If there are any common keys, it will be replaced with the chosen parameter set keys.

A reset button returns to original parameters.

6.4.2.2.6.4 Using Link Connector Nodes

You can use a dummy node that helps to service connecting functions. If a paragraph in studio got deleted, this type of widget will be placed instead of the paragraph widget in pipeline canvas Node name for deleted node 'Connector'. This helps if the links are not broken. You can drag and drop this node but will not be associated with any paragraph. If required, you can delete it and provide name and description for it.

To use Link connector nodes, follow these steps:

1. Click Tasks ☰ to navigate to Script Template.

2. Select Link Connector and click ⊕ . This is a dummy node with no paragraph is created/ associated on node save. During execution, this is used in to execute API, but will not get executed. It behaves like non-physicalized paragraph widget on execute.

    This node can be used as a dummy start node or connector node.

6.4.2.2.6.5 Executing 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, follow these steps:

1. Click Execute to view Execute Pipeline window.

   Figure 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.

■ **Scenario Creation:** Scenarios 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. Each scenario consists of parameters with a specific value to each parameters. 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 scenario and description which is not mandatory.

You can save Key Value parameter set using the Save Scenario option. To Save Scenario, enter the Threshold Value and Description in the Scenario window.

■ **Importing Scenario:** These saved scenarios can be imported during the execution. It will replace the values of the parameters from the chosen scenario.

To import scenario, follow these steps:

i. Click "Open from saved scenario".

The Threshold Code window is displayed.

ii. Select the scenarios from the available threshold code. You can import multiple scenarios in the same execution instance. In that case, if there are any common keys, value will be replaced with that from the latest scenario imported.

4. You can add new parameters using **Add** ⊞ . 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 scenarios 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 72: Example of Pipeline Designer
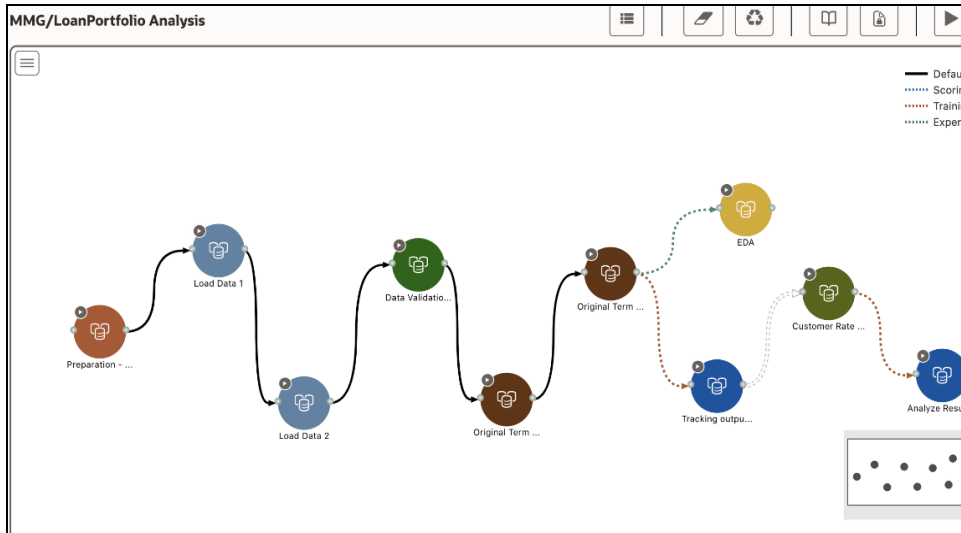
- Here, if you want to execute this Notebook for scoring purpose, then the flow will be executed till EDA paragraph with default paragraphs. To perform this, Click Execute Notebook and select Links as Scoring.

Figure 73: Execute Pipeline
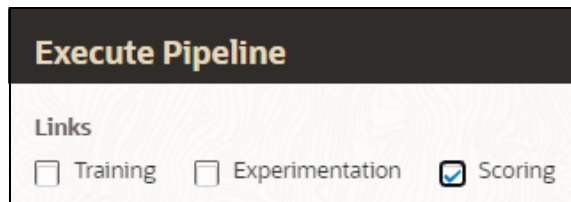
- Similarly, if you want to execute this Notebook for training purpose, then the flow will be executed till Analyze Result paragraph with default paragraphs. To perform this, Click Execute Notebook and select Links as Training.

Figure 74: Execute Pipeline
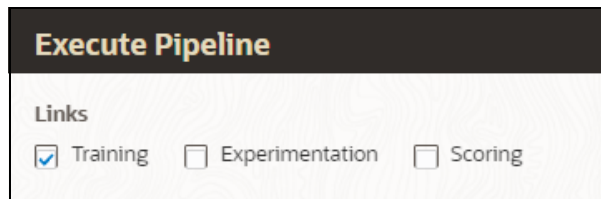
6.4.2.2.7   Execution Summary

This section of Pipeline Designer shows the summary of executed pipelines.

Figure 75: Execution Summary
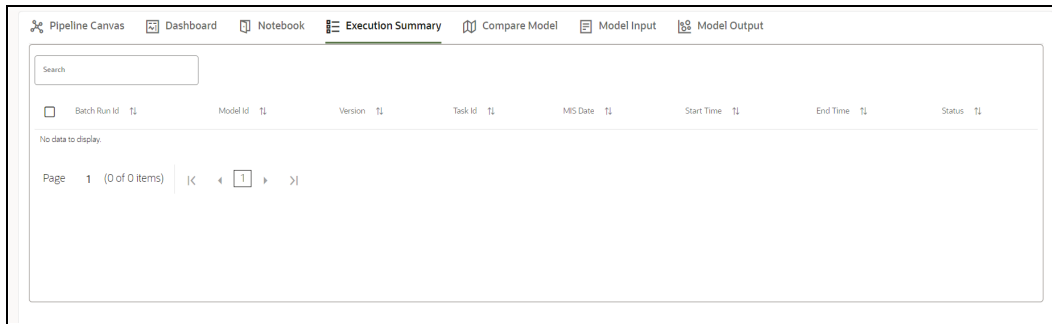
Comparing Execution

The Compare option allows you to compare the two executions.

To compare, follow these steps:

1. Navigate to Execution Summary window.

2. Select the executions using the corresponding check-boxes.

3. Click Compare ⬛. The Execution Comparison window is displayed.

6.4.2.2.8 Dashboard

a. This section of Pipeline Designer shows the Model Dashboard to view the execution output of Model. You can also execute the Models from here.

b. Note:

- For Published models, the 'Dashboard' is the default tab.

Figure 76: Dashboard
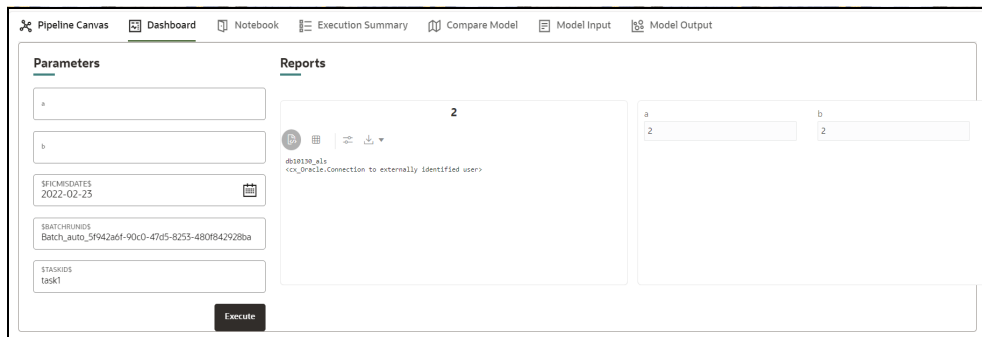
> **NOTE**   There is no Cancel button in Settings tab of reports in the Dashboard tab. You can press 'Escape' key to close the Settings.

6.4.2.2.9 Compare Models

This section allows you to compare the models with Champion model.

To compare, follow these steps:

1. Navigate to Pipeline Designer window.

2. Click **Compare Model** tab.

   This shows the following comparison details:

   - Model Properties

   - Model Input

   - Model Output

   - Model Script

   - Audit Trail

   Figure 77: Compare Model
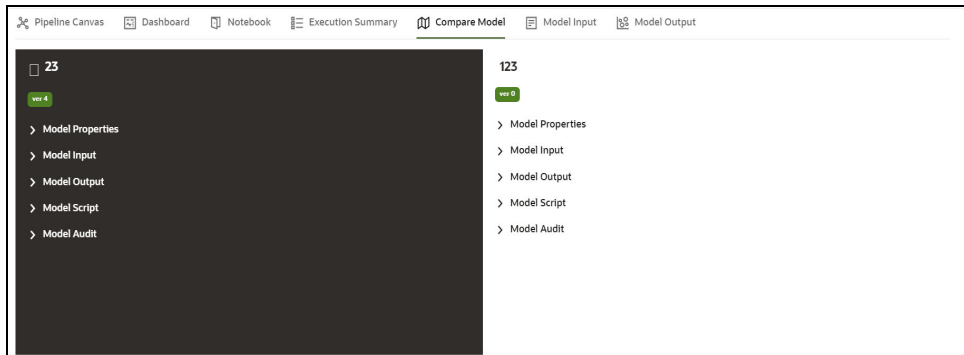
6.4.2.2.10 Model Input

This section shows the input details those were given for Model execution.

Figure 78: Model Input
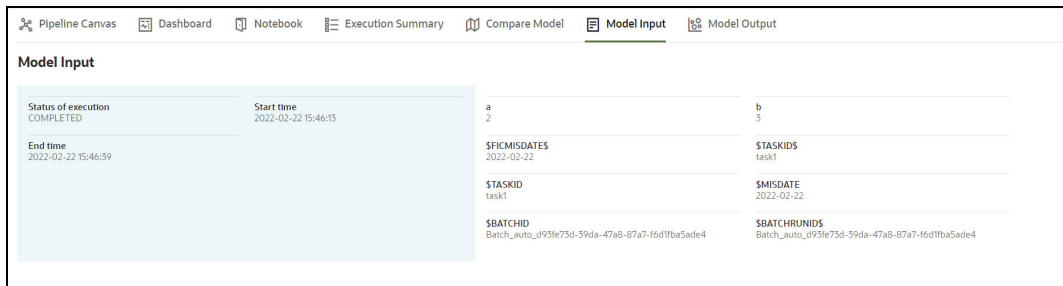
6.4.2.2.11 Model Output

This section shows the output details of model post execution.

Figure 79: Model Ouput
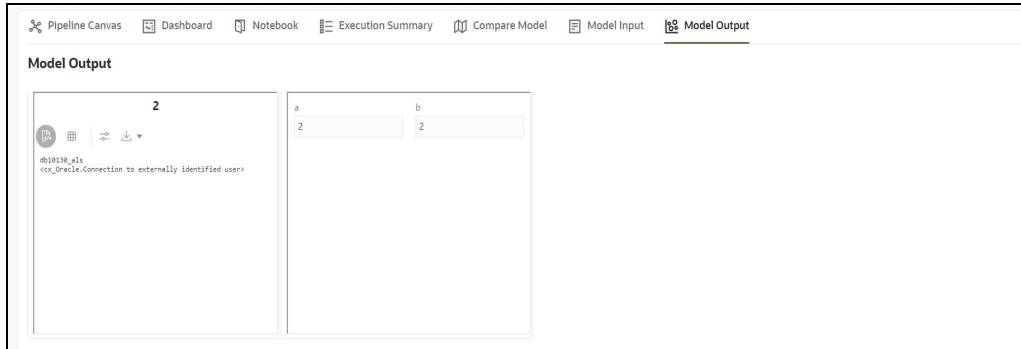
6.4.3   Publishing Models (Scoring)

After creating the Draft Models, publish the Notebooks, which have the Model script.

To create a Scoring Model, follow these steps:

1. Create a Draft Model. See the Creating Draft Models section for more information.

2. Click [···] next to corresponding Draft Model and select Publish Data Studio [icon] to open Open Model window.

Figure 80: Publish Model
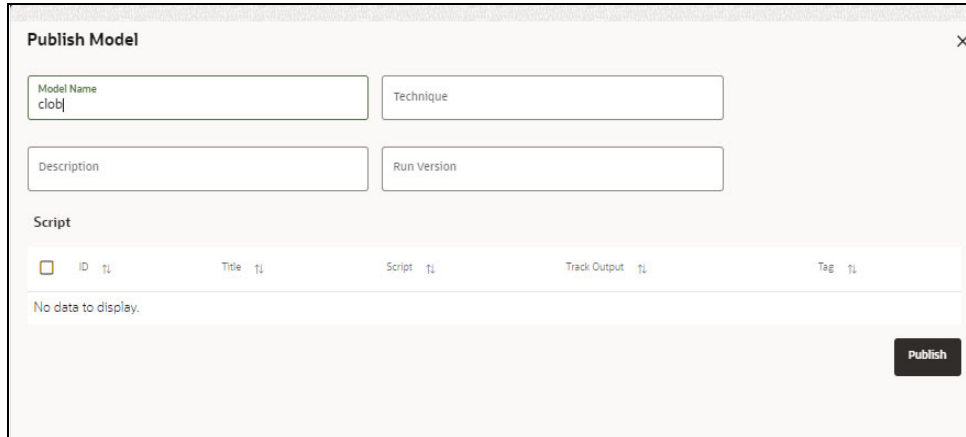
3. Enter the details as shown in the Table 28.

Table 28: Publish Model Details

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 rub version.

Table 28: Publish Model Details

Field or Icon	Description
Variable Mapping	The table displays the OFSAA variables and datasets used in the creation of the Training Model.
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**.

6.4.4 Viewing Model Details

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

To view model details, 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. Click **Model Pipeline** to display the **Model Pipeline** 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 81: Mouse over the Model
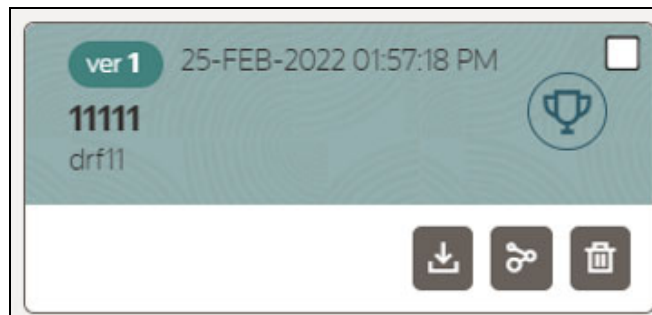
    The icon actions are listed in the following:

    a. Download Model Data. See the Import a Workspace Model Data into a New Model section for more details.

    b. Open in Pipeline Designer. See the Create Paragraphs using Pipeline Designer section for more details.

c. Delete Model. See the Deleting Objectives and Draft Models section for more details.

6.4.5 Understanding Model Governance

After comparing models, you must understand the Model Governance system in OFS CS. 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 82: The Request Drop-down List
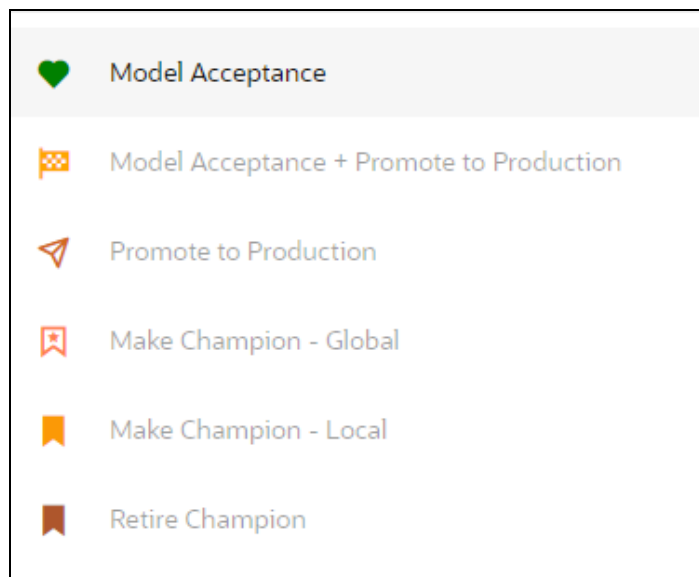
The values in the drop-down list are active based on the type of user (Requester, Reviewer, or Approver) and the phase that the model is in (accepted, promoted to production, global champion, and so on). Let us look at these with a few examples.

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.

6.4.6   Request Model Acceptance

After comparing models, move the selected models to acceptance. Only a user with either Requester or Reviewer roles can request for model acceptance. The model moved to acceptance is then available to users with the Approver role, who can promote to production. See the Understanding 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. Double click 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 83: Model Details Window
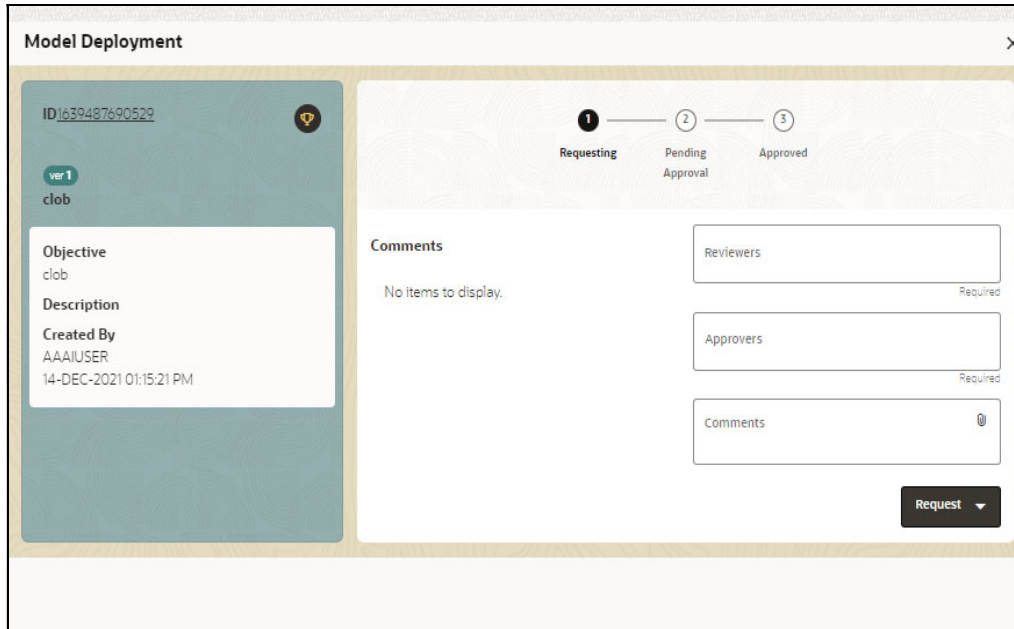
5. Select the reviewer group from the **Reviewers** drop-down list.

6. Select the Approver group from the **Approvers** drop-down list.

7. Enter comments in **Comments** and click 📎 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 84: Model Approval Progress Indicator
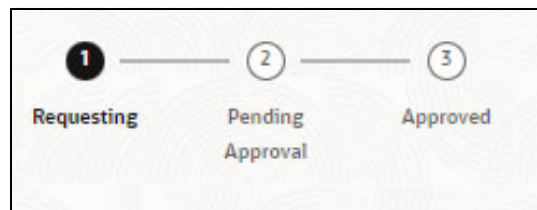
    ▪ Click the type of request from the drop-down list:

    — 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 when signed in, who must either accept the request or reject it.

6.4.7 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, 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 Understanding Model Governance section before you start here.

To review a model, 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. Click **Model Pipeline** to display the **Model Pipeline** 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.

6.4.8 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 Understanding Model Governance section before you start here.

To approve or reject models, 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. Click **Model Pipeline** to display the **Model Pipeline** 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.

6.4.9   Deploying 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, 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. Click **Model Pipeline** to display the **Model Pipeline** window.

   The window displays folders that contain models and model records in a table.

3. Select the model to deploy and click to display the details of the model. If a model is a champion, the icon is displayed.

6.5   Executing 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 Using Scheduler Service section.

6.5.1   Defining a Task

OFS CS supports following three Components during the Task creation:

- Model
- Populate Workspace
- Custom

Figure 85: 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. |
> |---|---|

6.5.2   When Component is Model

The following window shows the Task Parameters for Model Component.

Figure 86: 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

6.5.3 When Component is Populate Workspace

If you select Component as Populate Workspace, then to use the data population, enter the following parameters

- Additional Parameters
- Table Filters
- Global Filters

For more information, see the Populating the Workspace section.

Figure 87: Create Task
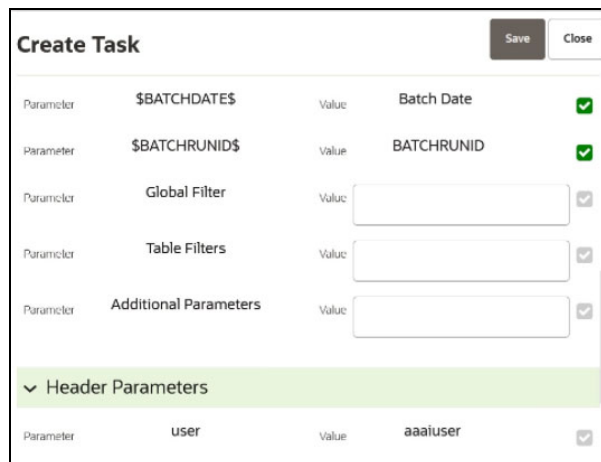
6.6 Import a Workspace Model Data into a New Model

The model data from existing models in OFS CS in .dmp format and the existing model data in OFS CS 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 Creating Draft Models section for instructions on how to proceed further.

To import model data, 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. Click **Model Pipeline** to display the **Model Pipeline** 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 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 88: Model Details - Import Dump
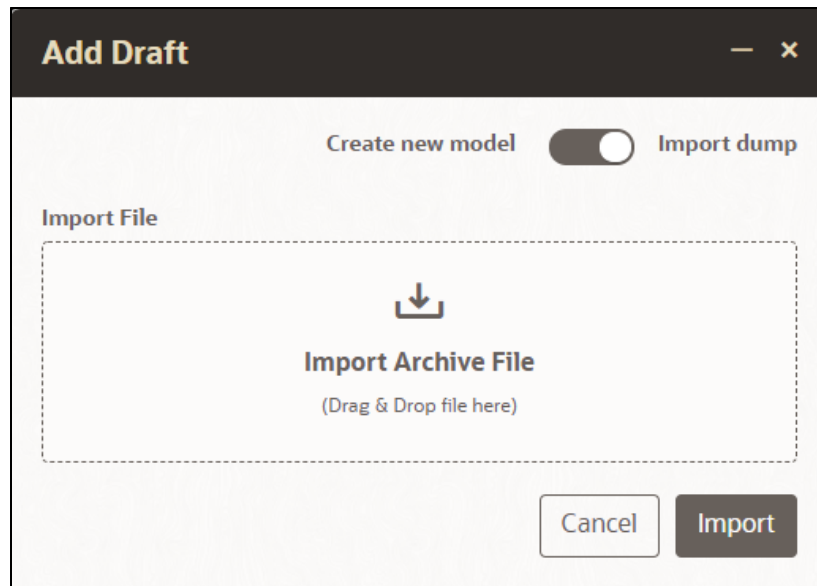
7. Drag the toggle switch to select **Import Dump**.

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.

9. Click **Import**. The following page is displayed.

Figure 89: Import File

10. Enter the details for **Draft Name** and **Description.**

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

12. Click **Import**. A new model is created by importing the model data dump of another model.

13. Click **Delete** in the **File Name** field to delete the imported file, if required.

    The imported file is deleted and then click **Import**. The empty model will be loaded.

> | **NOTE** | Once the imported file is deleted, the user can import the new file and click **Import** to create a new model. |

6.7 Using View Models

The View Models feature launches the OFS Compliance Studio window. You can edit models on this window.

To use View Models, 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. Click **Model Pipeline** to display the **Model Pipeline** window.

The window displays folders that contain models and model records in a table.

3. Click ![icon] next to corresponding Model and select Open in Pipeline Designer.

   See the Create Paragraphs using Pipeline Designer sections for details on how to use the OFS CS Data Studio.

6.8 Editing Models

The editing of models created versions that are different from the previously saved model and the cycle of Understanding Model Governance applies to any edited model. You can edit models from the OFS CS 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.

6.9 Deleting Objectives and Draft Models

To delete Objectives and Models that exist in the Objectives, follow these steps:

1. Click **Launch Workspace** ![icon] next to corresponding Workspace to Launch Workspace to display the **Dashboard** window with application configuration and model creation menu.

2. Click **Model Pipeline** ![icon] to display the **Model Pipeline** window.

   The window displays folders that contain models and model records in a table.

3. Click ![icon] next to corresponding Model and select Delete Draft ![icon] .

4. Select two or more objectives or models from the records.

5. Click ![icon] to display the Delete Draft dialog box.

6. Click **Delete** to delete or click **Cancel** to cancel.

Figure 90: Delete Objects
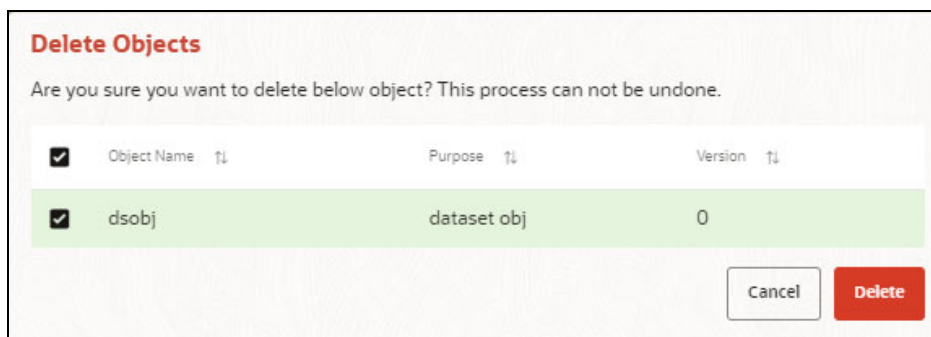
7   Managing Graph Pipelines

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:

- Define the nodes and edges for a graph model, including defining the properties of nodes and edges and their data types

- Define underlying data pipelines to source and standardize the data for nodes and edges from:

  - Oracle database

  - CSV

- 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.

Compliance Studio comes with a pre-configured Financial Crime Graph Model which can be used in models or applications such as Investigation Hub. The pre-configured model pipelines can source data from the Financial Crime Data Model and ICIJ but can be easily augmented with additional sources of data as required.

- The following figure is an example of the preconfigured Financial Crime Graph Model with Nodes and Edges:

Figure 91: Example of Graph Model
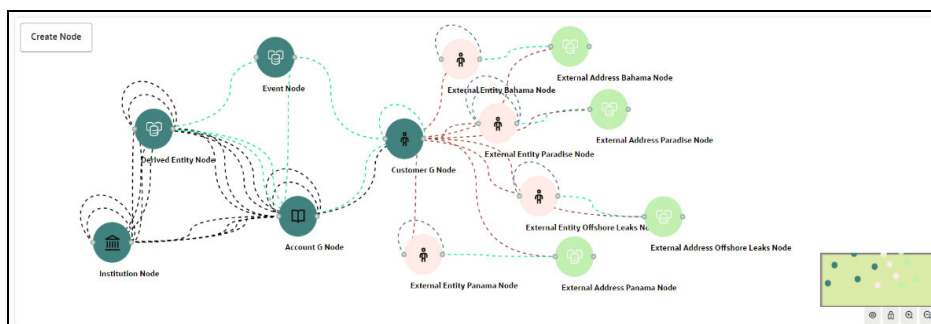
The Graph Model defines the nodes and the relationship between them:

- **Node:** It represents a single entity with its attributes. The entity can be a record from the 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:** Edge is 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:** 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.

  A Transaction edge will have properties like Transaction ID, Date, Amount, From/To account, etc.

See the Adding a Graph Pipeline for creating Graph, Nodes, Edges, and Scheduling.

7.1 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 corresponding Workspace to Launch Workspace to display the **Dashboard** window with application configuration and model creation menu.
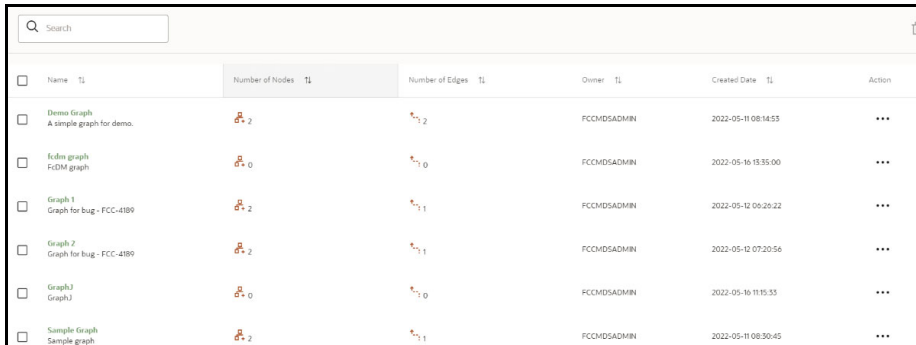
2. Click the **Graph Options** icon .

   This displays the graph pipeline in a table.

   Figure 92: Graph Summary Page

   The Table 29 provides descriptions for the fields and icons on the **Graph Summary** page.

   Table 29: 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.

Table 29: Graph Summary Details
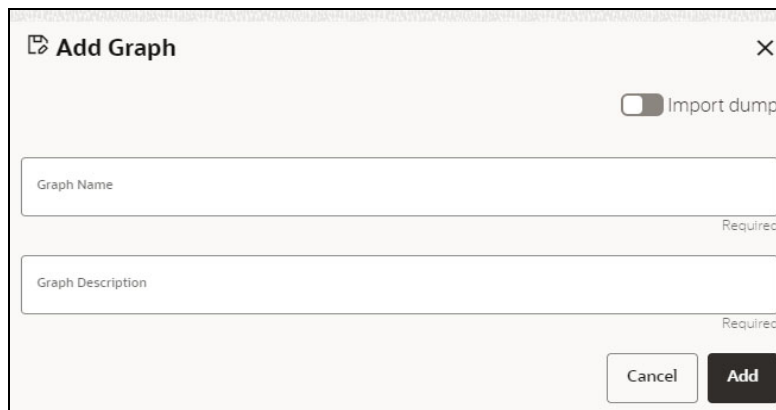
Field or Icon	Description
Number of Edges	The number of edges present in the graph pipeline.
Owner	The owner of the graph pipeline.
Created Date	The date on which the graph pipeline was created.
🗑	Click to delete multiple graph pipelines.
Add	Click to create a new graph pipeline. See Adding a Graph Pipeline section.
...	Click **Action** icon to View/Edit/Delete/Create Population Schedule functions on the selected graph pipeline.

7.1.1   Adding a Graph Pipeline

To add a graph pipeline, follow these steps:

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 93: Add Graph Window

3. In the **Graph Name** field, enter a name for the graph pipeline.

    The name must be unique to a particular workspace.

4. In the **Graph Description** field, enter the description for the graph pipeline and click **Add**.

    Or

5. Drag the toggle switch to select **Import Dump.** The following page is displayed.

Figure 94: Import Dump
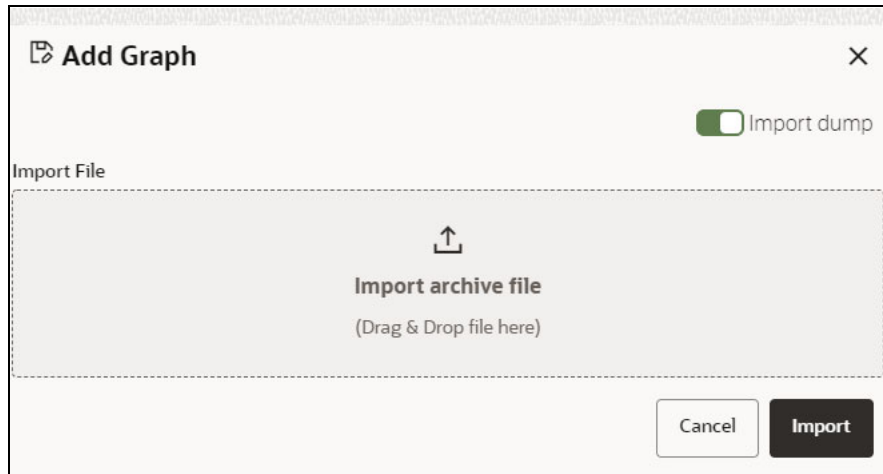
6. Drag and drop the file into the **Import archive file** field or click in the box to open the file selector dialog and select a file.

7. Click **IMPORT**. The following page is displayed.

Figure 95: Graph Edit page
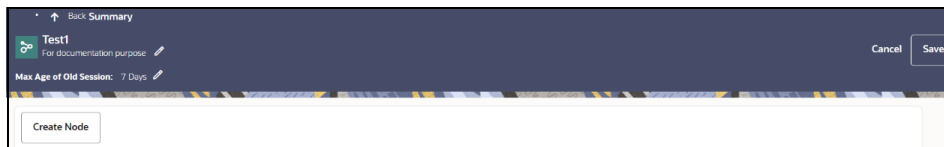
The Max Age of Old session for a graph is 7 days by default. You can modify the description and Max Age of Old session on the above page.

The user session of the Graph Pipeline will get refreshed after the set timeline.

8. You can perform the following:

a. Creating a Node

b. Creating an Edge

c. Toggle drawing which will enable or disable the user from dragging components

d. Zoom In and Zoom Out

7.1.2 Creating a Node

To create a Node, follow these steps:

1. Select the graph pipeline for which you want to create a Node, click on the **Action** button, and select the **Edit** option.

2. On the Graph Summary page, click **Create Node**.

The **Node Details** page is displayed.

Figure 96: Node Details

3. In the **Logical Name** field, enter a name for the Node or click **Setting** icon [icon].

The **Setting** pop-up window is displayed.

Figure 97: Setting

4. 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.

5. 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.

6. The selected Node is displayed on the graph pipeline. To add attributes to the Node, click Add

   icon .

   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 98: Add Node Modeler Attribute(s)

7. In the **Logical Name** field, enter the logical name for the modeler attribute.

   This field is mandatory.

8. 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.

9. Enter the data and click **Ok**.

   The Node attributes are displayed on the Node Details page below.

   Figure 99: Search
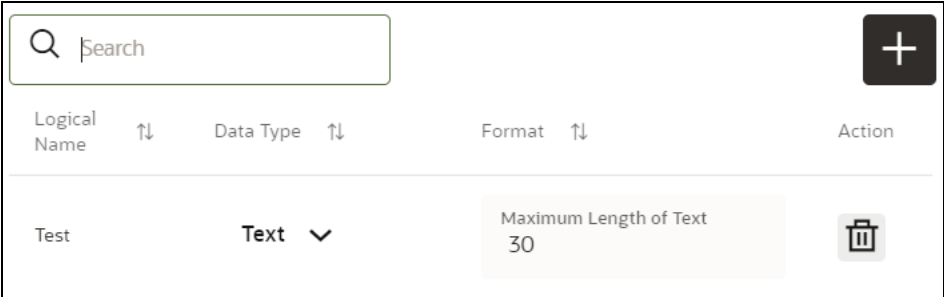
   You can add any number of modeler attributes based on your requirements by clicking on the Add icon.

10. Under the **Advanced Features** group, select the modeler attribute from the Column Identifier drop-down. All the created modeler attributes are displayed under this drop-down.

    Figure 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) group.

Figure 101: Manage Pipeline(s)
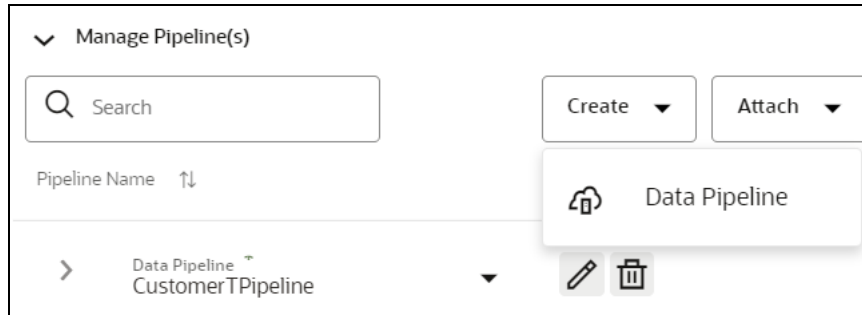
11. Click **Create** drop-down and then select the **Data Pipeline**. The **Create Pipeline** window is displayed.

Figure 102: Create a New Data Pipeline

12. Select the **Create New Data Pipeline** option and then click **Ok**. The Pipeline Designer page is displayed.

- To create a new 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	The Persist of the Data pipeline of the corresponding edge should be defined with the prescript. See the **Prescript Condition** section in the OFS Compliance Studio Administration and Configuration Guide.

Figure 103: Create an Index
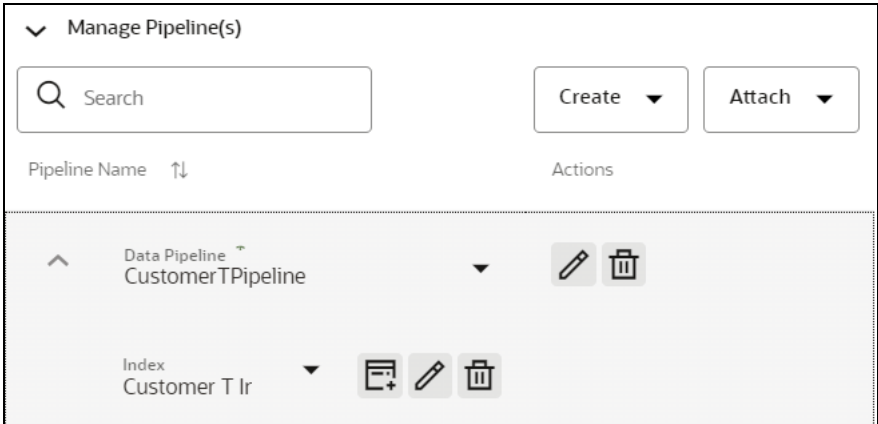
- To define an index for the data pipeline, see the Defining an Index section.

13. By clicking on the **Attach** drop-down, you can add any number of data pipelines to the Node-based on your requirements.

14. Click **Ok**.

The Node is created with pipeline Facebook and an account attached to it.

Figure 104: Graph with Node and Edge
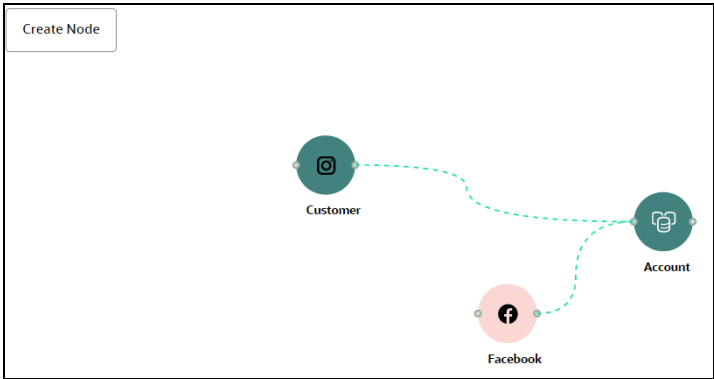
7.1.3 Creating an Edge

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 105: Graph Page
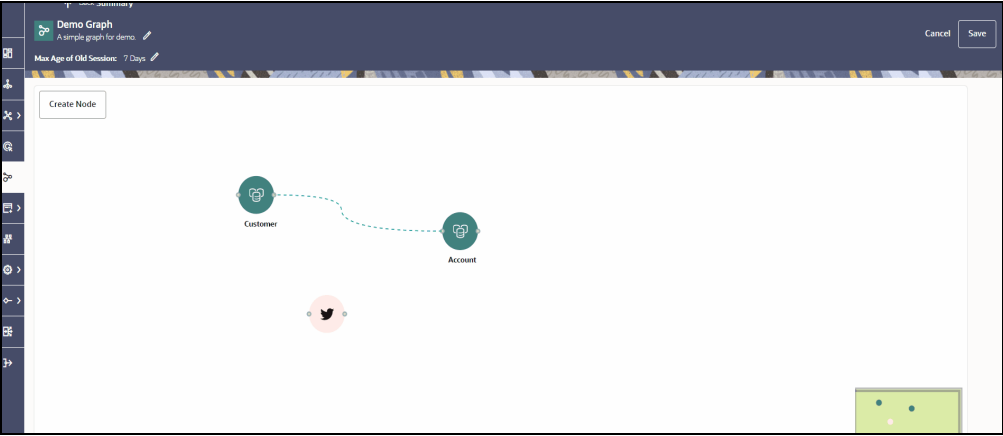
To create an edge, perform these steps:

1. Hover on the Node to create a relationship between them.

   The **Edge Details** screen is displayed.

Figure 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 107: Setting
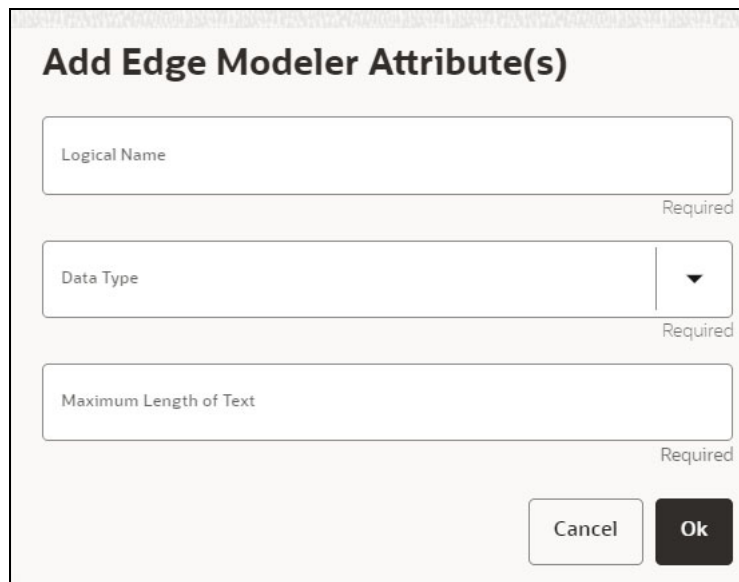
3. Select the required edge from the options displayed and choose the required color for the edge using the **Select Color** option.

4. Click **Ok**.

   The selected edge is displayed on the graph pipeline

5. To add attributes to the edge, click **Add** icon ✚ .

   The **Add Edge Modeler Attribute(s)** screen is displayed.

   Figure 108: Add Edge Modeler Attribute(s)

   - In the **Logical Name** field, enter the logical name for the edge attribute.

     This field is mandatory.

6. 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.

7. Enter the data and click **Ok**.

The edge attribute is displayed on the Edge Details page.

Figure 109: Search
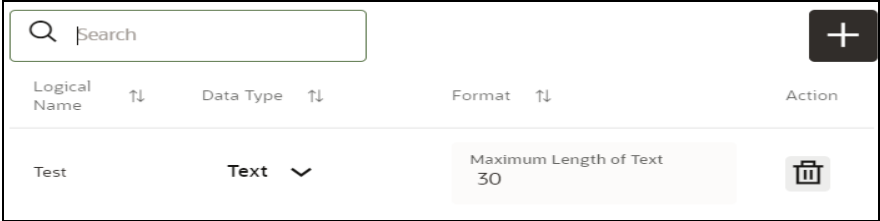
By clicking on the **Add** icon, you can add any number of edge attributes based on your requirements.

Or

Figure 110: Edge details
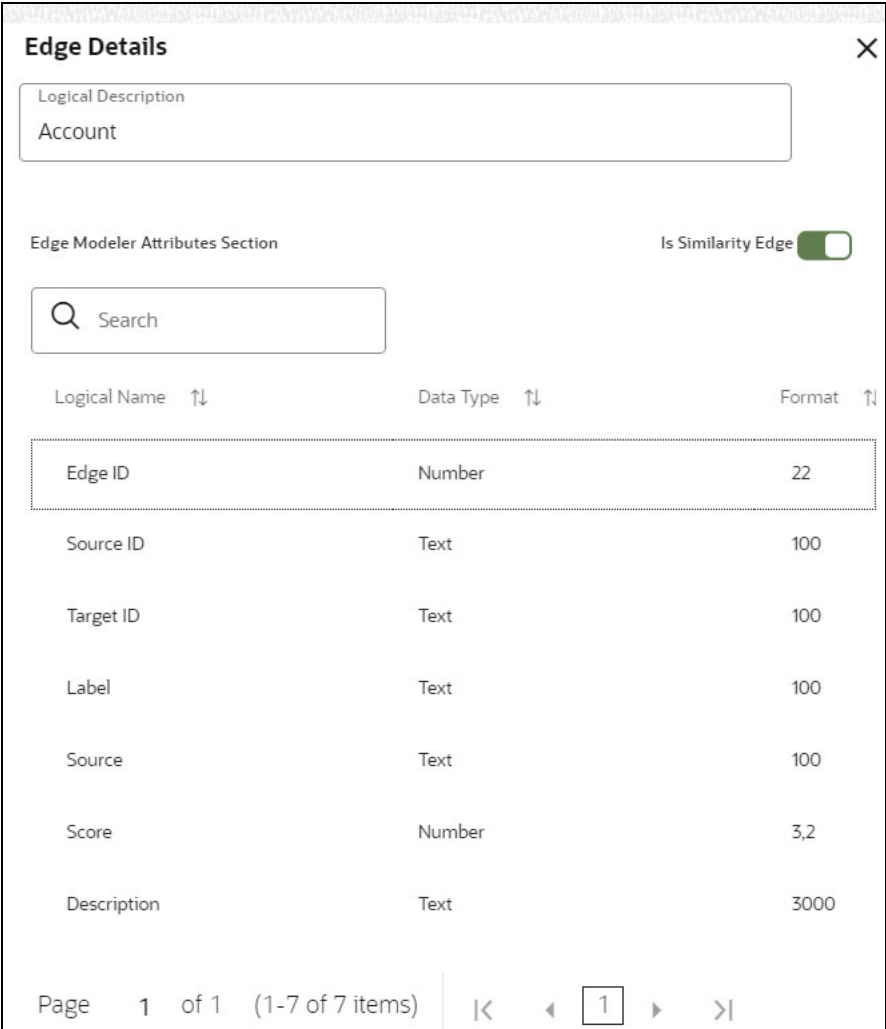
8.   Drag the toggle button to **Is Similarity Edge** in the Edge Modeler Attributes Section to add the pre-defined set of attributes available in the edge.

Similarity edges in 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 enable 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 edge latest 7 days of data will be maintained in this edge.

10. Enter the retention count and select the retention duration from the drop-down.

The available retention durations are Days, Weeks, and Months.

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 111: Pluggable Attribute
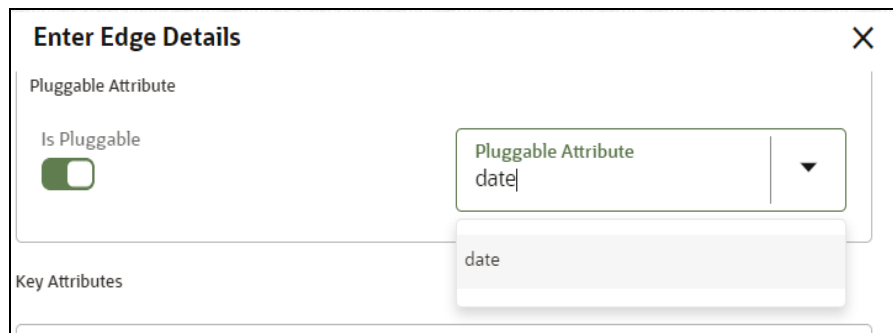
12. Under the **Key Attributes** field, select the Column, Start, and End identifier from the respective Identifier drop-down. Without these attributes, an edge cannot be created.

All the created edge attributes are displayed under this drop-down.

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 112: Manage Pipeline(s) for Edge
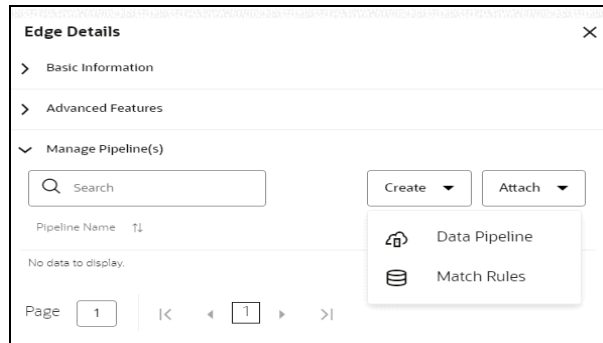
14. Click **Create** drop-down and then select the **Data Pipeline**. The **Create Pipeline** window is displayed.

Figure 113: Create a New Data Pipeline

15. Select the **Create New Data Pipeline** option and then click **Ok**. The Pipeline Designer page is displayed.

   ▪ To create a new 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	The Persist of the Data pipeline of the corresponding edge should be defined with the prescript. See the **Prescript Condition** section in the OFS Compliance Studio Administration and Configuration Guide.

16. Click **Create** drop-down and then select the **Match Rules** from the Manage Pipeline(s) group. The **Create Pipeline** window is displayed.

Figure 114: Create a New Match Rule

17. Select the **Create New Match Rule** option and then click **Ok**. The **Create Match Ruleset** window is displayed.

   ■ To create a new Match Rule, see the Creating Ruleset (Match/Merge) 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. |

   The selected data pipeline/match rule is displayed.

Figure 115: Manage Pipeline(s) for Edge

   By clicking on the **Attach** drop-down, you can add any number of data pipelines/match rules to the Node-based on your requirements.

18. Click **Ok**.

   An Edge is created between the Customer, Facebook, and Account.

Figure 116: Graph page
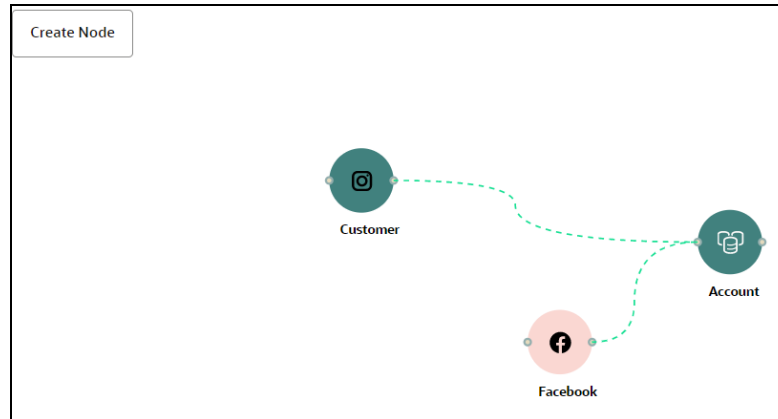
7.1.4 Zoom In and Zoom Out

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.

- Click ⊕ to zoom in on the graph pipeline.

- Click ⊖ to zoom out the graph pipeline.

- Click ◉ to view the data pipelines.

7.1.5 Editing 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. In the **Graph Description** field, enter the description for the graph pipeline and click **Save**.

   The following page is displayed.

   Figure 117: Graph page
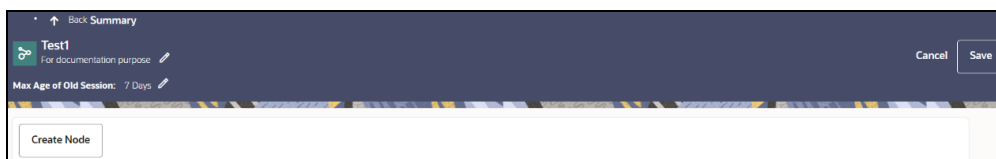
By default, the Max Age of Old session for a graph is 7 days. You can modify the description and Max Age of Old session on the above page.

7.1.6   Deleting a Graph Pipeline

To delete a graph pipeline, perform these steps.

1. On the **Graph Summary** page, select the graph pipeline you want to delete, click the **Action** icon, and select **Delete**.

    Figure 118: Graph Summary Page
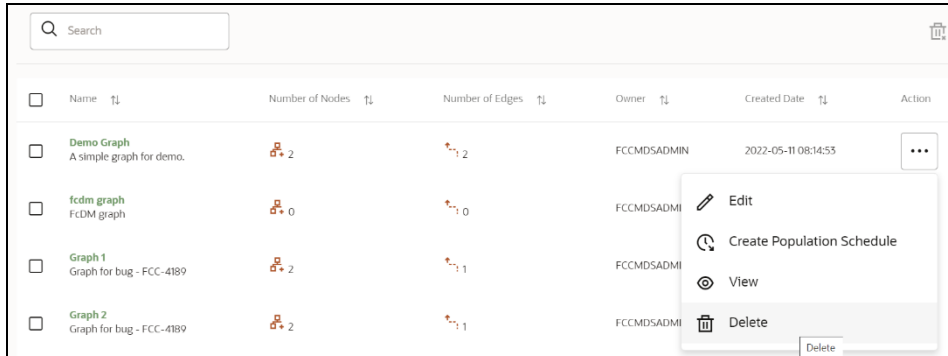
    To delete multiple graph pipelines, select the graph pipelines which you want to delete and click

     .

7.1.7   Editing 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 119: Edit Node
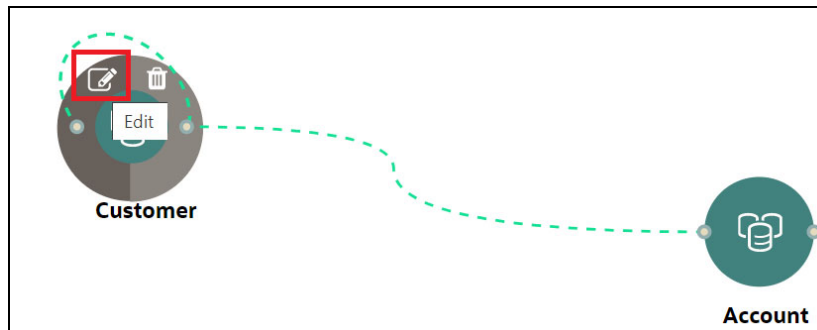
3. Perform the steps from 3 to 12 in the Creating a Node section.

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:
	1. Navigate to **Edge Details** > **Manage Pipeline(s)** > **Attach**. Attach the new match rule, which is not attached to the node details.
	2. Click **Edit** icon and delete the attribute in the Ruleset Details, which is removed from the Node Details page.
	3. Click **Save**. The attribute is removed.
	4. Click **Delete** in Manage Pipeline(s) of the Edge Details page to remove the attached match rule.
	5. After saving in the Edge Details page, navigate to the Node Details page and delete the attribute again.
	6. Click **OK**. The attribute is removed from the node.

7.1.8   Deleting a Node

To delete 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 **Delete** icon.

Figure 120: Delete Node
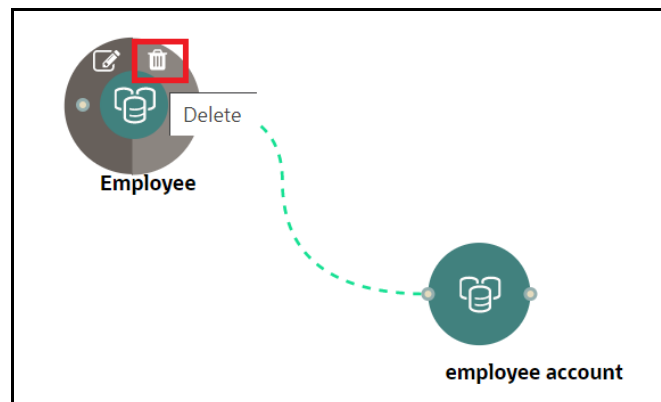
A confirmation dialog is displayed.

3. Click **Delete**.

The Node is deleted.

7.1.9   Creating a Population Schedule

To create a population schedule, perform these steps.

1. In the **Graph Summary** page, select the graph pipeline for which you want to create a population schedule and click **Action** icon.

2. Select **Create Population Schedule** option.

    The following screen is displayed.

    Figure 121: Graph page
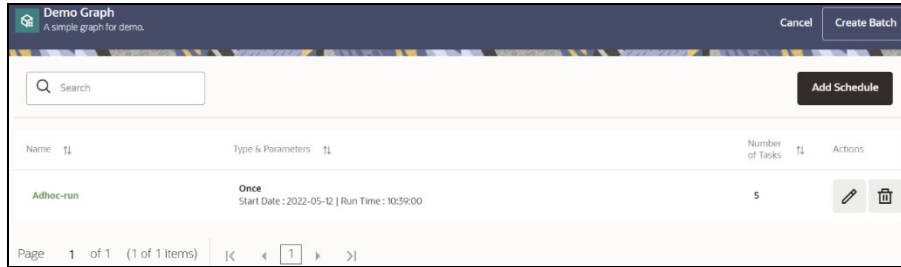
3. Click **Add Schedule**.

    The **Add Schedule** pop-up window is displayed.

    Figure 122: Add Schedule

4. In the **Schedule Name**, enter the desired name.

5. 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 cron schedule type, click the **help** icon.

6. Click inside the **Select Task** field and select the tasks. You can select multiple tasks.

7. Click **Next** and add the values for the selected task, and the values are mandatory fields.

Figure 123: Add Optional Parameters
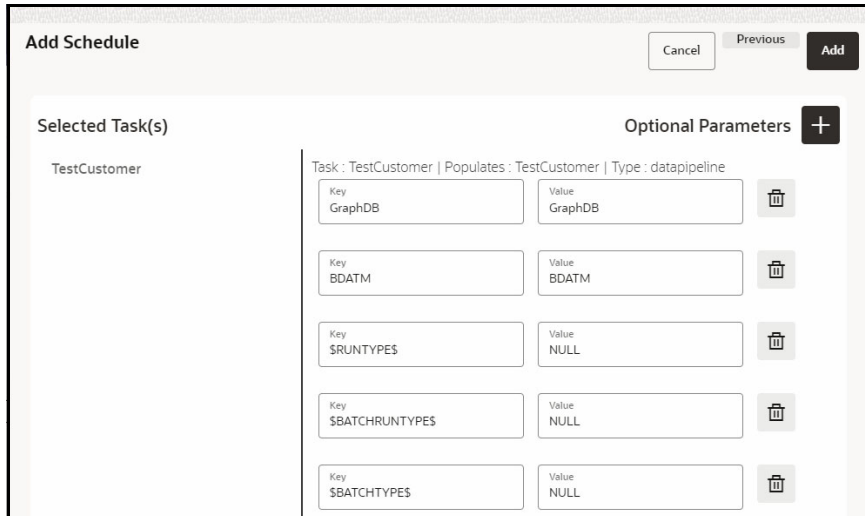
8. Click **+** icon to add optional parameters for the selected task.

   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 Table 30:

Table 30: 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
$batchRunType$	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.	

9. 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 Ruleset (Match/Merge) section.

Figure 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 Table 31:

Table 31: 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
basedOnPipelineId	Match rules can be executed based on pipeline Ids or ruleset Ids. The value should be true while running from the Graph.	true
runType	Indicates matching job that can be executed as part of graph or Entity Resolution. It should be set as Graph.	graph

10. Enable the **Refresh Graph** option and click **Add** icon to refresh the Graph to add optional parameters.

> **NOTE** The Key and Value parameters are not required.

Figure 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.

11. Click **Create Batch**.

    It will save the schedule and navigates you to the Graph Summary page.

    OR

    Click **Cancel**.

    It will navigate you to the Graph Summary page without saving the changes.

    Figure 126: Graph page
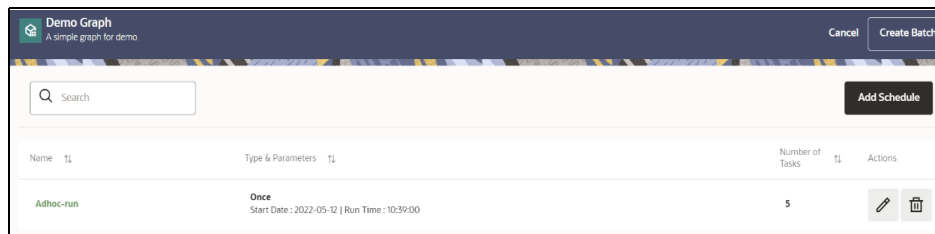
12. Execute the Batch.

8   Using 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 program execution.

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 processes 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.

Monitor: The Scheduler Service enables you to monitor your executions by using a web-browser. It provides real-time feedback on the status of the current encoding job and lists the jobs pending in the batch. You can Cancel or Restart the service when required.

8.1   Scheduler Service Dashboard

On accessing Scheduler Service using your browser, the **Scheduler Service Dashboard** window is displayed.

Figure 127: Scheduler Service Dashboard
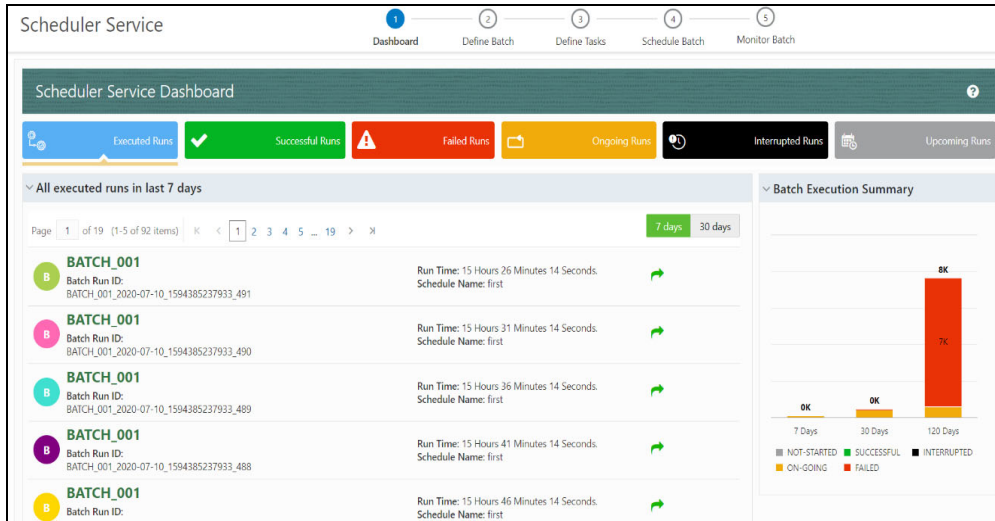
8.2   Header Details

The Header Details in the **Scheduler Service** window is displayed as shown and it has the following options.:

- 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: The User details are displayed, you can view the help, set you profile preferences.

Figure 128: Header Details
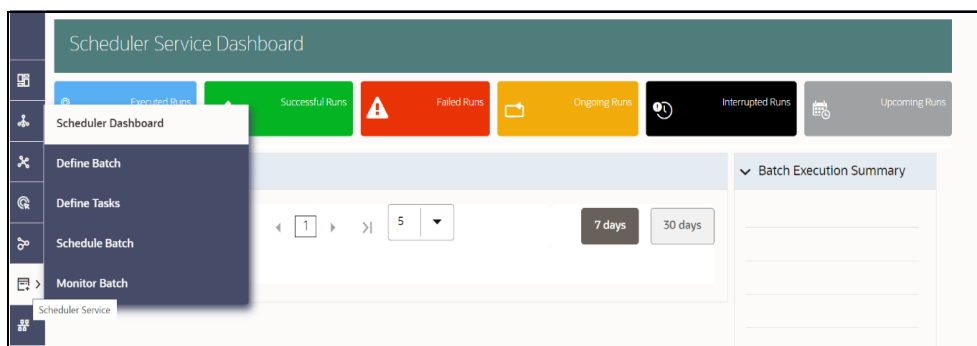
On the left side of the window, click the **Scheduler Service** icon to see the different option in the Scheduler Service.

8.3   Scheduler Service Dashboard

To access the Scheduler, In the OFSAA Applications window, select the [icon] Menu Navigation icon, and then select Batch Administration, and then select Scheduler to open the *Dashboard* is displayed.

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. You can click the [icon] icon corresponding to a Batch to monitor it.

- The **Batch execution Summary** pane displays te 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 the batches.

Figure 129: Batch Execution Summary
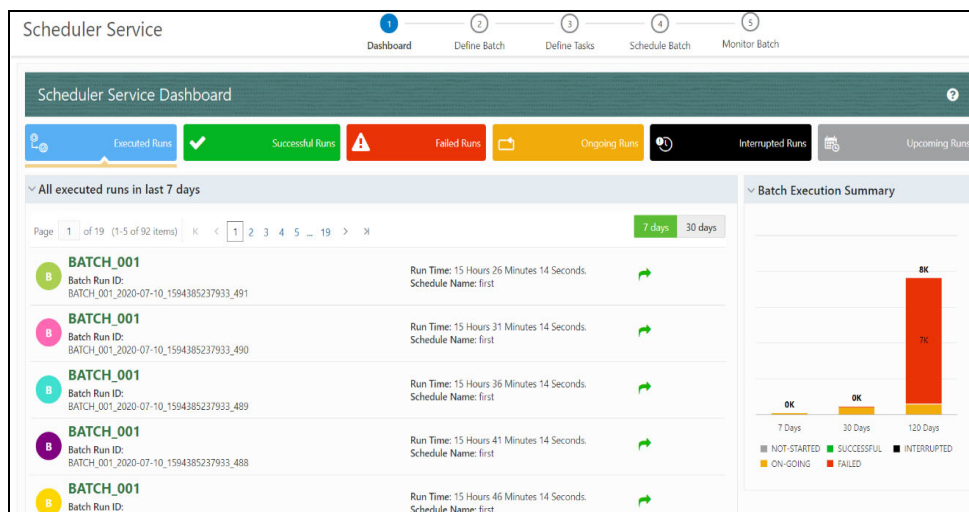
8.4   Define Batch

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 [icon] icon corresponding to the batch to proceed to create or edit tasks.

8.4.1   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.

8.4.1.1   Create a Batch

You can create a new batch in the Define Batch Window and schedule and monitor the batch that you created.

To create a new Batch, perform the following steps:

1. In the Define Batch window, click [+].

    The **Create a New Batch** window is displayed.

2. Specify the details as tabulated in the Create a New Batch window as described in the Table 32.

Table 32: Fields in the Create a New Batch window and their Description

Field	Description
Batch Details	
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	● 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.

3. From the Batch Parameters pane, click ![icon] to add a new Batch Parameter. By default, $FICMISDATE$ and $BATCHRUNID$ are added as Batch parameters.

> **TIP**   Enclose the parameter value for a run time with $ symbol. For example, $paramName$.

     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 ![icon] corresponding to the parameter.

4. Click **Save**.

The new Batch is created and displayed in the **Define Batch** window.

8.4.1.2 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 ![icon] 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.

8.4.1.3 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 ![icon] 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.

8.4.1.4 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 ✖ 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.

8.4.2 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.

8.4.2.1 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, perform the following steps:

1. In the Define Batch window, click ➕.

   The **Create a New Batch** window is displayed.

2. Select Batch Group option.

3. Specify the following fields:

   - Name
   - Description
   - Add Batches

   > **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.

4. Click **Save**.

   The new Batch Group is created and displayed in the **Define Batch** window.

8.4.2.2 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 ✏ 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.

8.4.2.3   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 📋 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.

8.4.2.4   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 ✖ corresponding to the Batch Group you want to delete.

3. Click OK in the confirmation dialog to confirm deletion.

8.5   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

8.5.1 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. 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

8.5.1.1 Add a Task

Adding a new task option allows you to add new tasks to a selected Batch definition.

To add new task, perform the following steps:

1. Click **Define Tasks** from the Header panel.

   The **Define Task** window is displayed.

2. Select the Batch 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.

4. Enter the details in the Table 33:

Table 33: Fields in the Create a New Task window and their Description

Field	Description
Task Details	
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.
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.

5. From the *Task Parameters* pane, click Add [image] 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 ([image])

   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 [image] corresponding to the parameter.

6. Click **Save**.

8.5.1.2 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 [image] 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**.

8.5.1.3 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 Menu [image] 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 [>]. You can press Ctrl key for multiple selections.

b. To select all the listed Tasks, click [»].

c. To remove a Task, select the task from the *Selected Tasks* pane and click [<].

d. To remove all the selected Tasks, click [«].

2. Click **Save** to update Task Precedence.

3. Click **Preview** to view the Precedence information.

8.5.1.4 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.

8.5.2 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. You can perform the following operation for the batch based on the task.

- Define Task Precedence

8.5.2.1 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 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 batch that you want to execute before the current task, from the **Available Tasks** pane and click [>]. You can press Ctrl key for multiple selections.

b. To select all the listed batches, click [»].

c. To remove a batch, select the task from the **Selected Tasks** pane and click [<].

d. To remove all the selected batches, click [«].

2. Click **Save** to update Task Precedence in the batches.

3. Click **Preview** to view the Precedence information.

8.6   Schedule Batch

The *Schedule Batch* window facilitates you to run, schedule, re-start, re-run the batches in the Scheduler Service. After you upload the data in the required format into the 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.

The **Schedule Batch** window allows you to perform the following operations for your batch and batch group:

- Batch
- Batch Group

8.6.1   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:

- Execute a Batch
- Schedule a Batch
- Re-start a Batch
- Re-run a Batch
- Edit Dynamic Parameters
- Task Definitions of a Batch

8.6.1.1 Execute a Batch

The Execute batch option allows you to run a batch instantaneously. To execute a batch, perform the following steps:

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.

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 or Include Tasks section.

6. If you want to hold/release some tasks, click **Hold Tasks**.

   For more information, see Hold or 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.

8.6.1.2 Schedule a Batch

You can schedule a Batch to run just for Schedule Once, Schedule Daily, Schedule Weekly, Schedule Monthly or Schedule Cron Expression for scheduling the batches. You can also have a user defined schedule to schedule and run a batch.

8.6.1.2.1 Schedule Once

To schedule a Batch to run once, perform the following steps:

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 drop down menu.

4. Enter a Schedule Name.

5. Click ▦ and select the date on which you want to run the Batch.

6. Click 🕐 and select the time at which you want to run the Batch.

7. Click **Schedule**.

8.6.1.2.2 Schedule Daily

To schedule a Batch to run daily, perform the following steps:

1. In the **Schedule Batch** window, click **Daily**.

2. Select the Batch or Batch Name you want to schedule daily from the Select drop down menu.

3. Enter a **Schedule Name**.

4. Click [icon] and select the start date from which you want to run the Batch.

5. Click [icon] and select the end date till which you want to run the Batch.

6. Click [icon] and select the time at which you want to run the Batch daily.

7. Click **Schedule**.

8.6.1.2.3   Schedule Weekly

To schedule a Batch to run weekly, perform the following steps:

1. In the **Schedule Batch** window, click **Weekly**.

2. Select the Batch or Batch Name you want to schedule weekly from the Select drop down menu.

3. Enter a Schedule Name.

4. Click [icon] and select the start date from which you want to run the Batch.

5. Click [icon] and select the end date till which you want to run the Batch.

6. Click [icon] and select the time at which you want to run the Batch.

7. Select the days on a week you want to run the Batch from the Select Days of the Week multi-select drop down menu.

8. Click **Schedule**.

8.6.1.2.4   Schedule Monthly

To schedule a Batch to run monthly, perform the following steps:

1. In the **Schedule Batch** window, click **Monthly**.

2. Select the Batch or Batch Name you want to schedule monthly from the Select drop down menu.

3. Enter a Schedule Name.

4. Click [icon] and select the start date from which you want to run the Batch.

5. Click [icon] and select the end date till which you want to run the Batch.

6. Click [icon] and select the time at which you want to run the Batch.

7. Select the months in a year you want to run the Batch from the Select Months of the Year multi-select drop down menu.

8. Select the day of the month that you want to run the Batch from the Select Day of the Month drop down menu.

9. Click **Schedule**.

8.6.1.2.5   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 icon next to the Cron Expression field.

To schedule a Batch based on Cron Expression, perform the following steps

1. In the **Schedule Batch** window, click **Cron Expression.**

2. Select the Batch or Batch Name you want to schedule from the Select drop down menu.

3. Enter a **Schedule Name**.

4. Enter the Cron Expression for your schedule. For more information about the Cron Expression, click icon next to the Cron Expression field.

5. Click **Schedule**.

8.6.1.3   Re-start a Batch

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, perform the following steps:

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 or Batch Name you want to re-start from the Select Name drop down menu.

4. Select the Batch Run ID.

5. Click **Re-start**.

8.6.1.4   Re-run a Batch

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, perform the following steps:

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 drop down menu.

4. Select the Batch Run ID.

5. Click **Re-run**.

8.6.1.5   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, 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.

8.6.1.6   Task Definitions of a Batch

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 from the **Schedule Batch** window, you can:

*   Exclude a task or include the excluded task

*   Hold a task or release the held task

8.6.1.6.1   Exclude or Include Tasks

You can exclude tasks or include the excluded tasks during Batch 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 ⟩ . You can press Ctrl key for multiple selections.

    b. To exclude all tasks, click ⟫ .

3. To include the excluded tasks:

    a. Select the required task from the **Excluded Tasks** list and click ⟨ . You can press Ctrl key for multiple selections.

    b. To include all excluded tasks, click ⟪ .

4. Click **Save**.

8.6.1.6.2 Hold or Release Tasks

You can hold tasks or release the held tasks during Batch 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/excluded for Batch execution.

To hold/release tasks, perform the following steps:

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 [>]. You can press Ctrl key for multiple selections.

   b. To hold all tasks, click [>>].

3. To release held tasks:

   a. Select the required task from the **Held Tasks** list and click [<]. You can press Ctrl key for multiple selections.

   b. To release all held tasks, click [<<].

4. Click **Save**.

8.6.2 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 group:

- Execute a Batch Group
- Schedule a Batch Group
- Re-start a Batch Group
- Re-run a Batch Group
- Edit Dynamic Parameters
- Task Definitions of a Batch Group

8.6.2.1 Execute a Batch Group

The Execute batch option allows you to run a batch instantaneously. To execute a batch group, perform the following steps:

1. Click **Schedule Batch** from the Header panel.

   The **Schedule** window is displayed.

2. Select the Batch Group from the Select Name drop down menu.

   For example, AMLDataLoad.

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 or Include Tasks section.

6. If you want to hold/release some tasks, click **Hold Tasks**.

For more information, see Hold or 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.

8.6.2.2   Schedule a Batch Group

You can schedule a Batch group to run just for Schedule Once, Schedule Daily, Schedule Weekly, Schedule Monthly or Schedule Cron Expression for scheduling the batches. You can also have a user defined schedule to schedule and run a batch.

8.6.2.2.1   Schedule Once

To schedule a Batch to run once, perform the following steps:

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 Group you want to schedule for once from the Select drop down menu.

4. Enter a Schedule Name.

5. Click ▦ and select the date on which you want to run the Batch.

6. Click 🕒 and select the time at which you want to run the Batch.

7. Click **Schedule**.

8.6.2.2.2   Schedule Daily

To schedule a Batch to run daily, perform the following steps:

1. In the **Schedule Batch** window, click **Daily**.

2. Select the Batch Group you want to schedule daily from the Select drop down menu.

3. Enter a **Schedule Name**.

4. Click ▦ and select the start date from which you want to run the batch group.

5. Click ▦ and select the end date till which you want to run the batch group.

6. Click ⏱ and select the time at which you want to run the batch group daily.

7. Click **Schedule**.

8.6.2.2.3 Schedule Weekly

To schedule a Batch to run weekly, perform the following steps:

1. In the **Schedule Batch** window, click **Weekly**.

2. Select the Batch Group you want to schedule weekly from the Select drop down menu.

3. Enter a Schedule Name.

4. Click 📅 and select the start date from which you want to run the batch group.

5. Click 📅 and select the end date till which you want to run the batch group.

6. Click ⏱ and select the time at which you want to run the batch group.

7. Select the days in a week you want to run the batch group from the Select Days of the Week multi-select drop down menu.

8. Click **Schedule**.

8.6.2.2.4 Schedule Monthly

To schedule a Batch to run monthly, perform the following steps:

1. In the **Schedule Batch** window, click **Monthly**.

2. Select the Batch Group you want to schedule monthly from the Select drop down menu.

3. Enter a Schedule Name.

4. Click 📅 and select the start date from which you want to run the batch group.

5. Click 📅 and select the end date till which you want to run the batch group.

6. Click ⏱ and select the time at which you want to run the batch group.

7. Select the months in a year you want to run the batch group from the Select Months of the Year multi-select drop down menu.

8. Select the day of the month that you want to run the batch group from the Select Day of the Month drop down menu.

9. Click **Schedule**.

8.6.2.2.5 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 🛈 icon next to the Cron Expression field.

To schedule a Batch based on Cron Expression, perform the following steps

1. In the **Schedule Batch** window, click **Cron Expression.**

2. Select the Batch Group you want to schedule from the Select drop down menu.

3. Enter a **Schedule Name**.

4. Enter the Cron Expression for your schedule.

   For more information about the Cron Expression, click ![info icon] icon next to the Cron Expression field.

5. Click **Schedule**.

8.6.2.2.6   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 Schedule Once, Schedule Daily, Schedule Weekly, Schedule Monthly or Schedule 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 Schedule Once, Schedule Daily, Schedule Weekly, Schedule Monthly or Schedule 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 ![plus icon] 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:

Schedule Batch

1. In the **Schedule Batch** window, you can select either Schedule Once, Schedule Daily, Schedule Weekly, Schedule Monthly or Schedule 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 Schedule Once, Schedule Daily, Schedule Weekly, Schedule Monthly or Schedule 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 ➕ 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 Schedule Once, Schedule Daily, Schedule Weekly, Schedule Monthly or Schedule 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 Schedule Once, Schedule Daily, Schedule Weekly, Schedule Monthly or Schedule 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 ➕ 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.

8.6.3   Re-start a 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, perform the following steps:

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**.

8.6.4   Re-run a 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, perform the following steps:

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 Group you want to re-run from the Select Name drop down menu.

4. Select the Batch Run ID.

5. Click **Re-run**.

8.6.5   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.

8.6.6 Task Definitions of a 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 or Include Tasks
- Hold or Release Tasks

8.6.6.1 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 [>]. You can press Ctrl key for multiple selections.

    b. To exclude all tasks, click [>>].

3. To include the excluded tasks:

    a. Select the required task from the **Excluded Tasks** list and click [<]. You can press Ctrl key for multiple selections.

    b. To include all excluded tasks, click [<<].

4. Click **Save**.

8.6.6.2   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, perform the following steps:

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 ⟩ . You can press Ctrl key for multiple selections.

   b. To hold all tasks, click ⟫ .

3. To release held tasks:

   a. Select the required task from the *Held Tasks* list and click ⟨ . You can press Ctrl key for multiple selections.

   b. To release all held tasks, click ⟪ .

4. Click **Save**.

8.7   Monitor Batch

The Monitor Batch enables you to view the status of executed Batch along with the tasks details. You can track the issues if any, on regular intervals and ensure smoother Batch execution. A visual representation as well as tabular view of the status of each Tasks in the Batch is available.

The **Monitor Batch** window allows you to perform monitoring for the following operations:

- Batch
- Batch Group

8.7.1   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. To monitor a batch, perform the following steps:

1. Click **Monitor Batch** from the Header panel.

   The **Monitor** window is displayed.

2. Select the **Batch 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.

  — **Task 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.

  — **Task 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. - If you have two batch run IDs for the first execution, ignore the execution status of run IDs and retrigger the batch one more time.

5. To restart the Batch Group, select Restart .

6. To rerun the Batch Group, select Rerun .

7. To interrupt the Batch Group, select Interrupt .

8. To view the log information about the batch, click the Log icon in the List View tab. The Log Viewer pop-up is displayed.

9. In the Log Viewer pop-up, the log information is displayed. You can click Download icon to download the log or click the Close icon to close the log information.

8.7.2   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. To monitor a batch group, perform the following steps:

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:

      i.   Batch Status: Displays the batch status, the different batch status are NOT-STARTED, ON-GOING, SUCCESSFUL, FAILED, INTERRUPTED, EXCLUDED, HELD, and UNDEFINED.

      ii.   Batch Start Time: Displays the batch start time details.

      iii. Batch End Time: Displays the batch end time details.

      iv.   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:

      i.   Batch Status: Displays the batch status, the different batch status are NOT-STARTED, ON-GOING, SUCCESSFUL, FAILED, INTERRUPTED, EXCLUDED, HELD, and UNDEFINED.

      ii.   Batch Start Time: Displays the batch start time details.

      iii. Batch End Time: Displays the batch end time details.

      iv.   Batch Details: Mouseover the task to display its status and details.

      v.   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.

9 Using 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, for example 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, 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. Click **Audit Trail** to view the time of the various actions performed on the model in the **Audit Trail** window.

   Figure 130: Audit Trail Window
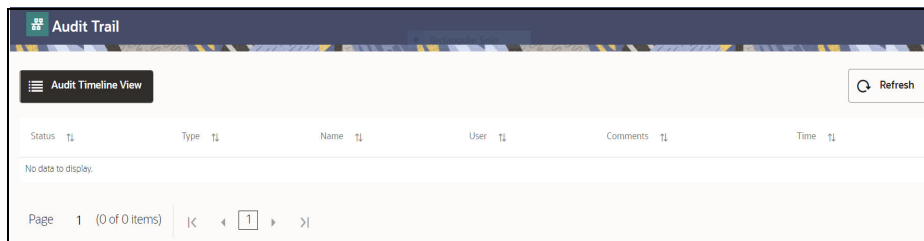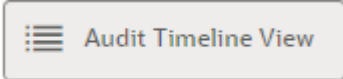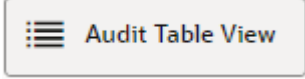
3. Click for Timeline View. Click to switch to the regular view. Click to refresh the Audit Trail window.

   Figure 131: Audit Trail Window Timeline View with Horizontal Time Axis
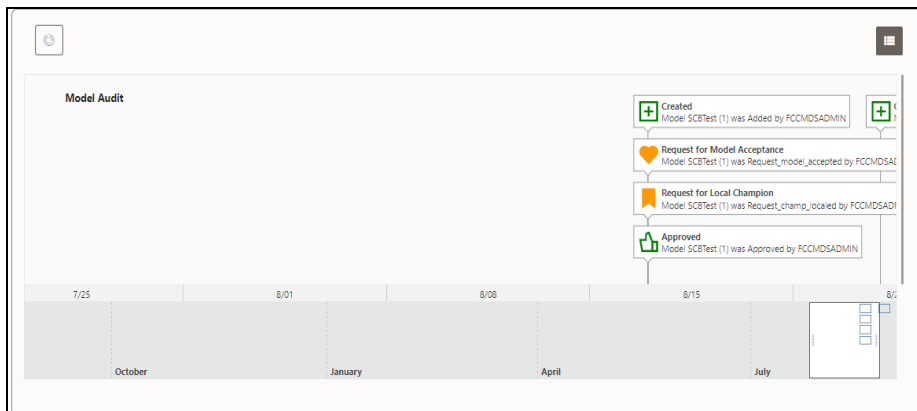
10 Using Model Actions

The Model Actions window allows you to view the list of models associated with 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 models champion locally or globally, and so on.
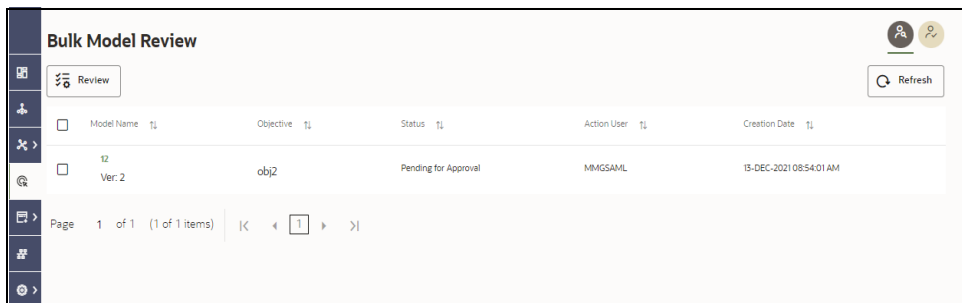
1. Click **Launch Workspace** next to corresponding Workspace to Launch Workspace to display the **Dashboard** window with application configuration and model creation menu.

2. Click **Model Actions** to view the list Models pending for action. For a Reviewer, sample Model Review window is showing below:

Figure 132: Bulk Model Review
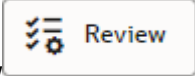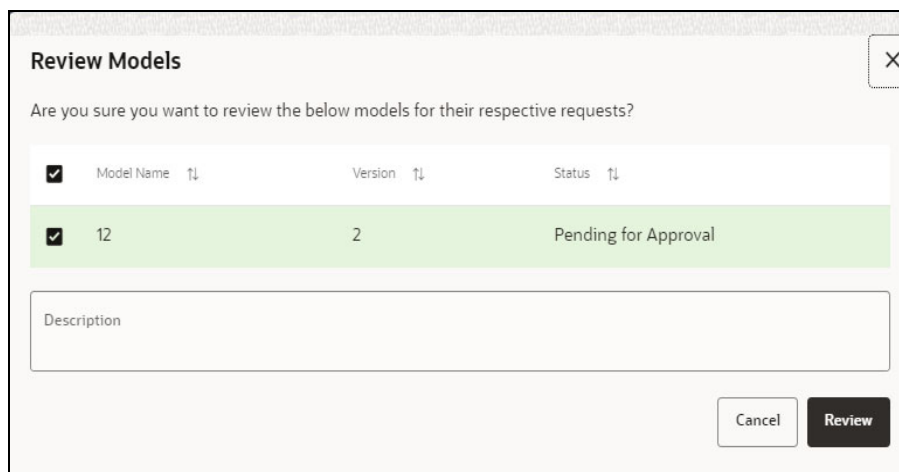
3. To review the model, follow these steps:

   a. Select the Model using the corresponding check box and click Review .

   b. Enter the comments in Comments field of the Review Models window and click **Review**.

Figure 133: Review Models window

4. For an Approver, sample Model Approve window is showing below:

Figure 134: Bulk Model Approval
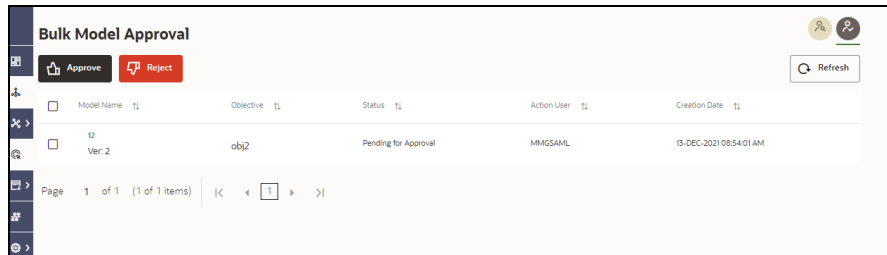
a. To approve or reject the model, follow these steps: Select the Model using the corresponding check box and click Approve or Reject .
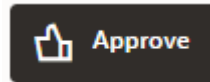
b. Enter the comments in Comments field of the Approve Models window or Reject Models window and click approve or Reject.

Figure 135: Approve Models
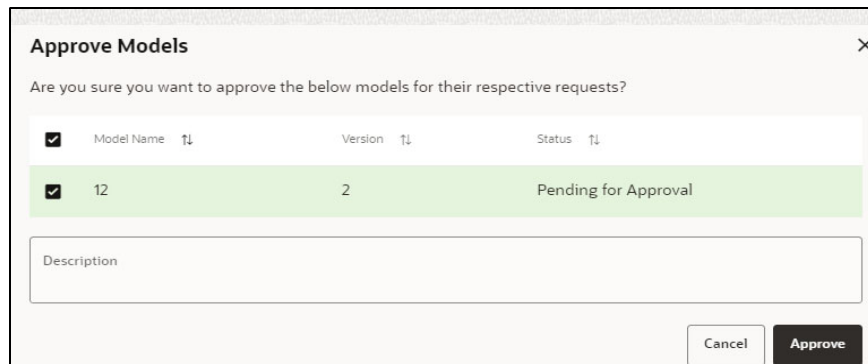
Figure 136: Approve Models window
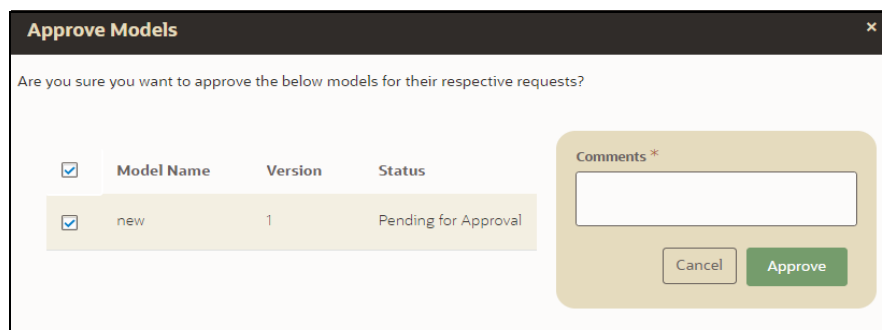
11   Using the Compliance Studio Workspace

The Compliance Studio Workspace provides core capabilities that are not model or sandbox-specific.

Topics:

- Compliance Studio Global Graph
- Entity Resolution
- Creating Rulesets
- Using Manual Decisioning
- Using Merge and Split Global Entities
- Using Provided Notebooks
- Create Event API
- Publishing your Notebooks from Notebook Server
- Mapping Graph Datasource
- Using Graph Definition
- Managing Data Pipeline

11.1   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 Hub. 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 to input other paragraphs in the notebook. The results of analytics algorithms are stored as the graph's transient properties of nodes and edges. 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 Rulesets.

ETL can be run in two ways:

1. Through graph pipelines, for more information see Managing Graph Pipelines section.

2. To configure and execute the ETL, follow these steps:

   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 137: Oracle Financial Crime Graph Model
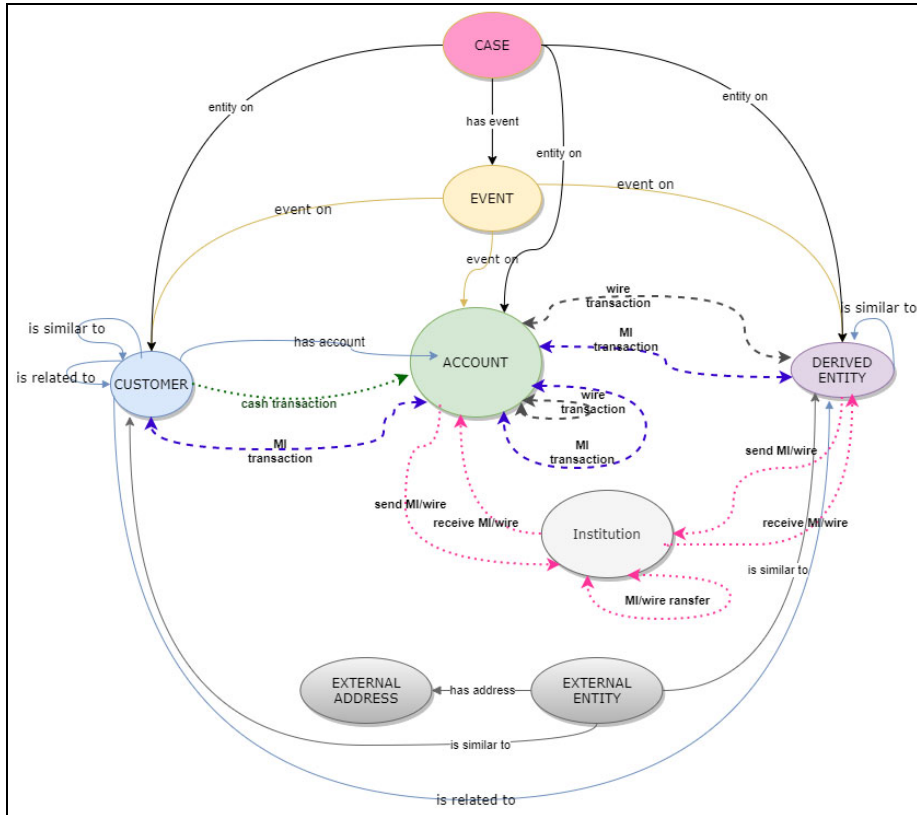
NOTE	The Case node in this Financial Crime Graph Model is loaded only when loading the FCDM data from Enterprise Case Management (ECM). The graph includes "CASE" nodes and "has event" edges when data is loaded from ECM.

11.1.1   Load Data into ICIJ Tables

To load data into ICIJ tables, see the 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, **Importing OOB Graph Definition and other related Metadata** section in the OFS Compliance Studio Installation Guide.

11.1.2   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 Hub. 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.

The following figure illustrates the out-of-box Global Graph Model for BD:

Figure 138: Out-of-Box Global Graph Model for BD
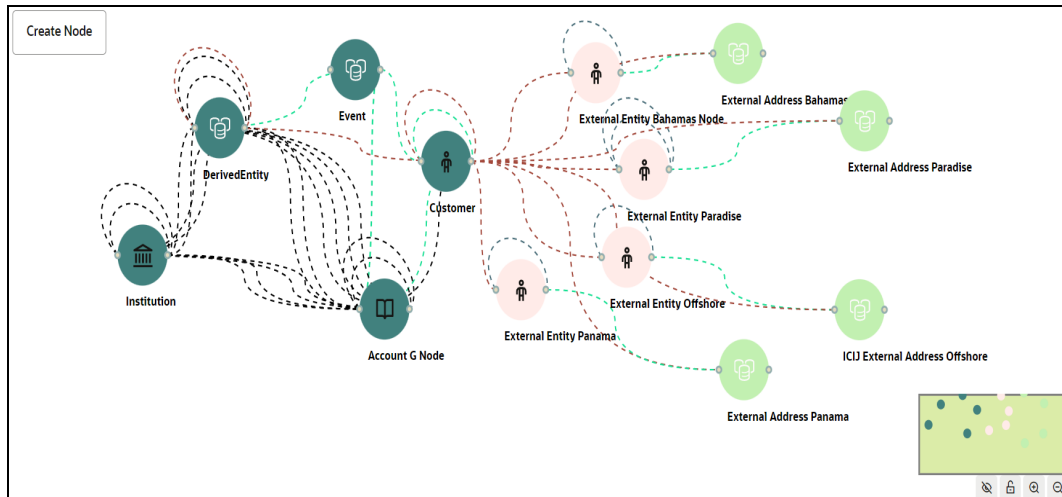
11.1.3 Pre-configured Graph FCGM Model

11.1.3.1 List of Nodes

The Table 34 describes the Node details:

Table 34: Node Details

Name	Description	Pipeline	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
External Entity Panama	ICIJ External Entity - Panama	Data Pipeline	External Entity Panama	Yes
External Address Panama	ICIJ External Entity Address - Panama	Data Pipeline	External Address Panama	Yes
External Entity Offshore	ICIJ External Entity - Offshore	Data Pipeline	External Entity Offshore	Yes
ICIJ External Address Offshore	ICIJ External Entity Address - Offshore	Data Pipeline	ICIJ Offshore Address	Yes

Table 34: Node Details

Name	Description	Pipeline	Pipeline Name	Is Active
External Entity Paradise	ICIJ External Entity - Paradise	Data Pipeline	External Entity Paradise	Yes
External Address Paradise	ICIJ External Entity Address - Paradise	Data Pipeline	External Address Paradise	Yes
External Entity Bahamas Node	ICIJ External Entity - Bahama	Data Pipeline	External Entity Bahamas	Yes
External Address Bahamas	ICIJ External Entity Address - Bahama	Data Pipeline	External Address Bahamas	Yes

11.1.3.2 List of Edges

The Table 35 describes the Edge details:

Table 35: Edge Details

Name	Description	Pipeline	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
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

Table 35: Edge Details

Name	Description	Pipeline	Pipeline Name	Is Active
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
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

Compliance Studio Global Graph

Table 35: Edge Details

Name	Description	Pipeline	Pipeline Name	Is Active
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
Panama EE Is Related to EE	Is ICIJ External Entity - Panama Related to Another	Data Pipeline	Panama EE Is Related to EE	Yes
Panama Address of EE to EA	Is ICIJ External Entity Address - Panama of an ICIJ External Entity - Panama	Data Pipeline	Panama Address of EE to EA	Yes
Offshore EE Is Related to EE	Is ICIJ External Entity - Offshore Related to Another	Data Pipeline	Offshore EE Is Related to EE	No
Offshore Address of EE to EA	Is ICIJ External Entity Address - Offshore of an ICIJ External Entity - Offshore	Data Pipeline	Offshore Address of EE to EA	No
Paradise EE Is Related to EE	Is ICIJ External Entity - Paradise Related to Another	Data Pipeline	Paradise EE Is Related to EE	No

Table 35: Edge Details

Name	Description	Pipeline	Pipeline Name	Is Active
Paradise Address of EE to EA	Is ICIJ External Entity Address - Paradise of an ICIJ External Entity - Paradise	Data Pipeline	Paradise Address of EE to EA	No
Paradise EE Is Linked to EE	Is ICIJ External Entity - Paradise Linked Another External entity - Paradise	Data Pipeline	Paradise EE Is Linked to EE	No
Bahamas EE Is Related to EE	Is ICIJ External Entity - Bahama Related to Another	Data Pipeline	Bahamas EE Is Related to EE	No
Bahamas Address of EE to EA	Is ICIJ External Entity Address - Bahama of an ICIJ External Entity - Bahama	Data Pipeline	Bahamas Address of EE to EA	No
Cust Is similar to Ext Panama	Is Customer Similar to External Entity - Panama	Match rule	Customer To External Source Panama BDG	Yes
Cust Is similar to Ext Address Panama	Is Customer similar to External Entity Address - Panama	Match rule	Customer To External Source Panama Address BDG	Yes
Cust Is similar to Ext Bahamas	Is Customer Similar to External Entity - bahamas	Match rule	Customer To External Source Bahamas BDG	No
Cust Is similar to Ext Address Bahamas	Is Customer similar to External Entity Address - Bahamas	Match rule	Customer To External Source Bahamas Address BDG	No
Cust Is similar to Ext Offshore	Is Customer Similar to External Entity - Offshore	Match rule	Customer To External Source Offshore BDG	No
Cust Is similar to Ext Address Offshore	Is Customer similar to External Entity Address - Offshore	Match rule	Customer To External Source Offshore Address BDG	No
Cust Is similar to Ext Paradise	Is Customer Similar to External Entity - Paradise	Match rule	Customer To External Source Paradise BDG	No
Cust Is similar to Ext Address Paradise	Is Customer similar to External Entity Address - Paradise	Match rule	Customer To External Source Paradise Address BDG	No
Cust is similar to cust	Is Customer is similar to customer	Match rule	Customer to Customer Match BDG	No
Cust Is similar to Derived Entity	Is Customer similar to Derived Entity	Match rule	Customer To Derived Entity Match BDG	No

Table 35: Edge Details

Name	Description	Pipeline	Pipeline Name	Is Active
Derived Entity Is similar to Derived Entity	Is Derived Entity is similar to Derived Entity	Match rule	Derived Entity To Derived Entity Match BDG	No

11.1.3.3 List of Indexes

The Table 36 describes the Out-of-the-box indexes for the match rules that are defined as similarity edge (Edge Details) in the FCGM.

Table 36: Pre-seeded Indexes

Index Name	Attributes		
	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
	status	Status	Text
	tax_id	Tax ID	Text

Index Name	Attributes		
	Name	**Display Name**	**Type**
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 Paradise Graph	address	Address	Text
	jurisdiction	Jurisdiction	Text
	country	Country	Text
	source	Source	Text
	label	Label	Text
External Entity Bahamas 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

Index Name	Attributes		
	Name	**Display Name**	**Type**
External Address Offshore Graph	address	Address	Text
	jurisdiction	Jurisdiction	Text
	country	Country	Text
	source	Source	Text
	label	Label	Text
External Entity Offshore Graph	source	Source	Text
	label	Label	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
External Address Panama Graph	address	Address	Text
	jurisdiction	Jurisdiction	Text
	country	Country	Text
	source	Source	Text
	label	Label	Text
External Entity Panama Graph	source	Source	Text
	label	Label	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

Index Name	Attributes		
	Name	**Display Name**	**Type**
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
	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

Index Name	Attributes		
	Name	**Display Name**	**Type**
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
External Entity Paradise Graph	business_domain	Business Domain	Text
	country	Country	Text
	state	State	Text
	city	City	Text
	name	Name	Text
	jurisdiction	Jurisdiction	Text
	original_id	Original_ID	Text
	label	Label	Text
	source	Source	Text
	address	Address	Text
External Address Bahamas Graph	address	Address	Text
	jurisdiction	Jurisdiction	Text
	country	Country	Text
	source	Source	Text
	label	Label	Text

11.1.3.4 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. The graph model expresses the conditional dependence structure between nodes and edges.

For information on Graph Data Model, see Graph Data Model.

For information on the node and edge properties of the Oracle Financial Crime Graph Model, see the Data Model Guide.

11.2 Entity Resolution

OFS Compliance Studio provides Entity Resolution (ER) capability. It allows firms to break through barriers in their data by gaining single views of their customers and their external entities and have the choice of monitoring them both under one consolidated Global Party.

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**

   Elasticsearch is a distributed search and analytical engine for all structured and unstructured data types in OFS Compliance Studio.

- **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, deduplicating 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 Rulesets section.

- **Grouping**

   It is used to Group (entity Ids or Customer Ids) them based on similarity links between entities 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 Rulesets 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.

11.3 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 of creating Match and Merge Rule (s) is as follows:

Figure 139: Process of creating Match and Merge Rule
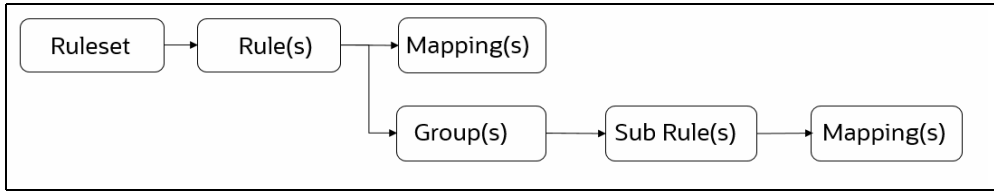
NOTE	• The sum of weightage for mappings and groups for each rule must be 1.
	• The sum of weightage for mappings for each sub-rule must be 1.

11.3.1 Using Ruleset (Match/Merge)

Use this section 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.

Perform the following to use the existing Ruleset:

1. Login to the OFS Compliance Studio application. The Workspace Summary page is displayed.

2. On the Workspace Summary page, select Launch workspace to display the **CS Production Workspace** window with application configuration and model creation menu.

Figure 140: Workspace Summary
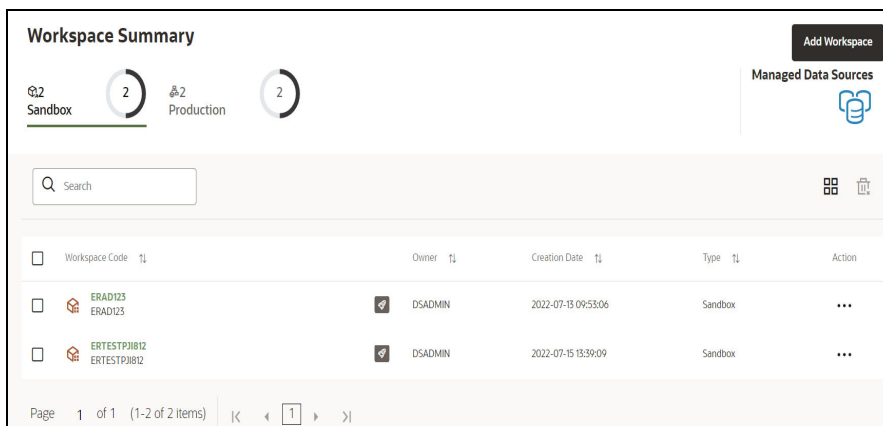
3. Hover the mouse over the Ruleset Details widget the following options are available.

Figure 141: Ruleset Details widget
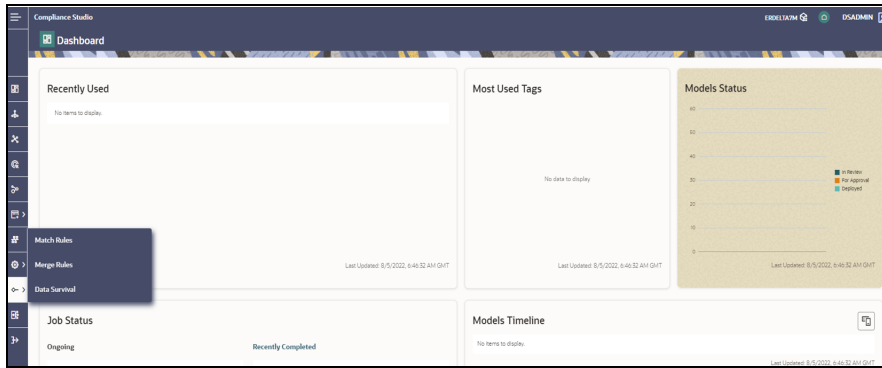
4. Click **Match Rules/Merge Rules**. The Match Rules/Merge Ruleset window is displayed according to the selection.

- Match Rules

  The following figure illustrates the OOB Match Rules for FSDF 808, FSDF 812, and Global Graph Nodes:

  Figure 142: Match Ruleset

- Merge Ruleset

  The following figure illustrates the OOB Merge Rules for FSDF 808 and FSDF 812:

Figure 143: Merge Ruleset
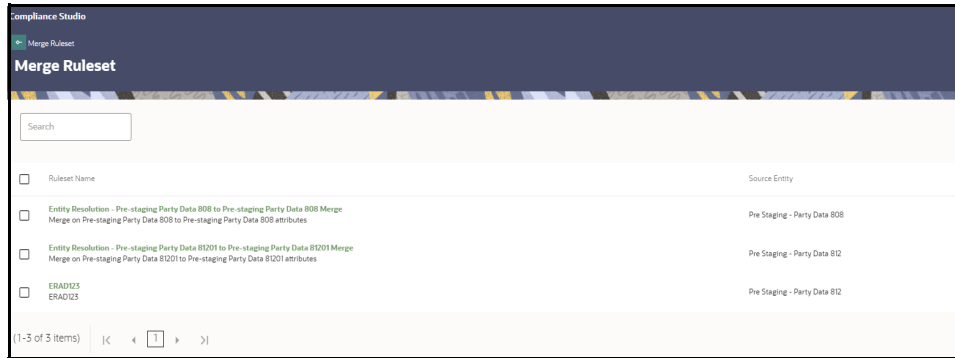
5. On the details page, perform the following actions:

- To enable the Ruleset, select the Ruleset and move the toggle switch to the right.

- To add the Ruleset, click **Add** [Add]. For more details, see Creating Ruleset (Match/Merge) section.

- To edit the Ruleset, click on the **Ruleset**, and the Ruleset page is displayed. On the Ruleset page,

  click **Edit** [Edit]. The page displays Ruleset details.

  You can modify the Ruleset details except for the following fields:

  - Ruleset Name

  - Scoring Aggregation Type

  - Source

  - Target (It is applicable only for Matching)

- 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 information on the fields in the Match/Merge Ruleset window, see the Creating Ruleset (Match/Merge) section.

11.3.2   Creating Ruleset (Match/Merge)

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.

Perform the following to create a new Ruleset:
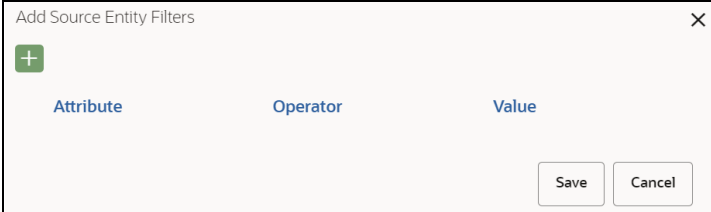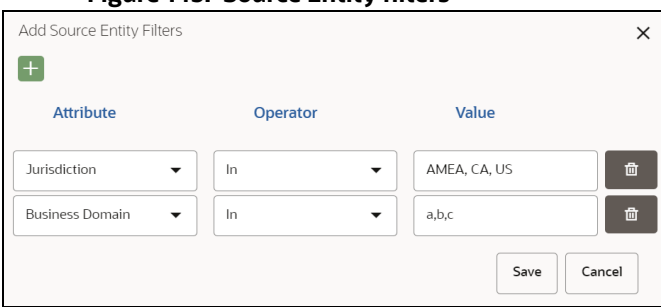
1. On the Ruleset page, click **Add** | Add |. The page displays as follows:

   ■ 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 Table 37:

Table 37: Create a Ruleset table

Item	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.

Item	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. It is an optional step. Multiple filter criteria are allowed in the filter. To add filter conditions, perform the following; 7. Click ⊤ icon. The Add Source entity Filter window is displayed. **Figure 144: Source Entity filters** 8. Click ➕ icon. The Attributes and conditions are displayed. **Figure 145: Source Entity filters** 9. 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. **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). 10. Click **Save** to save the filters. 11. Click **Delete** 🗑 icon to delete the corresponding filter.

Item	Description
Target	This is applicable only to Match Ruleset. Select the target entity. For example, customer, account, address, etc.
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. 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.
ER Type	The type of data that is to be matched. The values are as follows: • CSA_808 • CSA_811 • CSA_812 • Graph (this is applicable only for ETL) These may be nodes within the global graph or relational data to be resolved.
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**.

Item	Description
Manual Threshold	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.

1. Click **Add** . The page displays the attributes to create the rules for the new Ruleset.

   For information on creating Rule(s) in the Ruleset, see the Creating Rules in Ruleset section.

2. Click **Save** to save the Ruleset. You can save the Ruleset with or without creating Rules or click **Cancel** to cancel.

11.3.2.1 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 146: Creating Rules in Ruleset
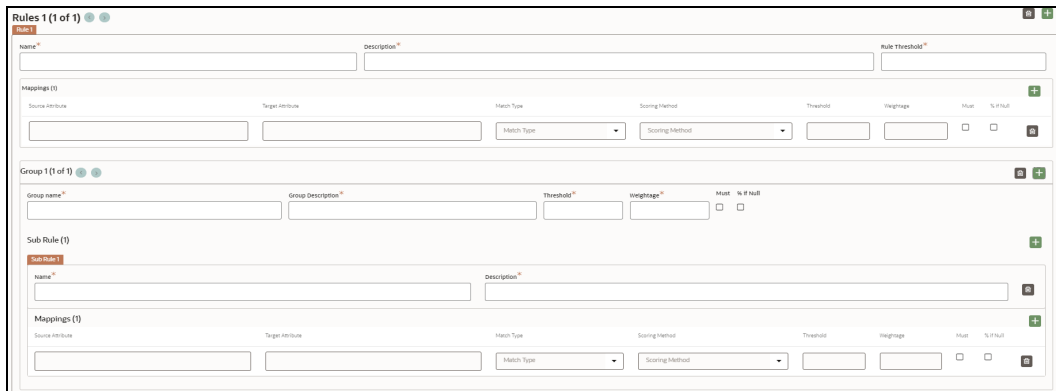
NOTE	• You can create multiple rules in the Ruleset, multiple mapping and groups in Rule, and multiple sub-rules and mappings within the Group.
	• The sum of the Weightage of Group(s) should be equal to the Weightage of the Rule.
	• You can navigate the Rules and Groups within the Rule using the Previous 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.

Perform the following to create Rule(s):

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.

    Table 38 provides descriptions for the fields on the page:

Table 38: Rules, Mappings, and Group table

Field	Description
Rules	
Name	The name of the Rule. **NOTE**: Rule Name field accepts only 255 characters.
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.
Mappings (Rules and Sub Rule)	
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.

Table 38: Rules, Mappings, and Group table

Field	Description
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.
Group	
Group name	The name of the Group.
Group Description	The description of the Group.

Table 38: Rules, Mappings, and Group table

Field	Description
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.
Sub Rule	
Name	The name of the Sub Rule.
Description	The description of the Sub Rule.

3. Click **Delete** 🗑 to delete the Rule.
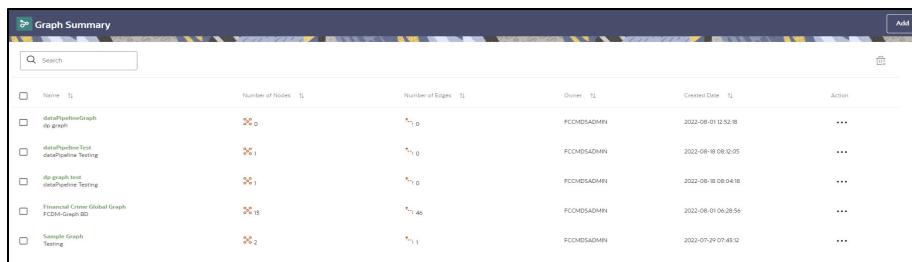
11.3.3   Cloning Ruleset (Match)

Use this section 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.

Perform the following to clone an existing match rule:

1. Navigate to **Workspace Summary > Graphs**. The Graph Summary page is displayed.

Figure 147: Graph Summary

2. Click <**Graph name**>. On the **Edge Details** page, click **Create** drop-down list and then click **Match Rules**. The Create Pipeline window is displayed.

Figure 148: Create Pipeline

3. Select **Clone From an Existing Match Rule** option and required pipeline from the **Select Pipeline** drop-down list and then click **Ok**. The following window is displayed.

Figure 149: Ruleset Details
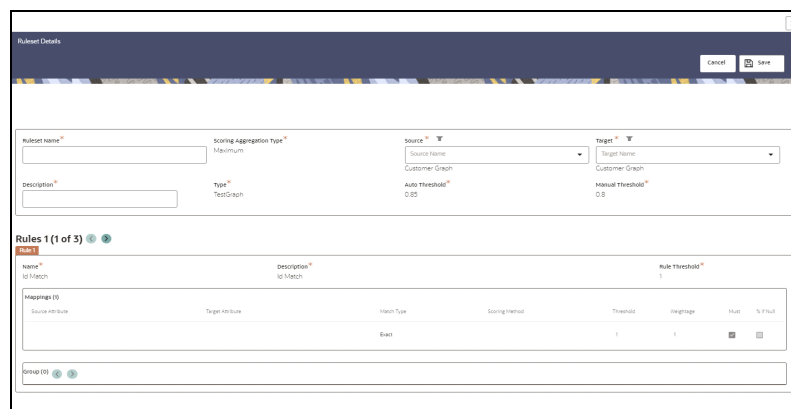
4. You can enter / select the Ruleset details for the following fields:

   ▪ Ruleset Name

   ▪ Source

   ▪ Target

   ▪ Description

   For information on the fields in the Ruleset Details window, see the Creating Ruleset (Match/ Merge) 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.

5. Click **Save**. The new match rule is cloned.

11.3.4 Data Survival Rules

The 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

11.3.4.1 Using Data Survival Rules

Use this section 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.
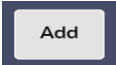
Perform the following to use the existing Ruleset:

1. On the **Workspace Summary** page, select Launch workspace to display the **CS Production Workspace** window with application configuration and model creation menu.

2. Hover the mouse over the Ruleset Details widget and click **Data Survival**. The Data Survival Rules page is displayed.

Figure 150: Data Survival Rules
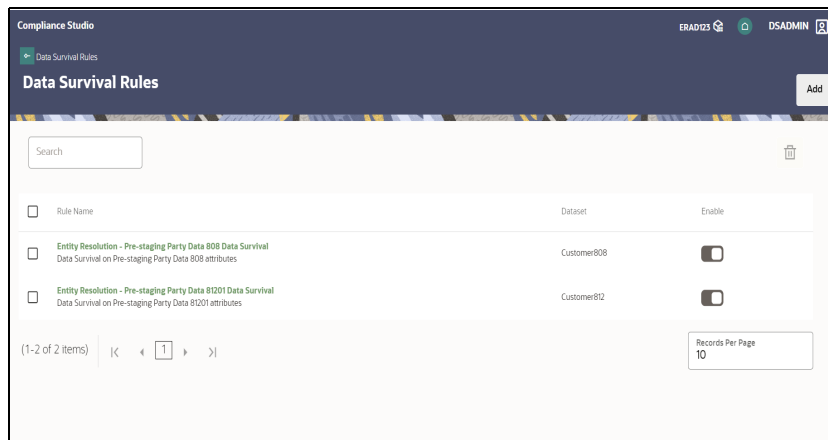
On the Summary page, perform the following actions:

— To enable the Rule, select the Rule and move the toggle switch to the right.

— To add the Rule, click **Add** . For more details, see Creating Data Survival Rules section.

— To delete the Rule(s), select the checkbox(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**.

The following figure illustrates the OOB Data Survival Rules for FSDF 808.

Figure 151: Data Survival Rules pre-standing tables for FSDF 808

The following figure illustrates the OOB Data Survival Rules for FSDF 812:

Figure 152: Data Survival Rules pre-standing tables for FSDF 812

3.   On the details page, perform the following actions:

To edit the Rule, click on the **Rule**, and the corresponding Rule page is displayed. On the Rule

page, click **Edit** ✎   Edit . The page displays Rule details.

You can modify the Rule details except for the following fields:

- ▪   Rule Name
- ▪   Description
- ▪   ER Type
- ▪   Dataset

NOTE	Fields marked with * are mandatory fields.

Figure 153: Data Survival Rule Page
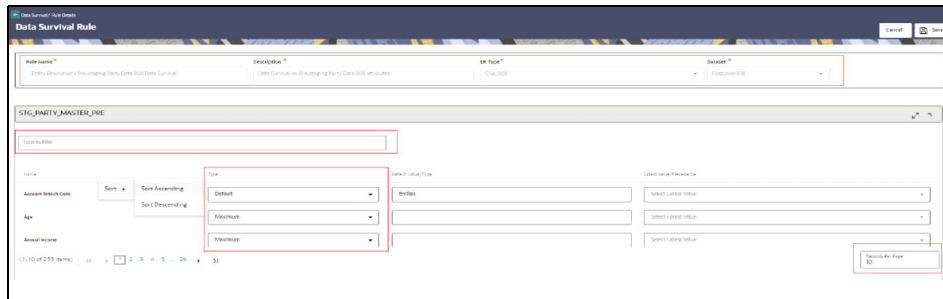
On this page, perform the following actions.

- 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. |

- Click **Save** to save the changes or click **Cancel** to revert the changes.

11.3.4.2   Creating Data Survival Rules

1. On the Data Survival Rules page, click **Add** [Add]. The page displays the attributes to create the Rule.

2. Enter the field values.

   Table 39 provides descriptions for the fields on the page:

Table 39: Data Survival Rules fields

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_808** **CSA_811** **CSA_812** These may be nodes within the global graph or relational data to be resolved.

Table 39: Data Survival Rules fields

Field	Description
Dataset	The name of the output data structure to which this data is saved. The values are as follows: • **Customer808** • **Customer811** • **Customer812**

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 154: Survival Function dropdown
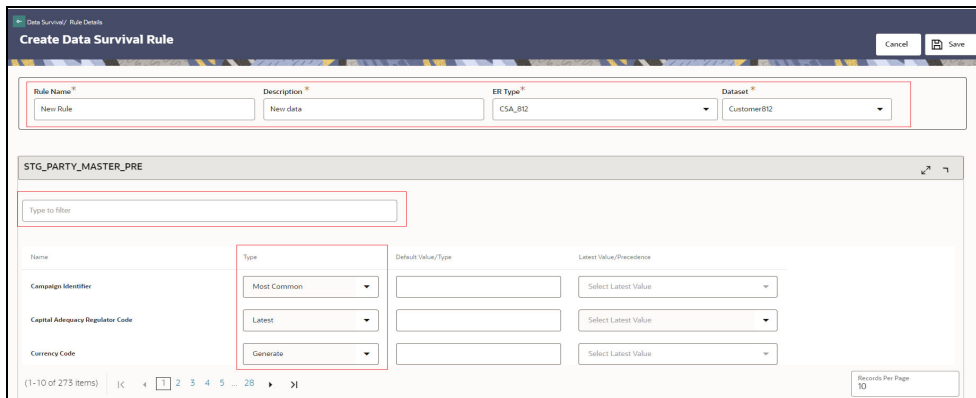
On this page, perform the following

- 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.

3. 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.

NOTE	This is applicable only for numbers attributes.

- **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.

NOTE	This is applicable only for numbers attributes.

- **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.

4. Click **Save** to save the Rule or click **Cancel** to revert the changes.

11.4 Using Manual Decisioning

The **Manual Decisioning** in Compliance Studio 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.

- **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.

- **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.

You can specify your **Manual Threshold** on the Match Ruleset Details page for Manual Decisioning.

Figure 155: Manual Threshold

To navigate the **Manual Decisioning** page, click the **Manual Decisioning** widget . The **Manual Decisioning** page is displayed.

The Manual Decisioning page will be displayed as follows:

Figure 156: Manual Decisioning

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 user privileges. If both icons are shown, you can choose whether to make a decision as an Analyst or Supervisor.

You must select the **ER Type** to display the records on the page. Other than **Graph**, records will be displayed only when you select the **Data Source** for **Entity Resolution (CSA8xx)** (ER Type).

The records on the Manual Decisioning page will be displayed on the **New** tab based on the selection of **ER Type**. The **Data Source** field is applicable only when **ER Type** is selected as **Entity Resolution (CSA8xx)**, and it is disabled in the case of **Graph**.

Records are displayed on one of the following tabs, depending on where they are in the Manual Decisioning workflow.

New (where no decision has yet been made)

- Pending Approved (where an Analyst has made an approve decision and this is to be confirmed or rejected by a Supervisor)

- Pending Rejected (where an Analyst has made a reject decision and this is to be confirmed or rejected by a Supervisor)

- Approved (where the decision has been approved)

- Rejected (where the decision has been rejected)

See the Manual Decisioning Workflow section for more details.

Figure 157: Manual Decisioning
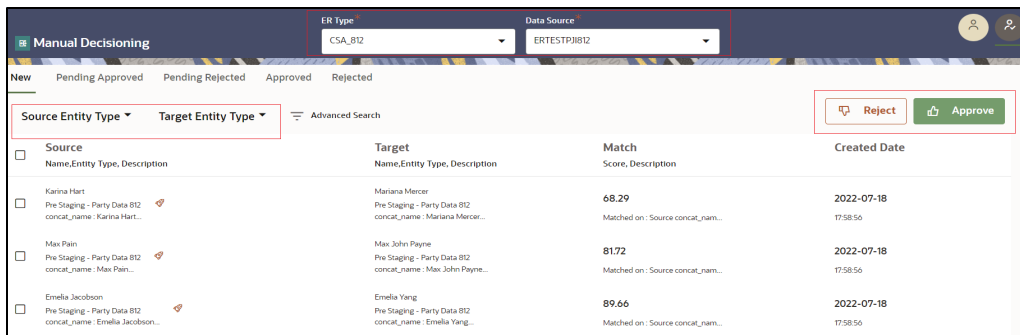
By default, ten edges or matches are displayed per page. Click the **Next arrow** to view the next set of edges or matches.

The following fields are available on the page:

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

Figure 158: Advanced Search screen
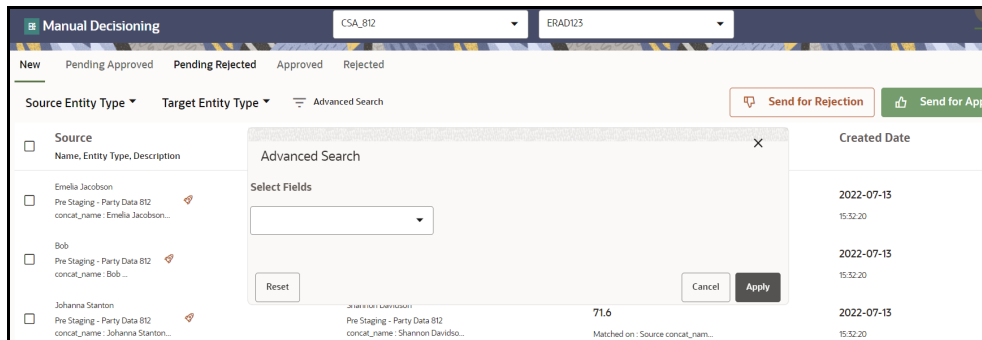
· To customize the view on the page, perform the following:

a. Select any column on the page.

b. Right-click and the following menu options are displayed:

Figure 159: Manual Decisioning screen
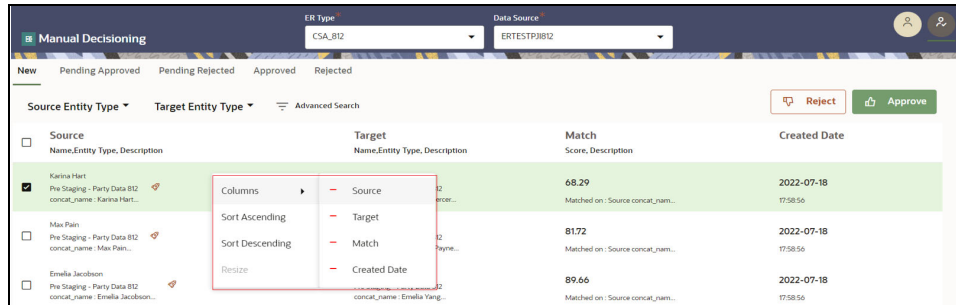
— **Columns:** You remove and add the column form list.

— **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**: You can resize the column by specifying the column width.

To filter the edges on the page, perform the following:

■ **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 that the view edge is different for **Graph** and **Entity Resolution (CSA8xx)**. For more information, see the Viewing the Edges from the Source section.

11.4.1   Manual Decisioning Workflow

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 160: Manual Decisioning workflow
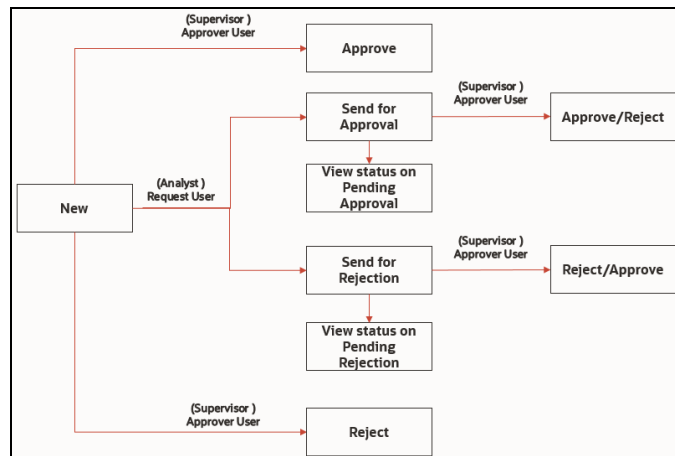
11.4.2   Searching for Manual Decisions

You can specify the Advanced Search using AND logic. It will display the edges only when they satisfy the search criteria.

To filter the edges in the Advance Search, perform the following:

1. Click **Advanced Search**. The Advanced Search window appears.

Figure 161: Advanced Search screen

2. Select the fields from the Select Fields drop-down list. You can add multiple search criteria.

Figure 162: Select the option in the search screen
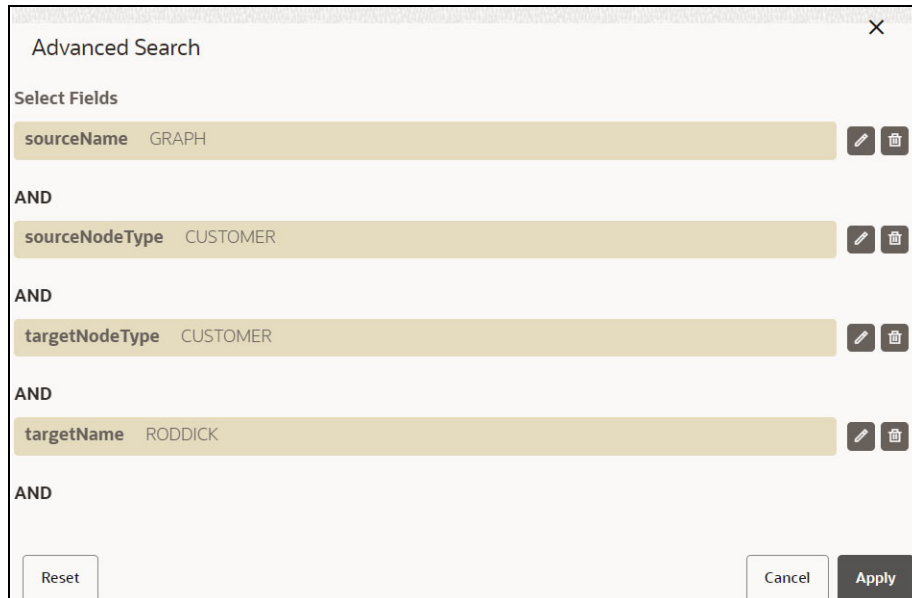
You can perform the following actions:

- Click **Add** icon to add the field value for the selected Field.

- Click the **Edit** icon to edit the field value for the selected Field

- Click the **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.

11.4.3   New, Pending Approved, and Pending Rejected Tabs

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).

See the Manual Decisioning Workflow section for more details.

11.4.4   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.
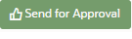
No further actions can be taken from these tab.

11.4.5   Taking a Decision on Similarity Edges

You can take the decision as an **Analyst** or **Supervisor**.

1. Select the edge and perform the following:

   ■ Analyst

      — Click on [Send for Approval] to send for approval.

      — Click on [Send for Rejection] to send for rejection.

   ■ 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 choose **Standard Comments** from the drop-down list for selected edge(s):
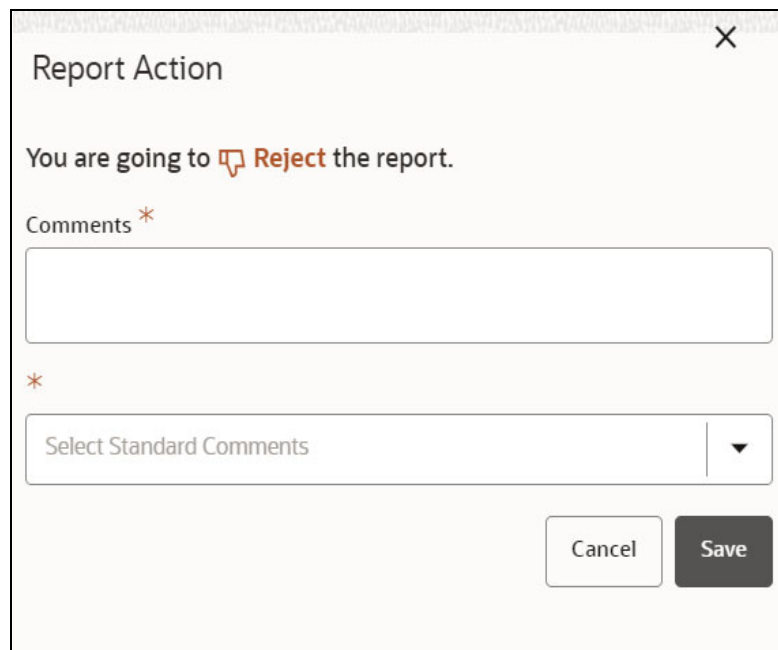
   ■ For Send for Approval or Approve:

      Figure 163: Send for Approval/Approve

■ For Send for Rejection or Reject.

Figure 164: Send for Rejection/Reject

· You can provide free-form comments for the selected similarity edge using the **Comments** text box to provide your own comments or choose 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.

11.4.6 Viewing the Edges from the Source

On the **Manual Decisioning** page, when you select the edge from the action tab (that is, **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 and validate it to include in your graph.

- Graph

Figure 165: Manual Decisioning
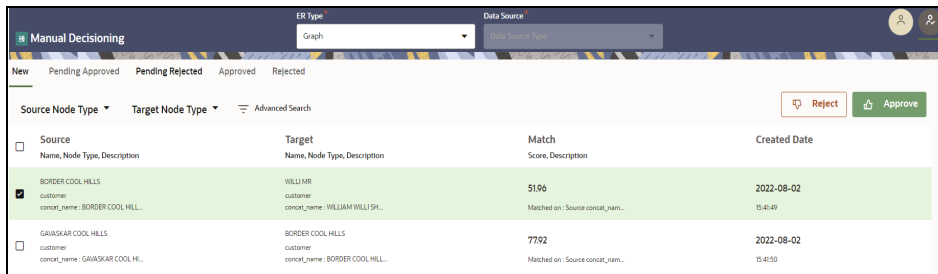
- Entity Resolution (CSA8xx)

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)**.

Figure 166: Manual Decisioning screen for Entity Resolution (CSA8xx)
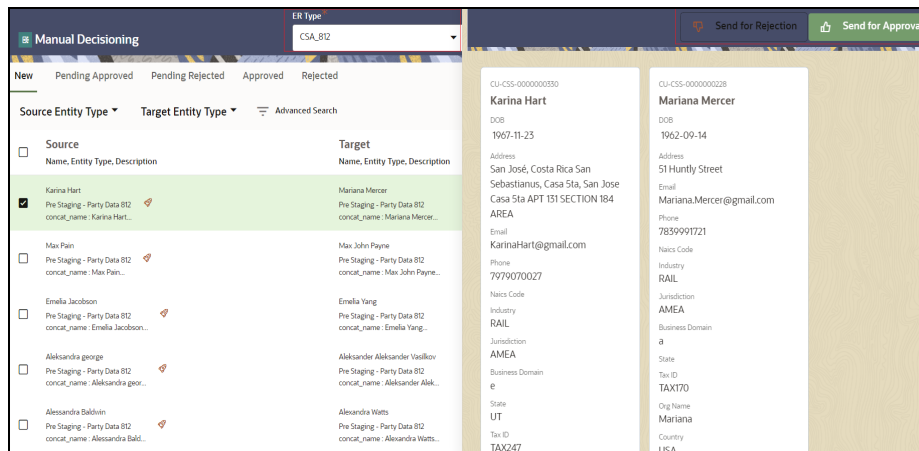
| NOTE | - If you are viewing the paginated graph, increase the Page Size of the graph in Manual Decision UI to avoid the Paginated graph.
- Navigate to **Settings** > **Visualization** > Increase the Page size to accommodate the graph on one page.
- You need to unlock the Manual Decision notebook from the Notebook server UI to view the graph. |
|---|---|

11.5 Using 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.

To navigate the Merge and Split Global Entities page, click **Merge and Split Global Entities** widget ⟨⟩. The Merge and Split Global Entities page is displayed.

The Merge and Split Global Entities page will be displayed as follows:

Figure 167: 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 168: Merge and Split Global Entities

.
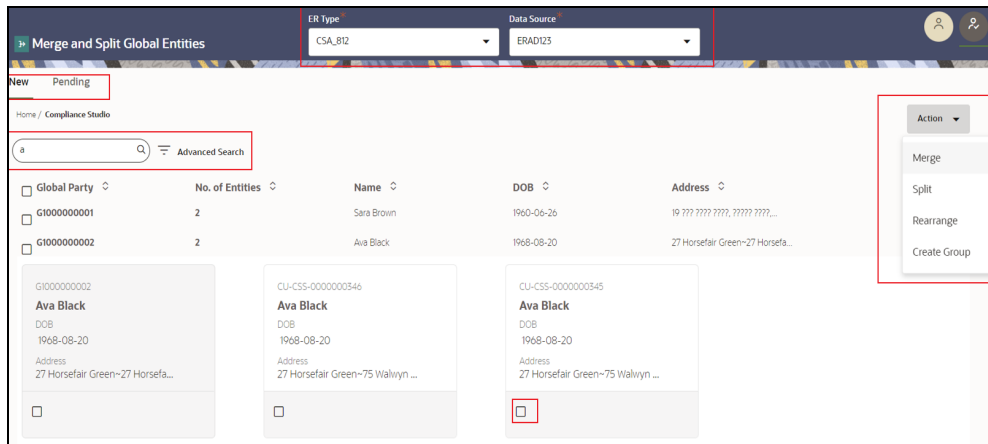
By default, the ten records are displayed per page. Click the **Next arrow** to view the next set of records.

You can filter the records using the search filter 🔍 based on a customer name or customer ID and **Advanced Search** based on search criteria.

NOTE	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 169: Card View
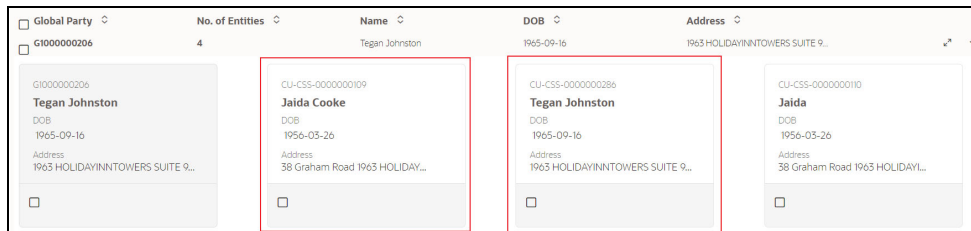
To help in making a decision, you can view the Expand the records to Full view (to compare the customer records with Global Party details)

Full View: Highlighted the comparison in the Customer details with respect to Global Party details. You can next customer details by clicking the **Next** and **Previous arrows**.

Figure 170: Next and Previous arrows

Analysts and Supervisors can perform the following actions:

- **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.
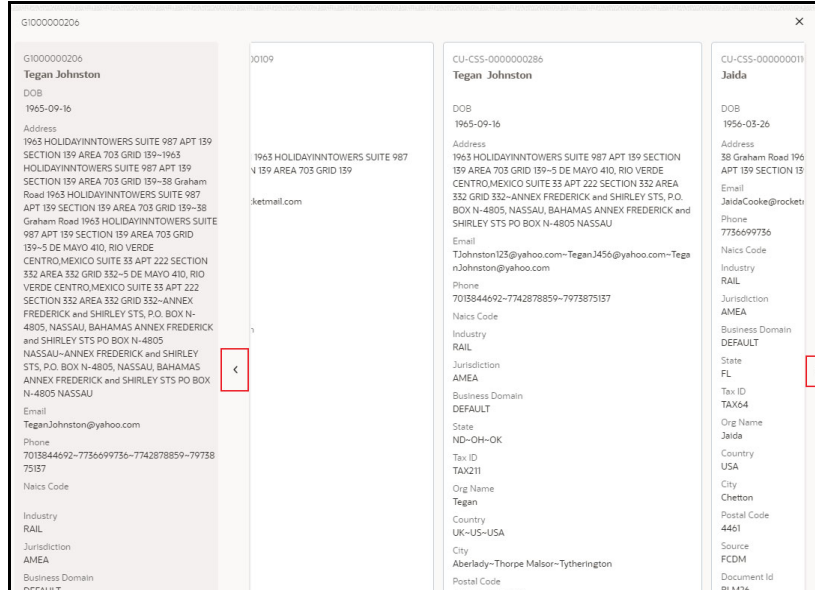
- **Create Group**: A new global party is created to which customers can be moved. New global party ids will be assigned.

Additionally, supervisors can take the following actions on the pending 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.

Supervisors can check the comment made by the analyst as follows:

1. Hover the mouse over the Global Party ID, and the Comment option is displayed.

   Click on the **Comments**. The Comments window is displayed with provided comments.

Figure 171: Merge and Split Global Entities
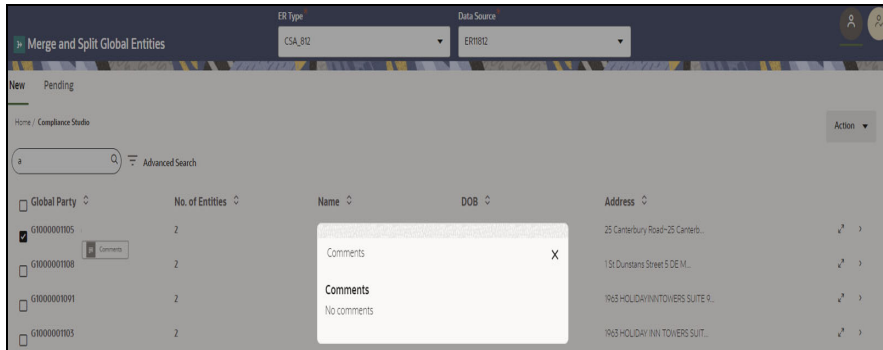
NOTE	Ensure that you have mapped the Data Source to the Workspace to list the data source values in **Data Source** drop-down list on the Manual Decisioning page. For more information on configuring the data sources, see Configuring Workspace Schema section.

11.5.1   Merge and Split Global Entities Workflow

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 actions taken by the Analyst.

- Approve
- Reject

The following figure illustrates the Merge and Split Global Entities workflow.

Figure 172: Merge and Split Global Entities workflow
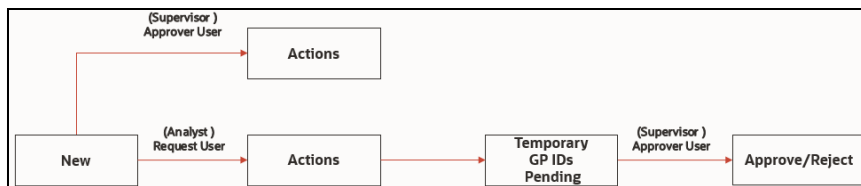
11.5.2   Searching for Merge and Split Global Entities

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 Advance Search, perform the following.

1. Click **Advanced Search**. The Advanced Search window appears.

Figure 173: Advanced Search screen

2.   Select the fields from the **Select Fields** drop-down list. You can add multiple search criteria.

Figure 174: Select the option in the search screen
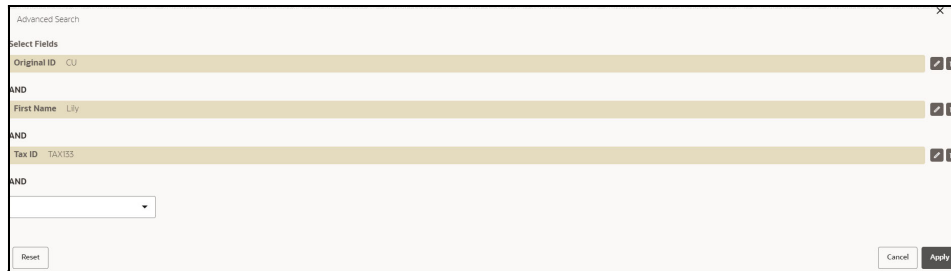
You can perform the following actions:

- Click **Add** icon to add the field value for the selected Field.

- Click the **Edit** icon to edit the field value for the selected Field

- Click the **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.

11.5.3   New Tab

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 in the **New/Compliance Studio** breadcrumb (Indicates the current page's position).

Use breadcrumb locator links to navigate back to Home after checking the action results.

The results are displayed for each action (Merge/Split/Create) in **New/<Actions>** with a new Global Party ID and the number of customers in **No. of Entities**.

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. |

When the Analyst takes action, the records will be moved as Temporary in the Pending tab that needs to be reviewed by the Supervisor.

The Supervisor can review the analyst's action on the Pending and take a further decision (Approve or Reject).

See the Merge and Split Global Entities Workflow section for more details.

Figure 175: Fields on New tab
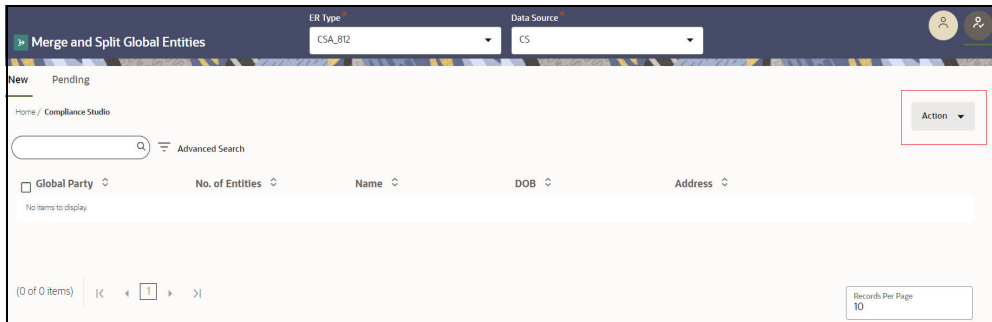
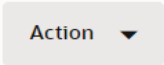Analyst or Supervisor can perform the following actions:

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 and select any one of the actions according to your requirement:

   ■ **Merge**: You must select at least two Global Parties.

     NOTE: It can be 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 **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 to provide your own comments or choose from the **Select Standard Comments** drop-down list.

Figure 176: Create a New Group pop-up
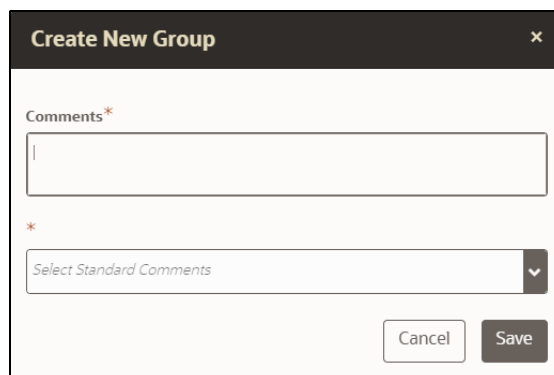
The available standard comments are as follows:

- 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

- Split Entity as DOB missing or conflicting

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.

5. The results are displayed to rearrange on **New/[Action]** Tabs as follows:

Figure 177: New/[Action] Tabs

The new Global Party has the selected records while creating the group, and existing the Global Party is updated with the number of records in **No. of Entities**.

11.5.4   Pending Tab

The Pending tab displays parties for which an analyst has taken action and which need to be approved by a Supervisor. This tab shows **Temporary Party, No. of Entities, Name, DOB,** and **Address** information**.**

Figure 178: Pending Tab
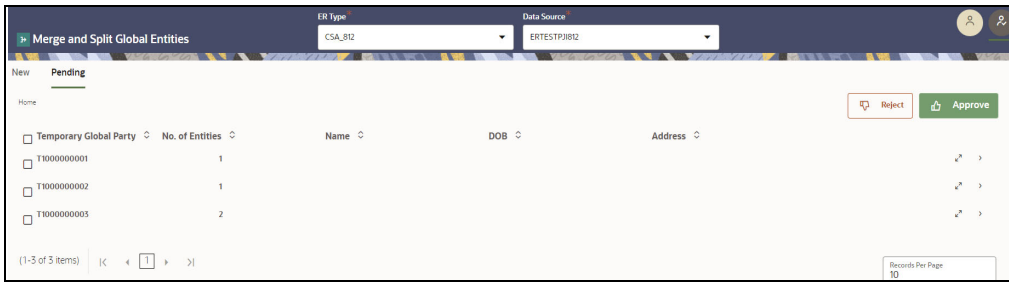
| NOTE | The Name, DOB, and Address data are not populated for new global parties until the change is approved, but the underlying party records can be seen.
| | This information will be added in a later release. |

To Approve or Reject as Supervisor on actions taken by the Analyst, perform the following:

1. Select the record(s) and perform the following:

    ■ Click on **Approve** to approve.

      The match is resolved as per the action the Analyst took.

    ■ Click on **Reject** to reject.

      The initial state of the Global Party before manual action will be retained.

2. A pop-up window appears with details of Global Party IDs and actions (split, merge, create, rearrange) for Approve and Reject:

    ■ For Approve:

      Figure 179: Approve
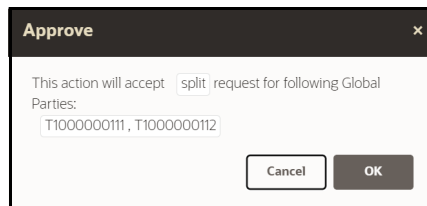
    ■ For Reject:

      Figure 180: Reject
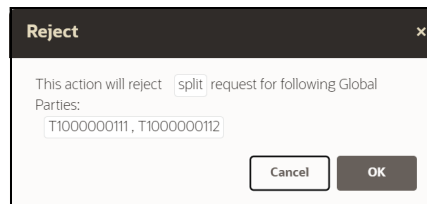
3. Click **OK** to approve the action or click **Cancel** to revert the changes.

11.6 Using Provided Notebooks

Compliance Studio has a collection of Notebooks that are packaged along with the license. These notebooks are deployed directly on the Notebook Server but can be exported and loaded as models if required.

11.6.1 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 7 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.

11.6.2 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.

11.6.3 Creating a Data Discovery Notebook or Model

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.

The sections in this chapter will quickly walk you through building your notebook or model based on the common scenarios explained in the following sections and get a quick insight.

11.6.3.1 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

11.6.3.2 Creating a Notebook for Your Financial Crime Discovery

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 181: Create a paragraph

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 182: 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.

Figure 183: Entities data

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.

Figure 184: Available Entities
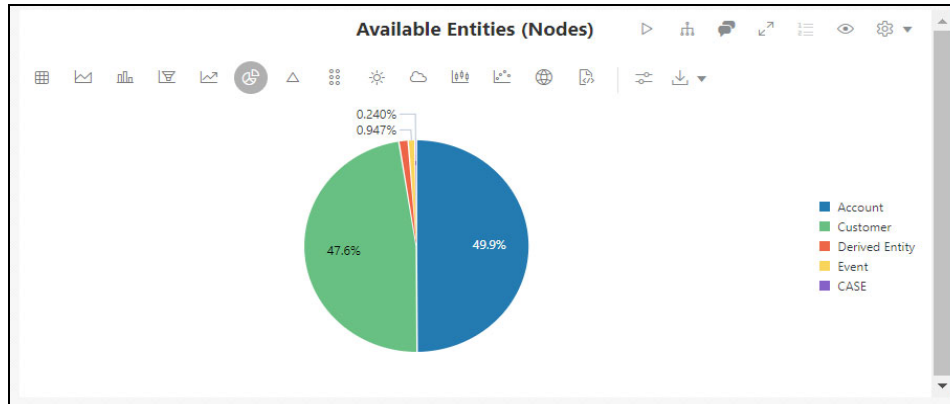
The visualizations displayed are easy to consume, allowing you the flexibility to choose the type of visual. The visualization, in this case, is a **Pie Chart**, which enables you to see the entities associated by their type based on the percentage.

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.

Figure 185: Execute the paragraph


```
%pgql
SELECT e.relationship,e.type,count(*) from g
MATCH (v1)-[e]->(v2)
GROUP BY e.relationship,e.type
```

In this example, the parameters included are:

- SELECT

  The `code e.relationship,e.type,coun(*)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 186: 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.

Figure 187: Available Edges
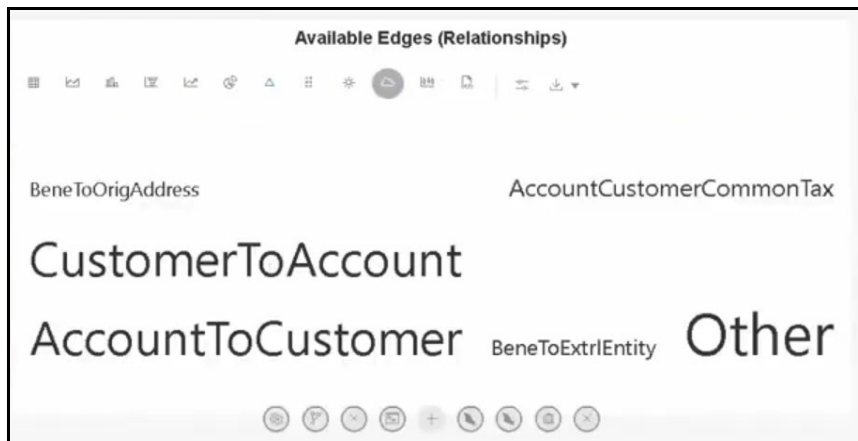
11.6.3.3 Loading a Financial Crime Graph

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, perform the following steps:

1. Create a paragraph and write the description of the analysis or your use case. For example, see the figure.

Figure 188: Description of use case

2. After executing the paragraph, the plain text is formatted and displayed as shown.

Figure 189: Plain Text displayed

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 figure.

Figure 190: Code to create another paragraph

3. Click **Execute Paragraph** to visualize your pattern.

Figure 191: Graph result
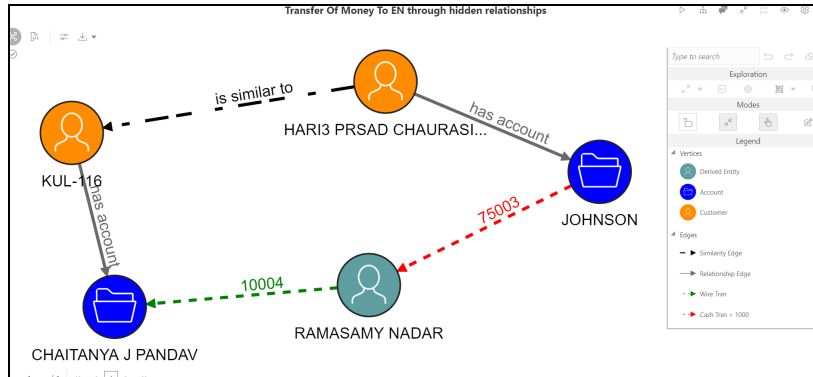
You can hide the code and add a title to the data that is fetched to make it presentable, consumable, and sharable with others. You can provide more highlights (such as labels and vertices), rearrange the nodes, and toggle the graph to enrich your graph, and you can open the paragraph in the iframe mode or expand the graph to make it more readable.

11.6.3.4   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 **General**, **Visualization**, **Highlights**, and **Smart Explorer**. For more information, see Customizing Graphs.

The following figure is customized using **Visualization** and **Highlights** in Graph Settings.

The graph is displayed as follows by providing Labeling changes, such as **Name Label** and **Edge Label**.

Figure 192: Graph settings
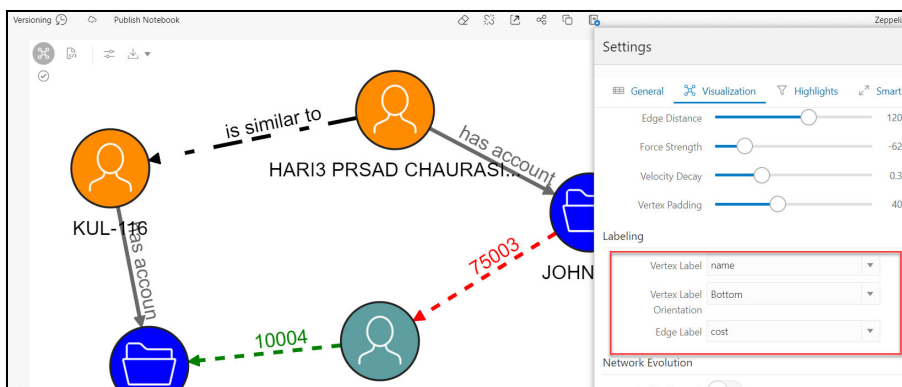
11.6.4   Highlighting the Nodes and Edges of the Graph

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.

11.6.4.1 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 193: Vertices settings
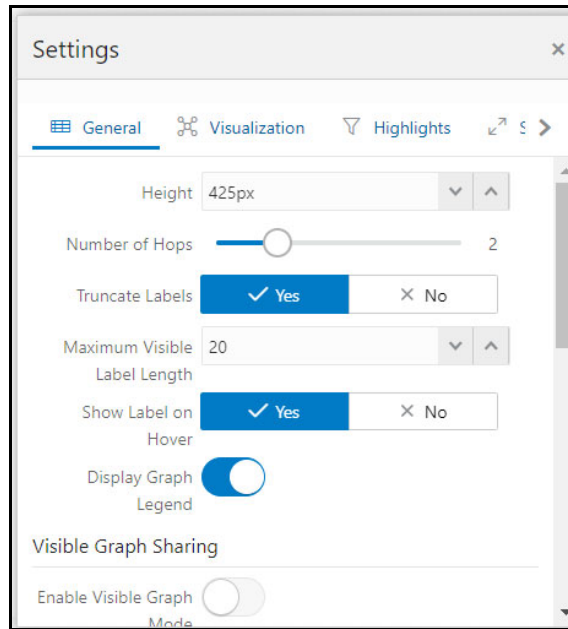
Based on the applied highlights, the graph changes are displayed as follows.

Figure 194: Graph Legend
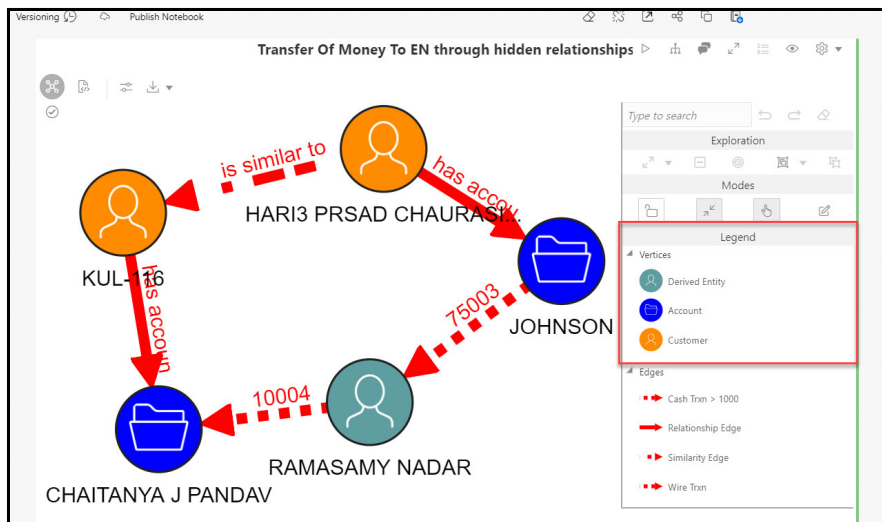
The graphs illustrate a pattern with a money flow from the different entities with a hidden relationship.

11.6.4.2 Edges

Edge can be highlighted by clicking the **New Highlight** and specifying the highlights' parameters. For example, the figure shows how you can differentiate the edges with different colors and animations to visualize money flow from the entities.

Based on the applied highlights, the graph changes are displayed as follows.

Figure 195: Graph Edges
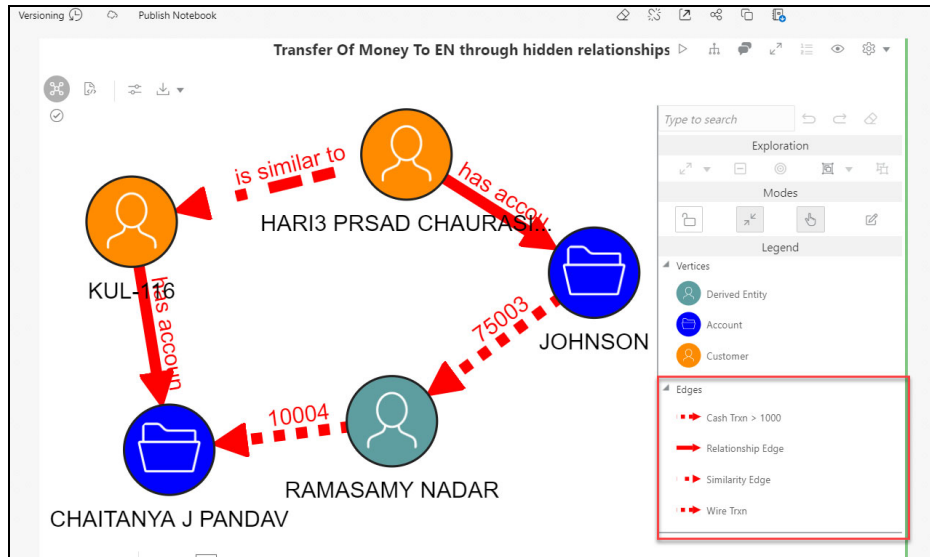
The graphs illustrate a pattern, where there is a money flow from the bank to the different accounts from a customer, with the flow of money highlighted with different colors. Blue represents the money flow from the account to the bank, and red represents the money flow from the bank to the account.

You can also specify the risk numbers to this pattern and visualize your risk. For example, if you use the risk parameters as highlights.

The graphs illustrate a pattern where there is a risk in the account, and when you click for **details**, the risk number is displayed.

Figure 196: Detailed view
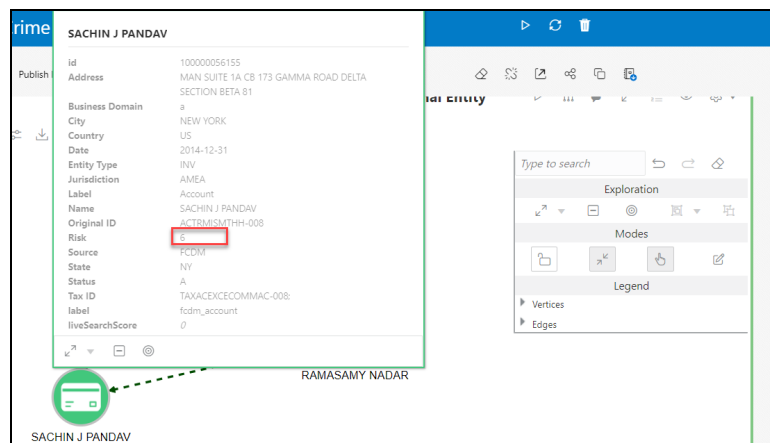
By exploring the options available within the graph customization, you can streamline your analysis and present your reports based on your use case.

11.7 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 1 Details and Table 2 Details.

11.7.1 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 40. One record is inserted for each focus (account or customer).

Table 40: 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.
V_DATA_ORIGIN	AMLBIG	Data origin, that is, MAN, DLY, etc.
V_STATUS_CD	NEW	Event status code. Default value: New.
D_EVENT_CREATED_DAT E	3/11/2019 15:02	Event creation date.
V_JURISDICTION_CD	AMEA	Jurisdiction.
V_BUSINESS_DOMAIN_C D	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.

Table 40: Table 1 Details

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	10	
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	
REMARK		
SCENARIO_CD	10000	
SCENARIO_MN	RMF Account(sql)	Name of scenario notebook.

Table 40: Table 1 Details

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.

11.7.2   Table 2 Details

Table 41 contains information on all the transactions involved in alert generation.

Table 41: 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.

11.8   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.

11.8.1 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.

11.8.2 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.

11.9   Mapping Graph Datasource

To map the graph data source, perform these steps:

1. Login to the Compliance Studio.

2. On the Workspace, click **Action** icon and select **Edit Workspace.**

   Figure 197: Edit Workspace
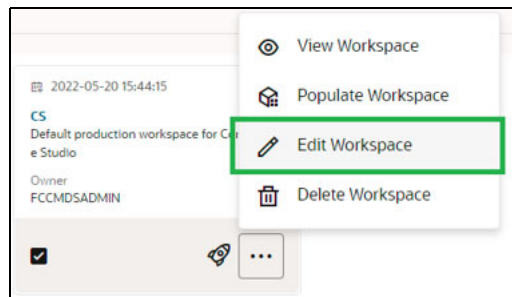
3. On the **Workspace Schema**, map the **graph** data schema.

   Figure 198: Workspace Schema

4. Click to go to the next step.

11.10 Using Graph Definition

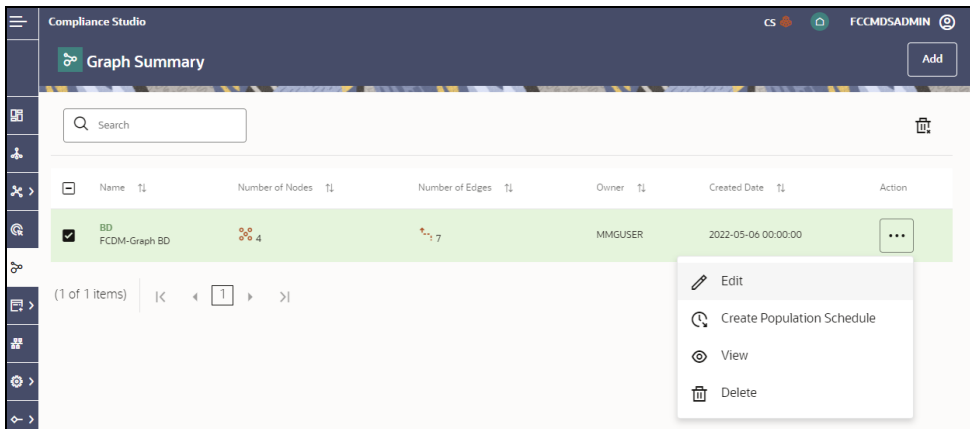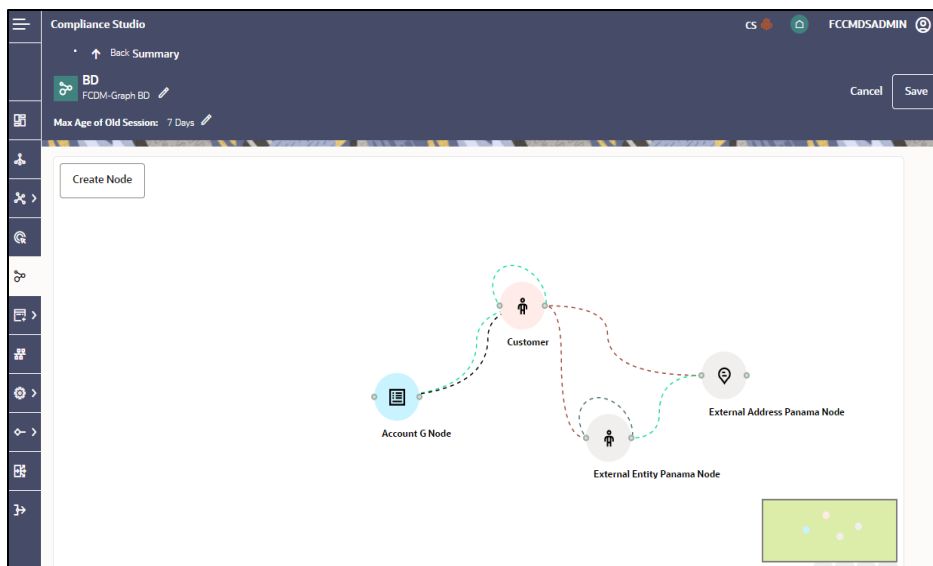To use the graph definition, perform these steps:

1. Log in to the Compliance Studio and click **Graph Options** icon.

2. Select the Out-of-the-box graph BD, and then click **Action > Edit**.
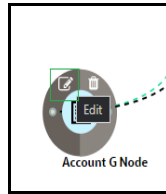
Figure 199: Graph Definition

3. On the **Graph Definition** screen, edit the Node and Edges to attach respective Data Pipeline.
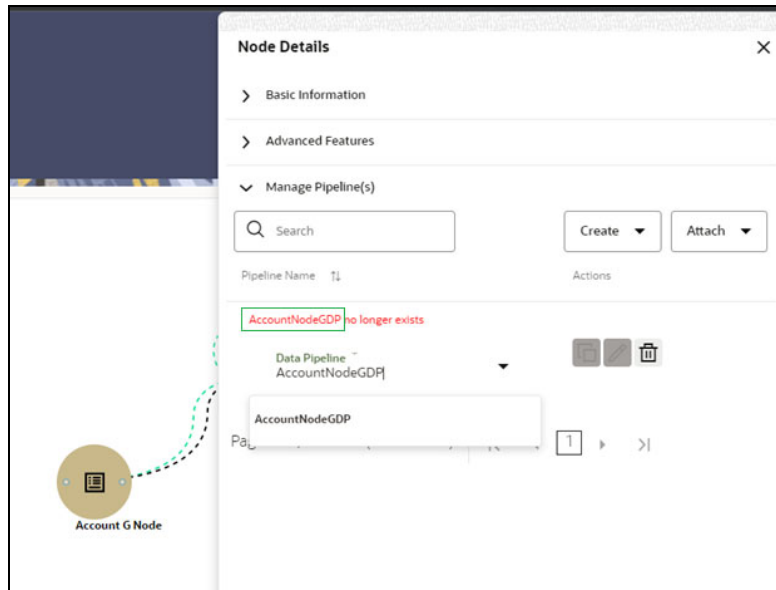
Figure 200: Create Node

a. To edit the Node, hover over the **Node** Icon and click **Edit**.
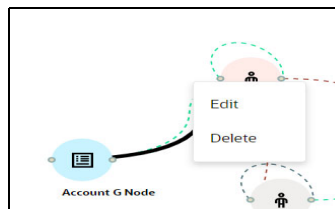
Figure 201: Edit Node

b. On the Node details page, expand Manage Pipeline(s) and map the respective Data pipeline by selecting the data pipeline name as shown below:
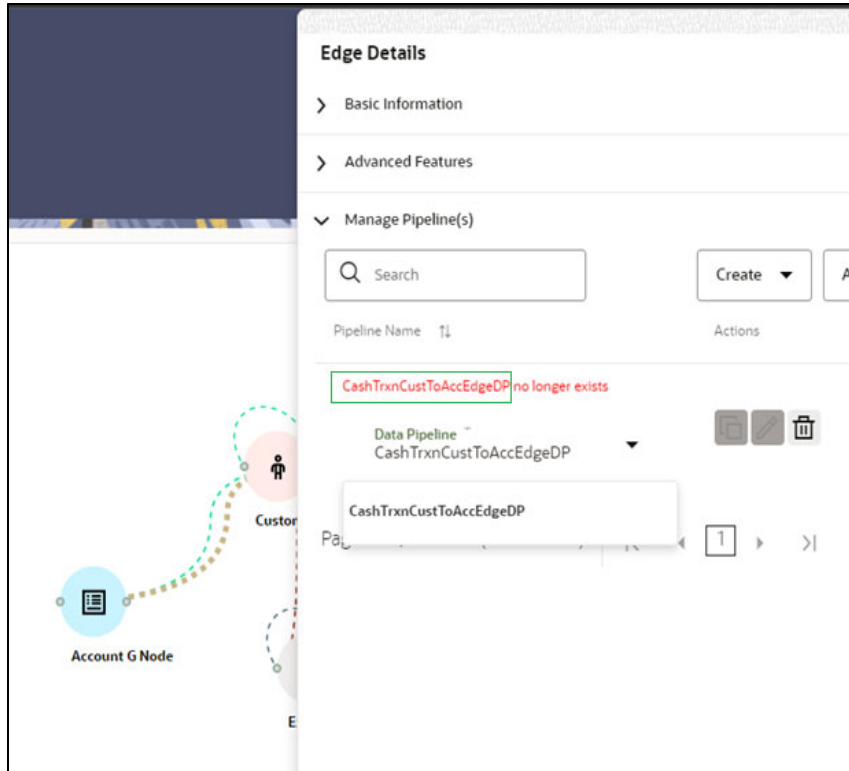
Figure 202: Node Details

c. Click **OK** to attach the data pipeline of other nodes.

d. To edit the Edge, select the edge and click **Edit**.

Figure 203: Editing the Edge

e. On the Edge details page, expand Manage Pipeline(s) and map the respective Data pipeline by selecting the data pipeline name as shown below:

Figure 204: Edge Details

f.   Click **OK** to attach the data pipeline of other edges.

g.   While editing the edge "Cash Trxn Cust To Acct Edge", open and close the Data-pipeline to update the attributes defined by the following steps:
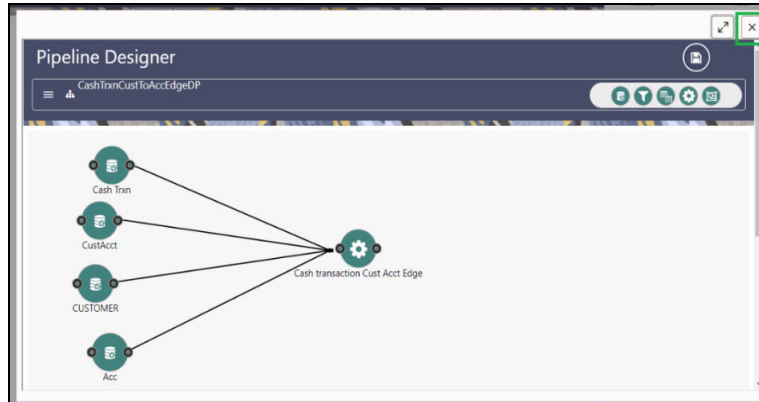
   i.   Click **Edit**.
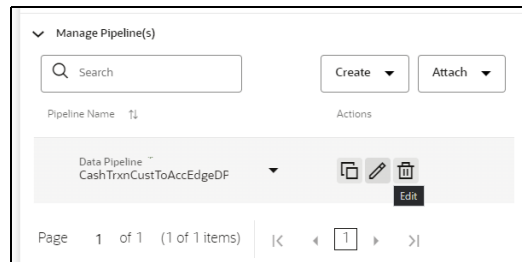
NOTE	The Match Ruleset drop-down list will be displayed if it is a Similarity Edge. Any ruleset that is available on this drop-down list should not be edited from the Match Ruleset window.

   ii.   Once the screen is loaded, click ✕ to close the page.

Figure 205: Pipeline Designer

iii. Once all the data pipelines are attached, click Save.

> **NOTE**: In case of error, follow the message to attach the missing data pipeline of node/edge.

4. Update the run-time parameter and create the batch by following the below steps:

| **NOTE**	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.

a. Click **Create Population Schedule**.

Figure 206: Action Column
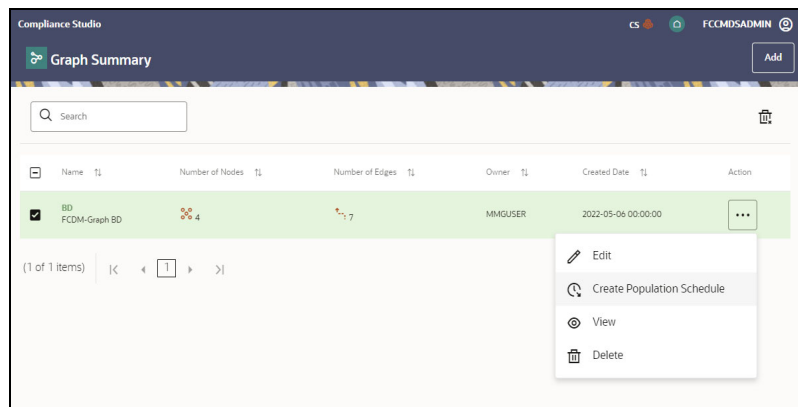
b. To edit the schedule of "Ad-hoc", click Edit under Actions column.

Figure 207: Add Schedule
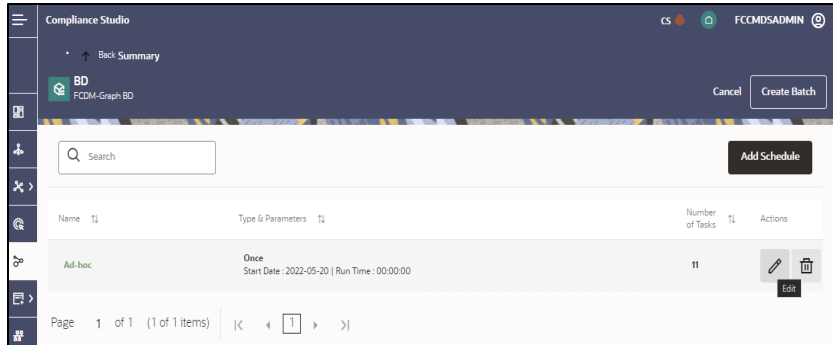
c. On the **Add Schedule** screen, select all the tasks and click **Next**.

Figure 208: Add Schedule Window
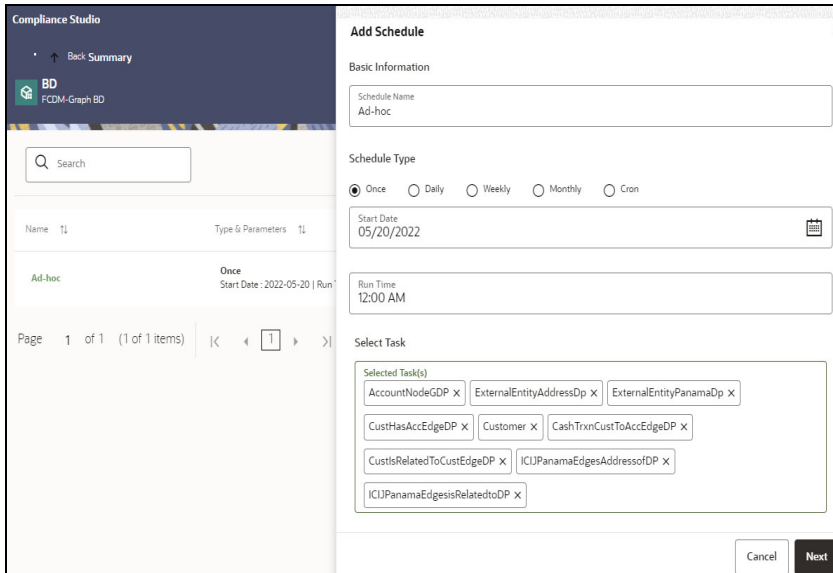
d. Update the value for "$BD_SCHEMA", "$GRAPH_SCHEMA" and "$GRAPH_DS" in the selected task(s) where the value will be respective schema wallet alias.

Figure 209: Add Schedule Window

e. Click **Add** and then Create Batch.

Figure 210: Create Batch

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 2nd- Aug.

The batch can be executed via Scheduler. For more information, see Using Scheduler Service section.

11.11 Managing Data Pipeline

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.

11.11.1 Widgets in Data Pipelines

Depending on the pipeline type, specific widgets are available in the widgets pane of the pipeline.

The Table 42 describes the widgets available in Data pipelines.

Table 42: Data Pipeline – Widgets and Descriptions

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.

Table 42: Data Pipeline – Widgets and Descriptions

	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.

11.11.1.1 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 details as described in the Table 43:

Table 43: Dataset Widgets and their Descriptions

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 43: Dataset Widgets and their Descriptions

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: 12. Click Add + next to the Enable DQ check option. 13. Under Master DQ, select one or multiple Primary Key options. All columns of the selected staging table are listed for you to select. 14. Under DQ Rules, select a column from the Available Columns list. This list contains all columns of the selected staging table. 15. 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.
Enable DQ check	16. Click Save to save your DQ rule. 17. 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.

6. You can perform certain common tasks in all the widgets, such as Edit, Delete, filter, etc. See Common Tasks for more information.

11.11.1.2 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 Table 44.

Table 44: Fields to Create Persist and their Descriptions

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 44: Fields to Create Persist and their Descriptions

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. See beginning at **Step 5** in the Creating Join section for information on Joining datasets.
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. Follow these steps in the **Map** pane:

   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 Common Tasks.

11.11.1.3 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.   Follow these steps in the **Output** pane:

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 Join Operators.

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 Common Tasks.

11.11.1.3.1 Join Operators

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.

11.11.1.4 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 Configuring Filters.

11.11.1.5 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 Common Tasks.

11.11.2   Creating a Pipeline

11.11.2.1   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 Table 45.

Table 45: Fields to Create a Pipeline and their Descriptions

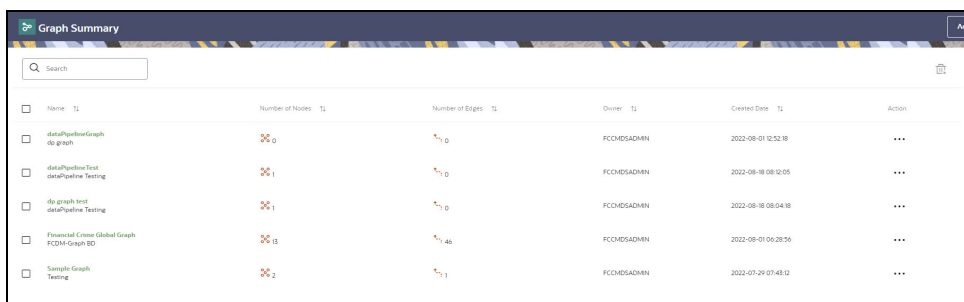| Field	Description
Name	Enter the name for the pipeline.
Description	Enter the description for the pipeline.
Add Search Tags	Enter the keywords for the pipeline. These keywords can be used as search tags while searching for a pipeline. Search tags are also used to group pipelines of the same type. These search tags appear as filters on the Pipeline page.
Type	Select the type of pipeline as either Scenario, Scoring, Data Loading, or Data.

4.   Click **Create**. A new pipeline is created and displayed on the Pipeline page. You can perform the required configurations in the newly created pipeline.

11.11.2.2   Cloning from an Existing Data Pipeline

To clone a pipeline, follow these steps:

1.   Navigate to **Workspace Summary** > **Graphs.** The Graph Summary page is displayed.

Figure 211: Graph Summary

2.   Click <**Graph name**>. On the <**Node/Edge Details**> page, click **Create** drop-down list and then click **Data Pipeline**. The Create Pipeline window is displayed.

Figure 212: Clone a Pipeline

3. Select **Clone From an Existing Data Pipeline** option and then select the required pipeline to be cloned from the **Select Pipeline** drop-down list.

4. Click **Ok** to clone the selected pipeline.

5. Navigate to the **Pipeline Designer** page. Enter the name in the Data Pipeline text box which has to be cloned.

6. Click **Save**. The new pipeline is created.

11.11.3 Common Tasks

11.11.3.1 Configuring Filters

You 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.

11.11.3.2 Defining Filter Conditions

You can define the filter conditions using one of the following:

- Expression Builder
- Tables
- Text

11.11.3.2.1 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.

11.11.3.2.2 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.

11.11.3.2.3 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.

11.11.3.3 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. For more information, see Defining Filter Condition Using Expression Builder.

3. Click **Add** . The New Runtime Parameter dialog box is displayed.

4. Provide the details as described in the Table 46:

Table 46: Fields on New Runtime Parameter and their Descriptions

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.

11.11.3.4 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 ![icon]. A dialog box is displayed.

4. Modify the required details.

5. Click ![icon] to save the changes. The widget details are modified.

11.11.3.5 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 ![icon].

4. The Pipeline Delete dialog box is displayed. Click Confirm. The widget is deleted.

11.11.4 External Service Widget

The ![icon] 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 Elastic Search if similar edges are found in the graph between two nodes. Currently, the loading to Elastic search is achieved by Data Pipeline – External Service and also this is used to load the data into Elastic Search Indexes.

Figure 213: External Service
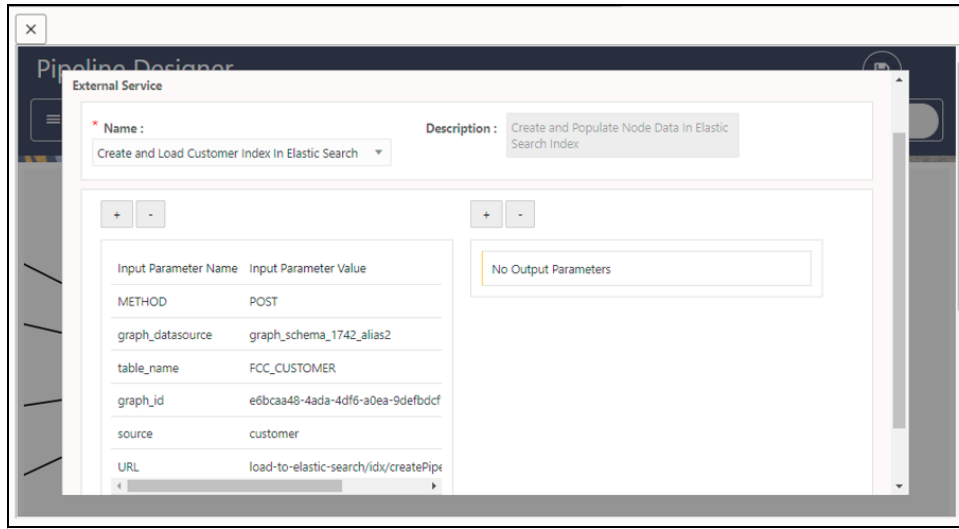
To configure an external service for Load-to-Elastic-Search, perform these steps:

1. Seed an entry for the Index in the **FCC_M_EXTERNALSERVICE** table.

    This table is present in the Compliance Studio Schema. A sample insert script is shown below, replace the placeholders and execute in the Compliance Studio Schema.

    Insert into FCC_M_EXTERNALSERVICE
    (N_COMPONENT_ID,V_STEP,V_TYPE,V_NAME,V_SERVICE_TYPE,F_STATUS,F_PRODUCT_PART,
    V_CODE,V_DESCRIPTION,V_SERVICE_NAME) values (<give unique number(Max of
    N_COMPONENT_ID + 1)>, '/idx/createPipelineIndexAndLoadData' , 'REST', 'Create and Load
    Index In Elastic Search - < Node Name >' , 'EESERVICE' , 'A' , 'C' ,
    'AmlLite_C_1592205797180_<n_component_id>' , 'Create and Populate Node Data In Elastic
    Search Index – < Node Name >' , 'load-to-elastic-search');

    Commit;

2. Drag an External Service widget from the menu. For more information, see Creating External Service.

3. Navigate to the Definition option from the widget and select the Name seeded in Step 1 (value for V_NAME).

    For example: Create and Load Index In Elastic Search - < Node Name >

4. Click + icon to add the input parameters.

Figure 214: External Service
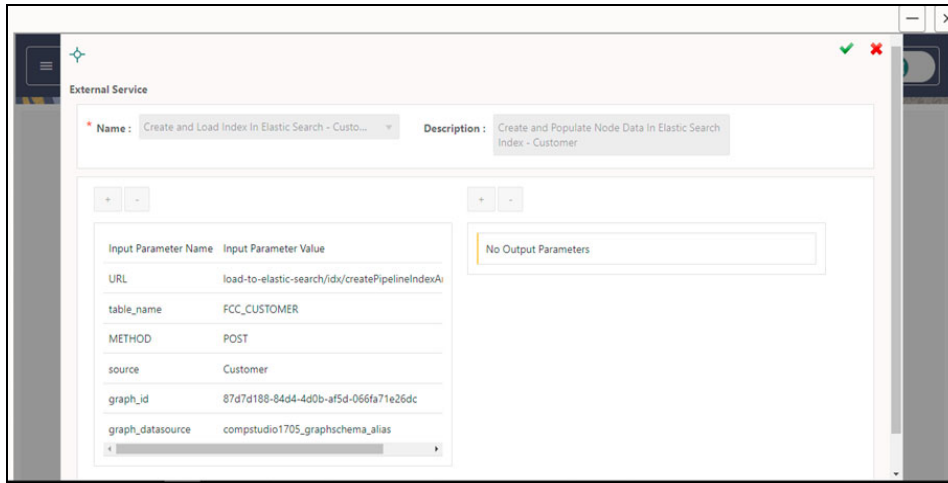
The Table 47 describes the parameters required for External Service.

Table 47: External Service Details

Name	Description	Value
METHOD	Call type to be Made from Data Pipeline Service to External Service	POST
URL	Target table name of the node, where elastic search index would populate data from.	load-to-elastic-search/idx/ createPipelineIndexAndLoadData
graph_datasource	Oracle wallet alias of a graph schema	
source	Source value for Node. Logical name of the Index.	
graph_id	Graph ID of the graph where the external service is part of.	

11.11.5 Defining an Index

An index is used to map the node attributes and load the data into Elastic Search Indexes.

NOTE	Attaching the External Service to the data pipeline before creating an index is mandatory only if you want to load the particular node on the elastic search.

To define an index, follow these steps:

1. Navigate to **Workspace Summary** > **Graphs**. The Graph Summary page is displayed.

Figure 215: Graph Summary

2. Click <**Graph name**>. On the Node Details page, click **Attach** drop-down list and then click **Data Pipeline**. The following page is displayed.

Figure 216: Node Details
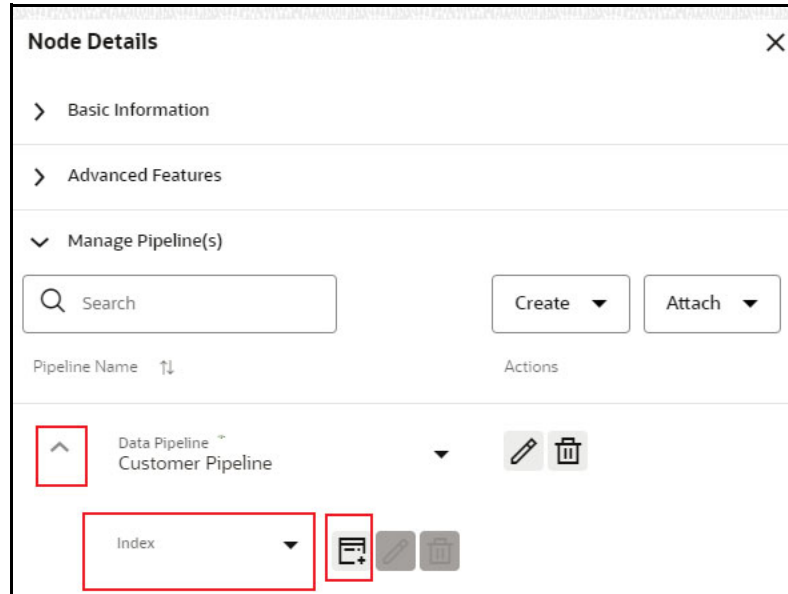

11.11.5.1   Creating an Index

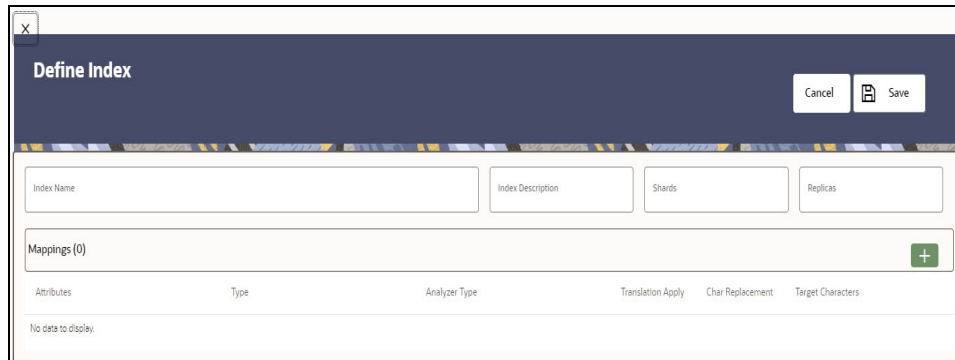To create an index, follow these steps:

1. Click **Show attached index** drop-down list to view the index. The following page is displayed.

Figure 217: Create an Index

2. Click **Create** button to add the index. The **Define Index** window is displayed.

**Figure 218: Define Index**

3. Enter the following fields:

   ▪ **Index Name**: Enter the unique index name across all the graphs. The field is limited to 30 characters.

   ▪ **Index Description**: Enter the description of the index. The field is limited to 255 characters.

   ▪ **Shards**: Enter the required integer value. Enter the value as 1.

   ▪ **Replicas:** Enter the required integer value. Enter the value as 3.

4. Click [+] icon to add mapping to the index. The **Mappings** section is displayed.

Figure 219: Mapping
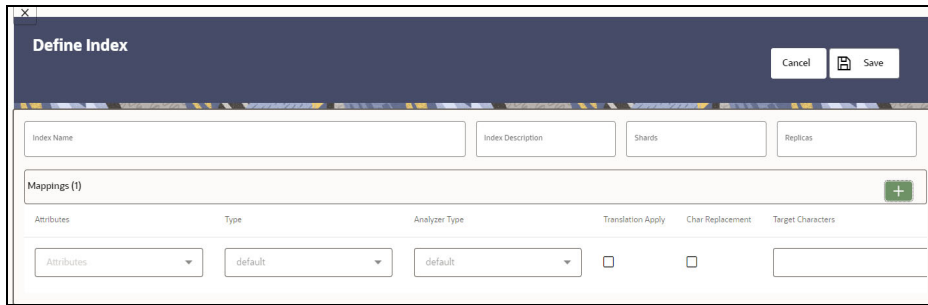
The Table 48 provides descriptions for the fields in the **Mappings** section.

Table 48: Mapping Details

| 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.
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** ● **BOOLEAN** ● **TEXT** **NOTE:** If you want to seed a new value in the Type, then enter the value as FCC_IDX_JSON_CONTENT in the metadata table.
Analyzer Type	It should be set as **default** for key attributes. For other attributes, you can select from the drop-down. The available options are: ● **ADDRESS** ● **ORGANIZATION** ● **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.
Translation Apply	Enable this option for language translation, if required.
Char Replacement	Enable this option for character replacement in the elastic search 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.

Table 48: Mapping Details

| Field	Description
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** button 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.

11.11.5.2 Editing an Index

To edit an index, perform these steps:

Figure 220: Edit an Index
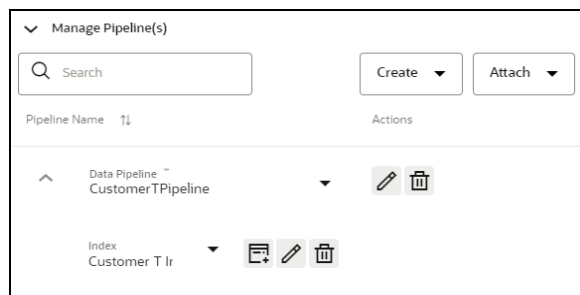
1. Click **Edit** icon to edit the required fields. The Define Index page is displayed.

Figure 221: Define Index with Edit Option
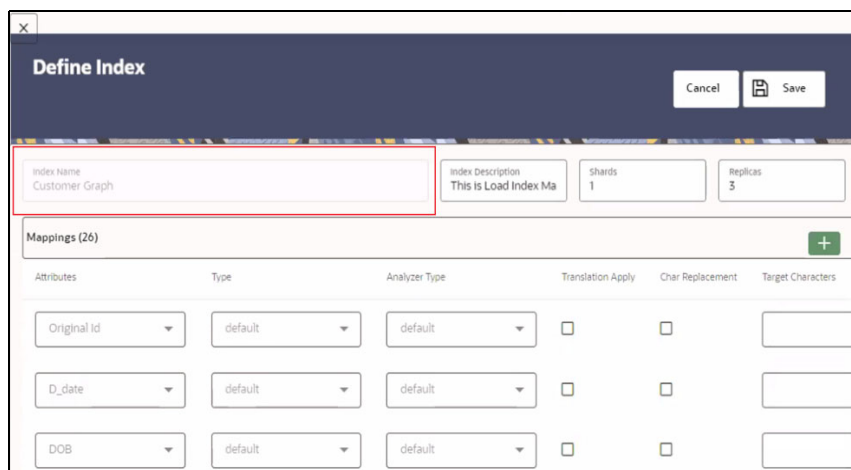
2. Edit the fields as required and click **Save**. The index is saved.

    NOTE: You cannot edit the Index Name.

11.11.5.3 Deleting an Index

1. Click **Delete** icon in the Manage Pipeline(s) group.

   The index is removed from the respective data pipeline.

12   Accessing Notebook Server Options

Topics:

- Configuring Interpreters
- Managing Tasks
- Managing Permissions
- Managing Credentials
- Configuring Graphs
- Configuring Templates

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

- Interpreters
- Tasks
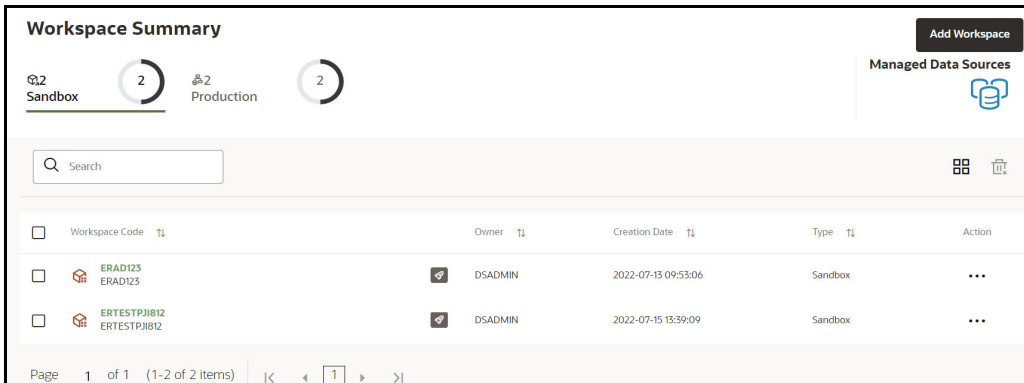- Permissions
- Credentials
- Templates

1. On the **Workspace Summary** page, select Launch workspace   to display the **CS Production workspace** window**.**

Figure 222:   Workspace Summary

2. Hover the mouse over the Data Studio Options widget   the following options are available:

- Interpreters
- Tasks
- Permissions
- Credentials
- Templates

12.1   Configuring 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. In Compliance Studio, 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.

Compliance Studio has ready-to-use interpreters such as fcc-python, jdbc Interpreter, etc. 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**   fcc-python, 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, follow these steps:

1. Click the **Interpreter** that you want to view from the list displayed on the LHS. The default configured interpreter variant is displayed on the RHS.

Figure 223: Interpreter's selection screen
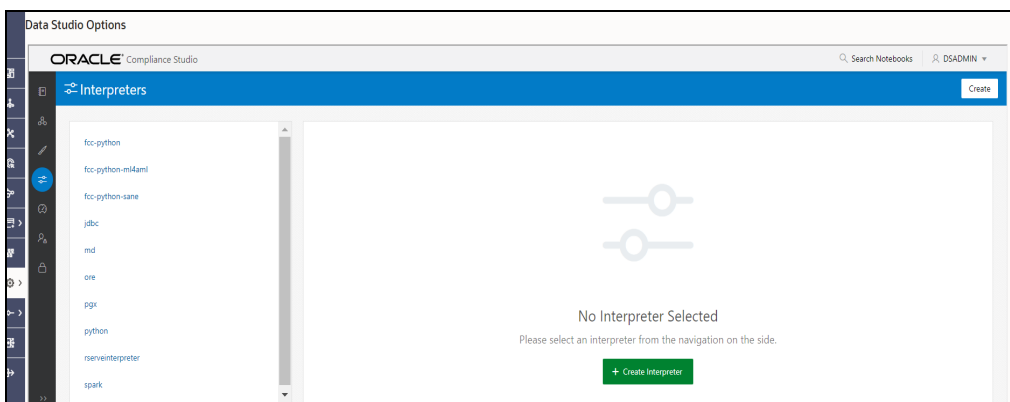
2. Modify the values in the fields as per requirement. For example, to modify a parameter's limit, connect to a different schema, PGX server, etc.

3. Click **Update**. The modified values are updated in the Interpreter.

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

12.2   Managing 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.

For more information on tasks, see the **Monitoring Tasks on Notebook Server** section in the OFS Compliance Studio Administration and Configuration Guide.

12.3   Managing Permissions

Use this section to 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.

12.4   Managing 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.

12.5   Configuring Graphs

The Notebook Server provides a pre-defined Financial crime graph that can be pre-loaded for most users and then states how to load different graphs. It also provides an intuitive way for creating graphs used in notebooks, where you can load graphs from external sources or create custom graphs. 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 to input other paragraphs in the notebook. The results of analytics algorithms are stored as the graph's transient properties of nodes and edges. Pattern matching can then be used against these properties.

The pre-defined graphs within the notebook can be accessed and used for the analysis. To access the pre-defined graphs within the notebook, select any out-of-the-box notebooks, make the necessary

changes for the PGQL code within the paragraph for the analysis you require, and run the paragraph. The query results are displayed as a graph.

For example, to view the Transfer of Money through hidden relationships, use the Financial Crime Graph Patterns notebook and execute the paragraph by providing the information about the entity, customer, and account.

Figure 224: Hidden Relationships graph
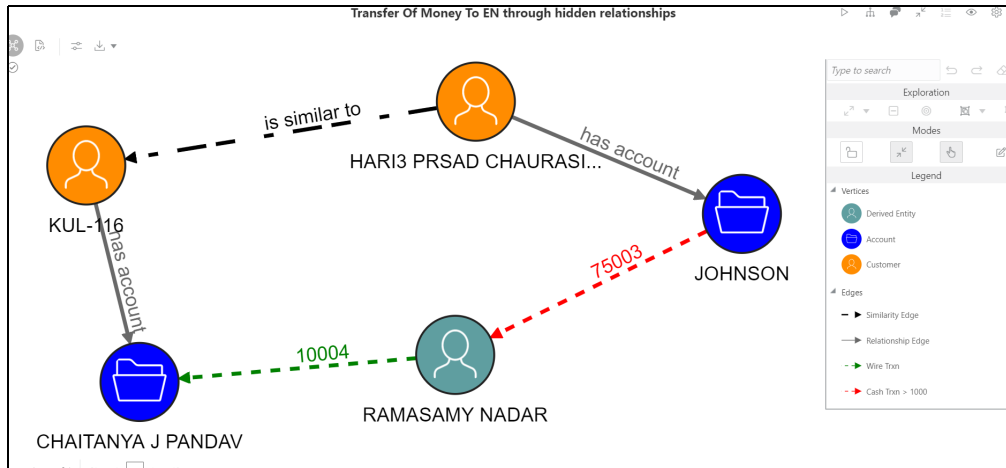
The graph configuration can be defined through a UI-based or JSON configurator. Graph configurations give you easy access to graphs using PGX-ALGORITHM, PGX-JAVA, and PGQL interpreters.

The graph can be created based on the following options:

- Creating Graphs by using the UI-Based Configurator
- Creating Graphs by using the JSON Configurator

12.5.1   Accessing the Graphs page

To create graphs based on the configurator, you need to access the Graphs page in the Notebook Server.

To access the Graphs page, follow these steps:

1. In the Notebook Server Workspace, click **Navigation Menu** in the upper-left corner. The menu items are listed.

2. Click **Graphs**. The Graphs page is displayed.

3. On the **Graphs** page, click **Create**.

    The **Create new Graph Configuration** dialog box is displayed.

4. In the dialog box, you can choose your **graph type**. The options available are **Homogenous** and **Heterogeneous**.

5. On the Create new Graph Configuration page, you can click:

    - **Wizard** to configure the graphs based on the UI configurator.

    - **Plain** to configure the graphs based on the JSON configurator.

12.5.2 Creating Graphs by using the UI-Based Configurator

To create the Graphs by using the **UI-Based Configurator**, follow these steps:

1. In the **Create new Graph Configuration** dialog box, click ✦ **Wizard** and specify the following information:

    - **Name**

      Define a unique graph name for this graph.

    - **Format**

      Notebook Server provides creating a graph from multiple formats, such as *Two tables*, *PGB*, *Edge list*, *Adjacency list*, *Flat file*, *GraphML*, *PG*, and *RDF*.

    - **Vertex Properties**

      Define the properties and types that your vertices must-have in the graph. For example, a Patient Vertex can have a **Name (string)** property.

    - **Edge Properties**

      Define the properties and types that your edges must-have in the graph. For example, a Patient_Medication Edge can have a **Dose (String)** property.

    - **Datastore**

      Choose the data source for the vertices and edges of the graph. Depending on the option, you must provide different information:

      — **File**

        Provide the Vertex URL List and the Edge URL List.

      — **Spark**

        Provide the Nodes and the Edges Table Name.

      — **RDBMS**

        Provide the username, password, the number of connections, Nodes, Edges Key Columns, and the From and To Node ID Columns.

      — **Advanced Mode**

        Click **Advanced Mode** to add more properties related to Vertex and Edge.

Figure 225: Create Graph Configuration

2. Click **Create**. The graph is created.

12.5.3   Creating Graphs by using the JSON Configurator

You can write a JSON configuration file based on the PGX JSON fields defined in the PGX graph configuration documentation to configure your graph.

To create the Graphs by using the **JSON Configurator**, follow these steps:

1. In the **Create new Graph Configuration** dialog box, click **Plain.**

2. Specify the name of your graph and the **JSON** code in the code field.

Figure 226: Add details in configuration
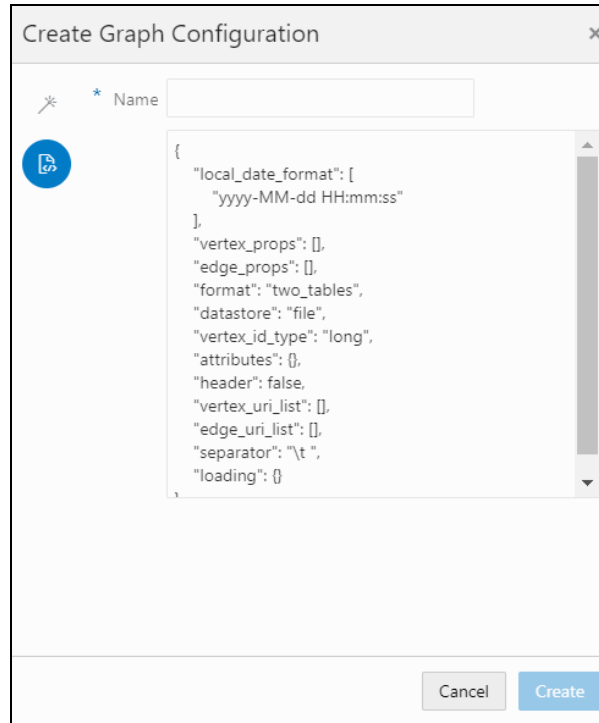
3.   Click **Create**. The graph is created. A sample code snippet is as follows:

```
{

"vertex_labels": false, "edge_label": true, "loading": {

"load_edge_label": true

},

"date_format": "yyyy-MM-dd HH:mm:ss", "vertex_props": [

{

"type": "string",

"name": "type"

},

{

"type": "string", "name": "religion"

},

{

"type": "string",

"name": "company"

},

{

"type": "string", "name": "musicGenre"
```

```
},
{
"type": "string",
"name": "show"
},
{
"type": "string",
"name": "name"
},
{
"type": "string",
"name": "country"
},
{
"type": "string",
"name": "team"
},
{
"type": "string",
"name": "genre"
},
{
"type": "string", "name": "occupation"
},
{
"type": "string",
"name": "role"
}
],
"edge_props": [
{
"type": "float",
"name": "weight"
}
],
"format": "edge_list", "vertex_id_type": "integer", "attributes": {},
```

```
"header": "",

"uri": "http://pgx.us.oracle.com/graphs/connections.edge_list",
"separator": "\t"

}
```

This newly created graph configuration can be used in a notebook for the following:

- Load a graph using the PGX interpreter.

  To load a graph, create a notebook and a paragraph with a PGX interpreter. Enter the code format given in the following example:

  ```
  graph = session.readGraphWithProperties(dataSourceName, 'graphName')
  ```

  Here, `dataSourceName` refers to the graph name that you have created.

- Query a graph using the PGQL interpreter.

  To query a graph, create a notebook and create a paragraph with a PGQL interpreter. Specify the code format given in the following example:

  ```
  "SELECT n,e,m FROM GRAPH_NAME MATCH (n) -[e]-> (m)"
  ```

Here `GRAPH_NAME` refers to the graph name that you have created.

12.5.4 Viewing the Details in a Graph

The graphs created and executed can be repositioned, collapsed, expanded, and then edges and vertex can be customized.

12.5.4.1 Repositioning Nodes

The Network Graph page allows you to move nodes around the screen, using the drag and drop feature, to reposition them.

To reposition nodes, follow these steps:

1. Navigate to the **Network Graph** in a notebook.

2. Select a node to reposition and click it.

3. Drag and drop the node to the required position.

> **NOTE**   The graph only uses a specific portion of the browser window to display the graph. Dragging a node beyond a certain point towards the right side of the browser hides the portion of the graph dragged beyond that point. However, you can use the Zoom Out feature on the Graph Toolbar to view the hidden portion again.

12.5.4.2 Dropping and Expanding Nodes

You can drop the nodes to hide all the outgoing links and nodes to which these outgoing links are connected from the collapsed node. You also similarly expand the nodes. The dropped node can be regenerated again by executing the graph. To drop a node, follow these steps.

1. Navigate to the **Network Graph** in a notebook.

2. Select a node to collapse and right-click the node. An option menu is displayed.

3. Select the **Drop** option from the menu. The node and its outgoing links are dropped.

NOTE	If any child node has at least one incoming link from any other node, the child node and its child network are not dropped. But the link from the dropped node to the child node is hidden, and the icon of the dropped node changes to indicate that the node is in a dropped state.
	On the Node menu of a dropped note, the drop option changes to Expand. The **Node** menu remains unchanged if the user drops a node but has no impact on the graph (if no part of the graph is hidden). There is no restriction on how many nodes can be dropped on a graph.

4. To expand the node, select **Expand** from the menu. The outgoing links are then restored on the page.

12.5.4.3 Viewing the Node Details

You can view the current information associated with the selected node. This is the same information displayed in this entity's **Entity Summary Historical Report** paragraph.

To view the node details, follow these steps:

1. Navigate to the **Network Graph** in a notebook.

2. Select a node and right-click. An option menu is displayed. The **Node Details** window is displayed with the current information associated with the selected node. This includes the Properties and Risk details of the node.

12.5.4.4 Deleting a Node

If you want to delete a node from a graph that is not required for your analysis, you can remove the unwanted node and view it. To delete a node, follow these steps:

1. Navigate to the **Network Graph** in a notebook.

2. Right-click on any node and click **Drop**.

12.5.5 Removing an Edge

You can remove an edge from a network graph to view the result in On-screen data. To remove an edge, follow these steps:

1. Navigate to the **Network Graph** in a notebook.

2. Right-click on any edge and click **Drop - Delete** selected vertices.

Figure 227: Removing the Edges
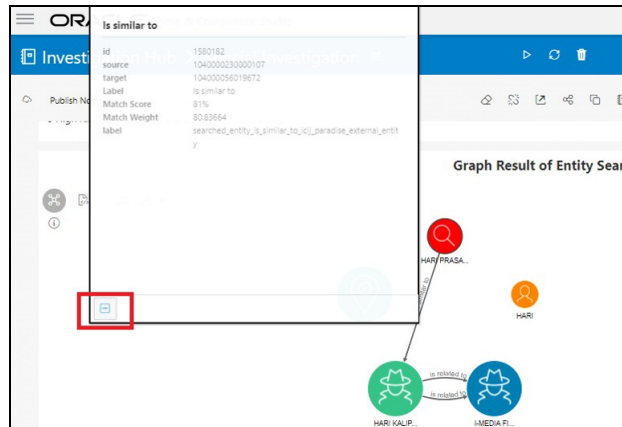
12.5.6   Customizing Graphs

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 **General**, **Visualization**, **Highlights**, and **Smart Explorer**.

The graph settings can be accessed by clicking ⊟ **Settings** in the paragraph. You can specify the properties in the Settings window to customize your graphs.

Figure 228: Graph Customization and Settings
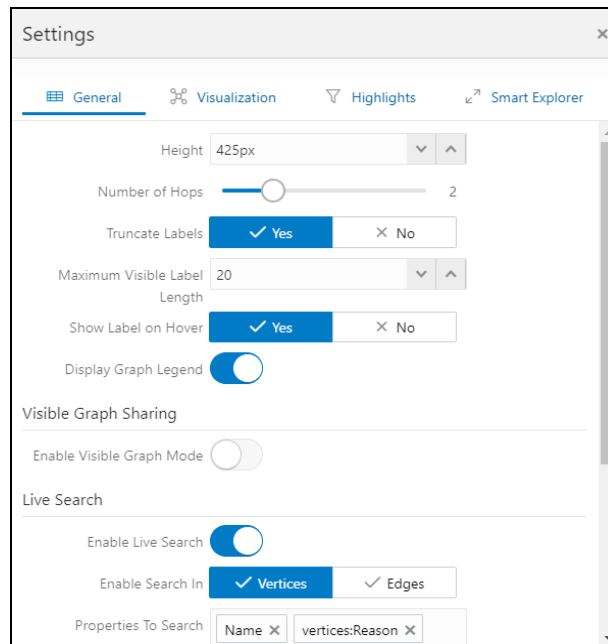
12.6   Configuring Templates

Perform the following steps to Create, Import, and set the FCGM Default **Template in the Template Dashboard**.

1. Click **CS** Launch Workspace to display the **CS Production Workspace** window.

2. Click Templates from Design Studio Options.

Figure 229: Templates
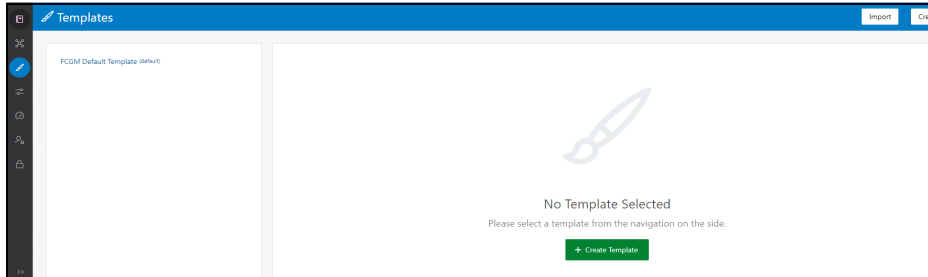
The Compliance Studio offers different formats to view the result after a paragraph's execution. Templates enable you to define parameters to customize the result formats. You can customize the visualization of the result by defining parameters in a template and then applying that template to a notebook.

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.

13   Appendix

Topics:

- Common Screen Elements in a Notebook
- Managing Resource Utilization
- Using Data Visualizations
- Dynamic Forms
- Quantifind Risk Report
- Using Scenario Conversion Utility

13.1   Common Screen Elements in a Notebook

The Table 49 describes the common screen elements in a notebook that can be used to perform quick actions when preparing and executing the notebooks.

Table 49: Common Screen Elements in a Notebook

Icon	Action or 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.
	Re-initialize session: Re-initialize the paragraph in a notebook. 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.

Table 49: Common Screen Elements in a Notebook
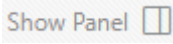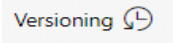
▷	Run Paragraphs: Executes all the paragraphs in a notebook in sequential order. For more information, see Running All Paragraphs.
↻	Reset Session: Resets any connection or code executed in a notebook.
🗑	Delete Notebook: Deletes a notebook.
	Clear Result: Clears results for all the paragraphs in a notebook. **NOTE**: This action clears all the results. You must rerun the paragraphs to view the results.
	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.
⧉	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>.
💾	Save Notebook: Saves the notebook to your computer as a DNSB file. For more information, see Exporting Individual Notebooks.
Layout 📖	Layout: Sets the preferred layout: Zeppelin or Jupyter.
Default Template 🖌	Default Template: Applies the overall look and feel of the notebook using the default template.
Default ▦	Default View: Switches between Default, Simple, and Report views.

Table 49: Common Screen Elements in a Notebook

|	
Show Panel ☐☐	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.
Versioning ⏲	Versioning: You can use this option to create versions for your notebook, which helps you analyze the changes based on the version control.

13.2   Managing Resource Utilization

If any of the Notebooks are no longer needed, you can invalidate the session of that notebook and free up the system resources.

To invalidate a session, perform the following steps:

1. Login to the CS application.

2. Navigate to **Model Pipelines** > **<Model lists>** > **Notebook**. The notebook is displayed.

    For example, **Model Pipelines** > **MLNameIncrementalTraining** > **Notebook.**

    Figure 230: Name Incremental Training Notebook
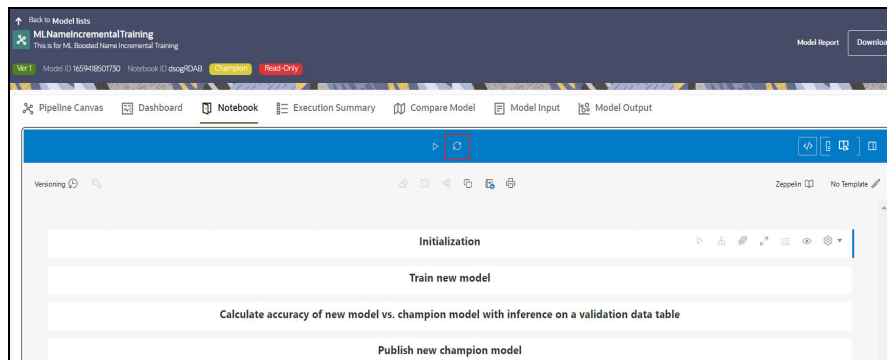
3. Click ⟳ to invalidate the session.

13.3   Using Data Visualizations

The following are the data visualizations that are available in the Notebook Server:

● Graph Visualization

● Table

● HTML or Markdown

● Text

● Area Chart

● Bar Chart

- Funnel Chart

- Line Chart

- Pie Chart

- Pyramid Chart

- TreeMap Diagram

- Sunburst Diagram

- Tag Cloud

- Box Plot

- Scatter Plot

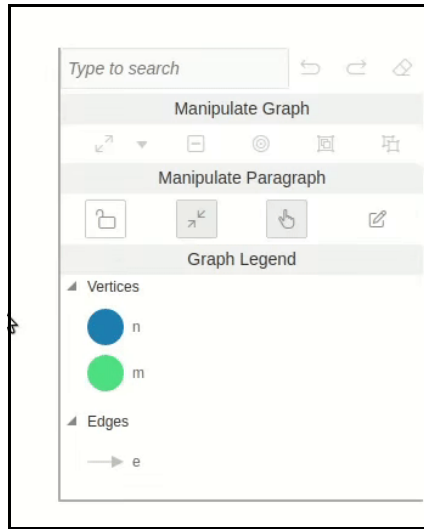- Map Visualize

13.3.1 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 231: Graph Visualization
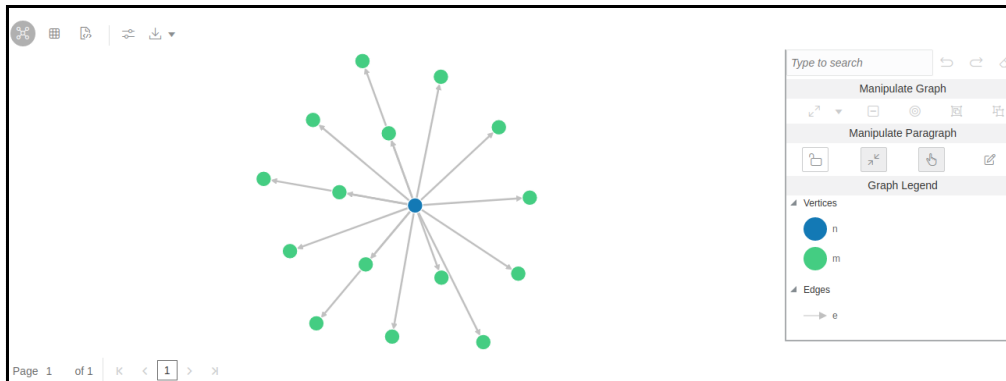
The panel can be resized and includes two states, default and minimized panel. These panels enable you to manipulate paragraphs and graphs and provide options to customize your graph.

Figure 232: Minimized Panel

13.3.2   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.
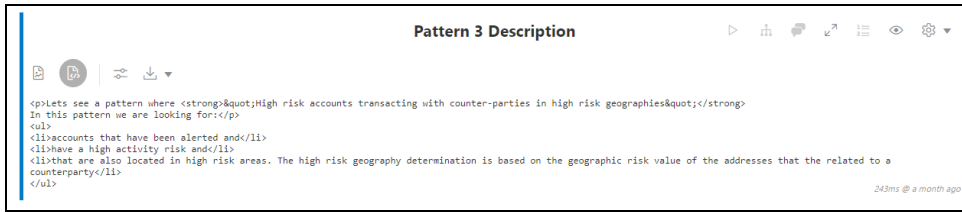
Figure 233: Table Diagram

13.3.3   HTML or Markdown

The Notebook Server allows you to visualize your data in the form of HTML or Markdown text.
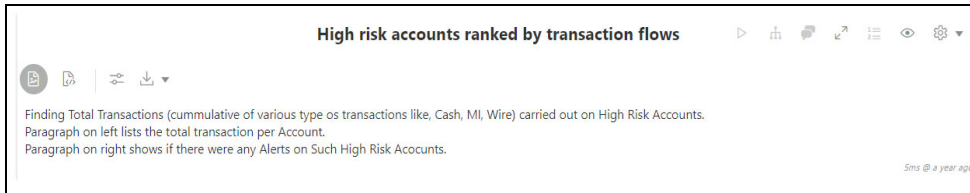
Figure 234: HTML Text

13.3.4 Text

The Notebook Server allows you to visualize your data in structured and presented text for you to customize the data.
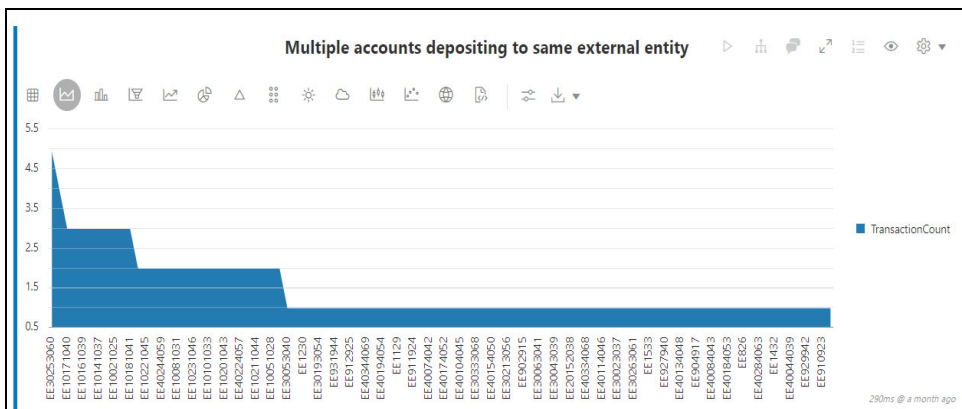
Figure 235: Text

13.3.5 Area Chart

The Notebook Server allows you to visualize your data in the form of an Area Chart.
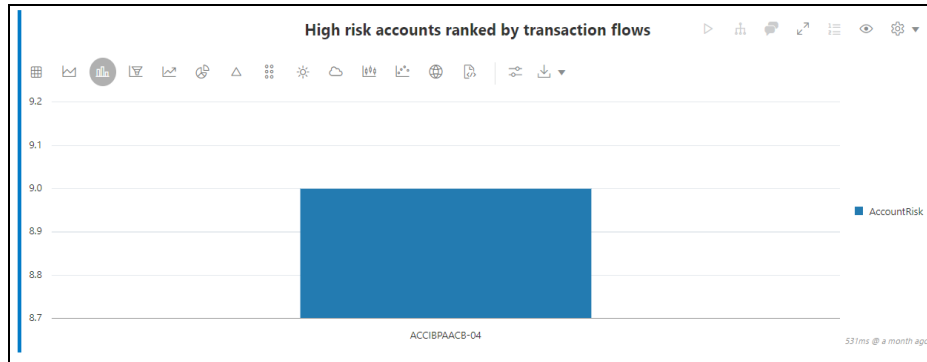
Figure 236: Area Chart

13.3.6 Bar Chart

The Notebook Server allows you to visualize your data in the form of a Bar Chart.
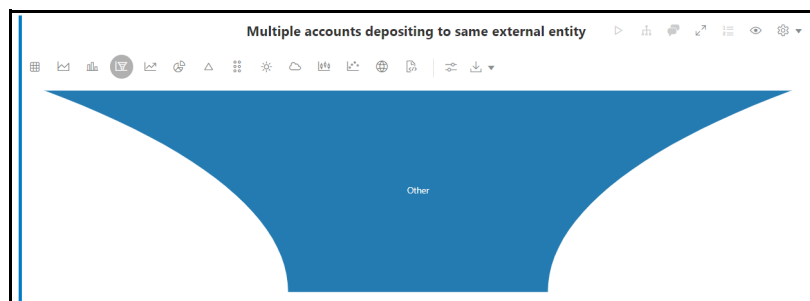
Figure 237: Bar Chart

13.3.7   Funnel Chart

The Notebook Server allows you to visualize your data in the form of a Funnel Chart.
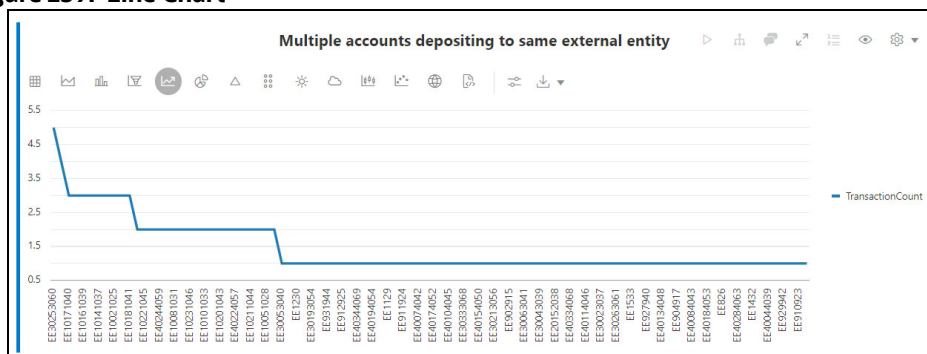
Figure 238: Funnel Chart

13.3.8   Line Chart

The Notebook Server allows you to visualize your data in the form of a Line Chart.
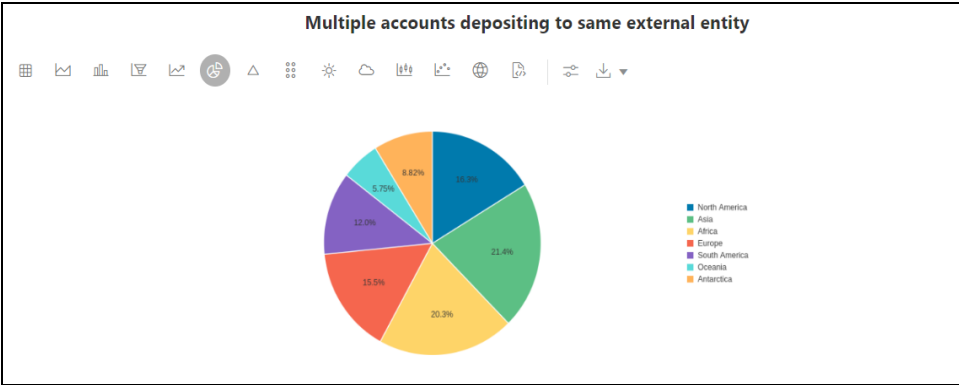
Figure 239: Line Chart

13.3.9   Pie Chart

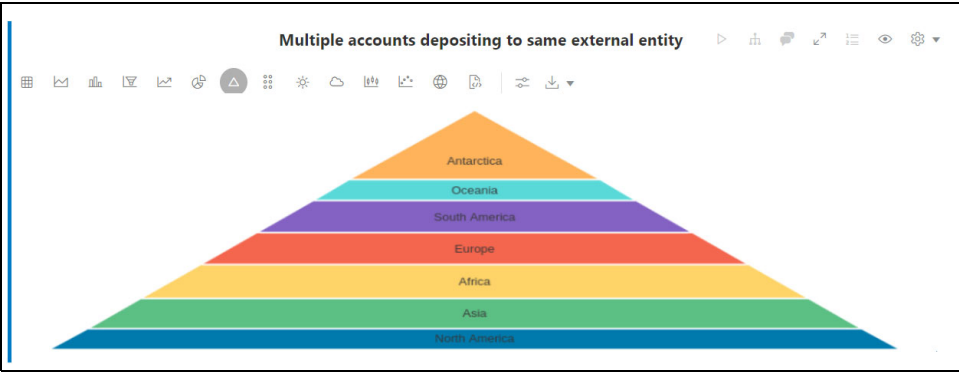The Notebook Server allows you to visualize your data in the form of a Pie Chart.

Figure 240: Pie Chart

13.3.10 Pyramid Chart

The Notebook Server allows you to visualize your data in the form of a Pyramid Chart.
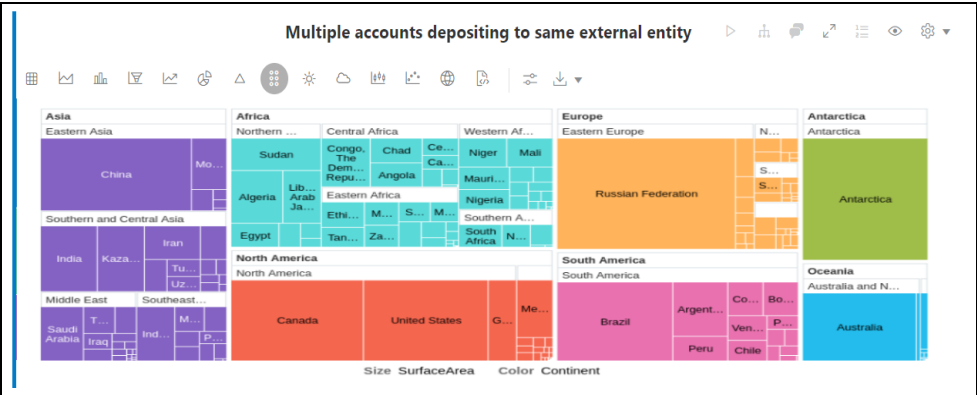
Figure 241: Pyramid Chart

13.3.11 TreeMap Diagram

The Notebook Server allows you to visualize your data in the form of a Tree Map Diagram.

Figure 242: TreeMap Diagram

13.3.12   Sunburst Diagram

The Notebook Server allows you to visualize your data in the form of a Sunburst Diagram.

Figure 243:   Sunburst Diagram
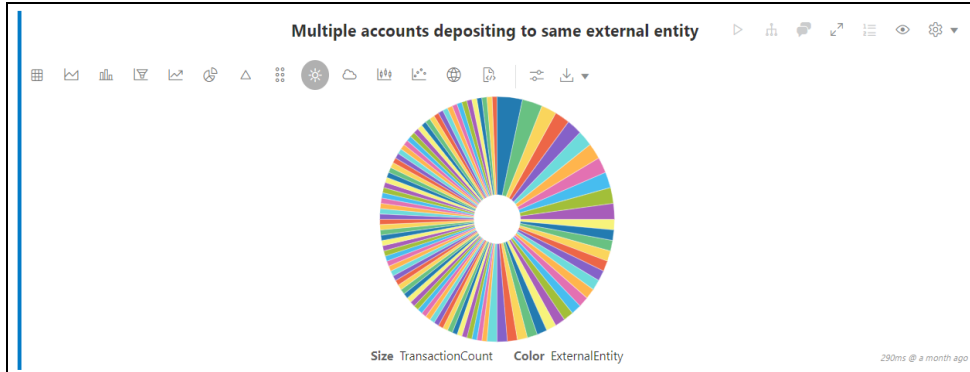
13.3.13   Tag Cloud

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 244: Tag Cloud
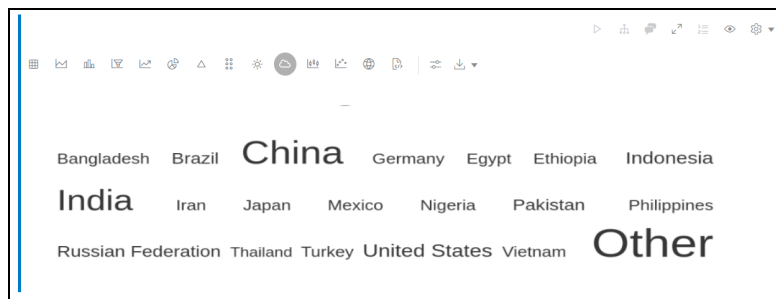
13.3.14   Box Plot

The Notebook Server allows you to visualize your data in the form of a Box Plot.

Figure 245: Box Plot
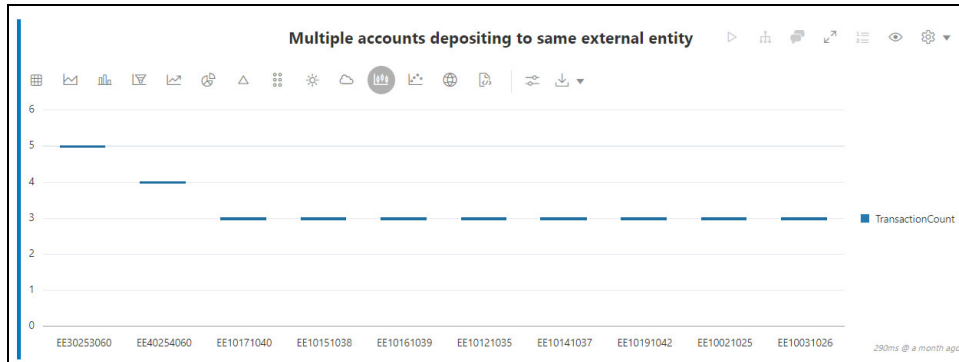
13.3.15 Scatter Plot

The Notebook Server allows you to visualize your data in the form of a Scatter Plot.

Figure 246: Scatter Plot

13.3.16 Map Visualizer

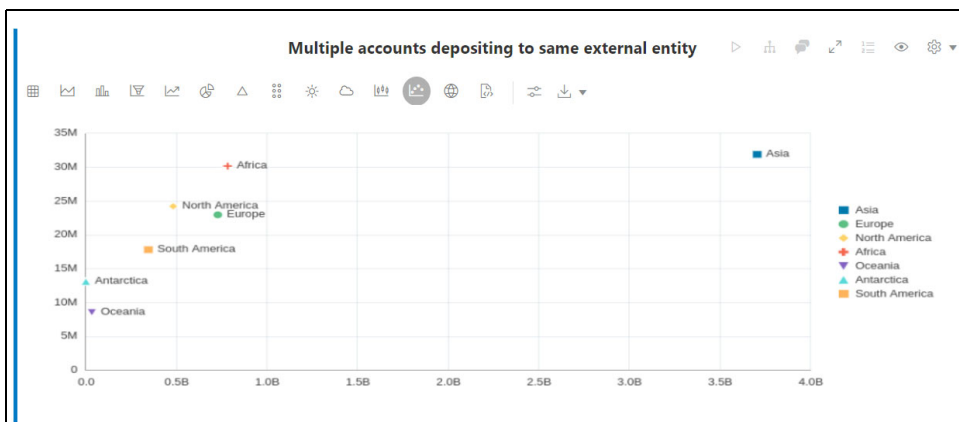The Notebook Server allows you to visualize your data on top of a Map.

Figure 247: Map Visualizer
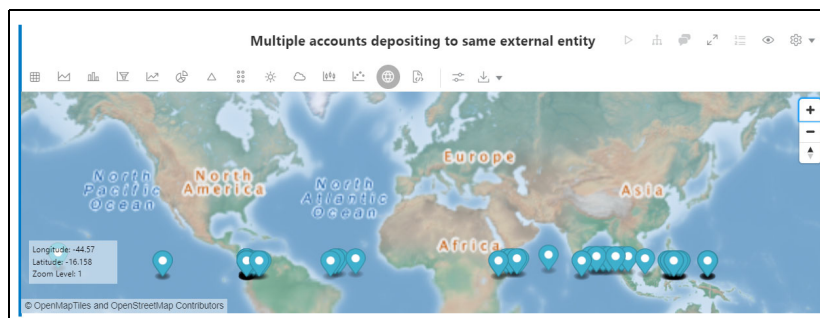
13.3.16.1 Setting up the Visualizations

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

Figure 248: Settings Window
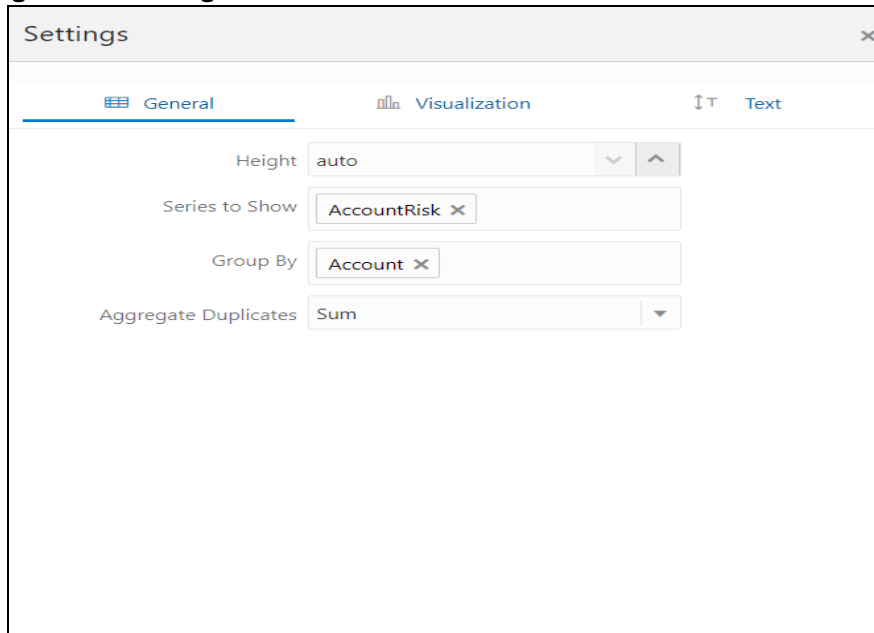
13.3.16.2 Download As

The displayed visuals can be downloaded using the ⬚ **Download As** option. The available download formats are RAW, SVG, and PNG.

Figure 249: Download As an option
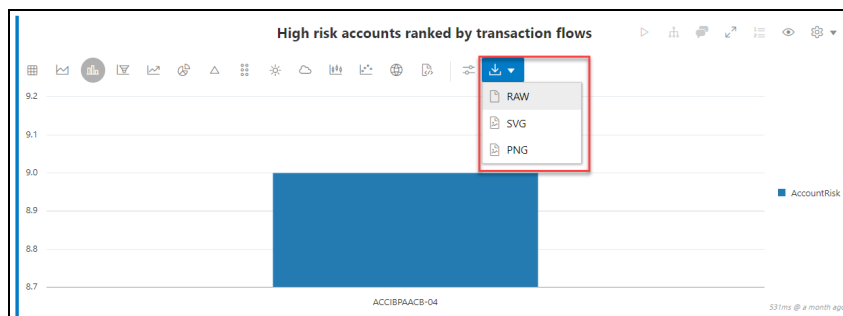
1. Navigate to the **Workspace Summary** page.

2. Navigate to **Model Pipelines** > <**Model**> > <**Object**>. The following Page is displayed.

Figure 250: Pipeline Designer
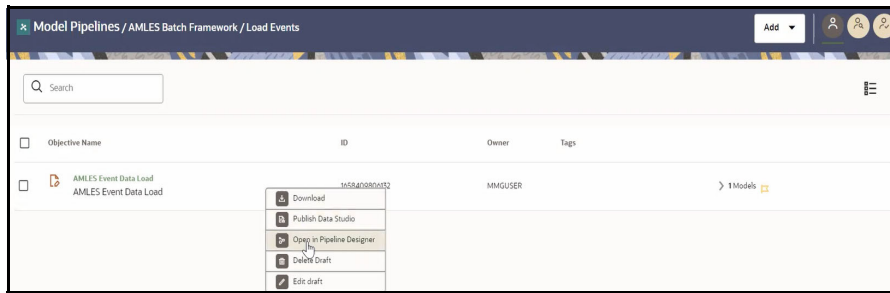
3. Click the ⊙ icon next to <Objective Name> to view the list of options and then select the **Open in Pipeline Designer.** The following page is displayed.

Figure 251: Notebook
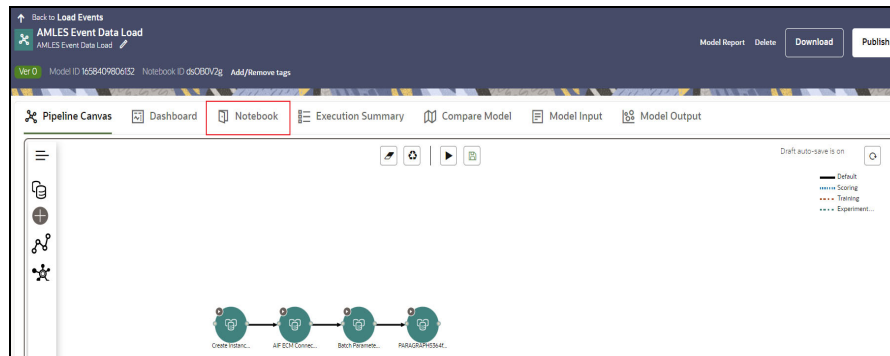
4. Click on the **Notebook** tab to view the invalidate session button.

5. Click the **Invalidate Session** [icon] button to invalidate the session and free up resources.

Figure 252: Invalidate Session button
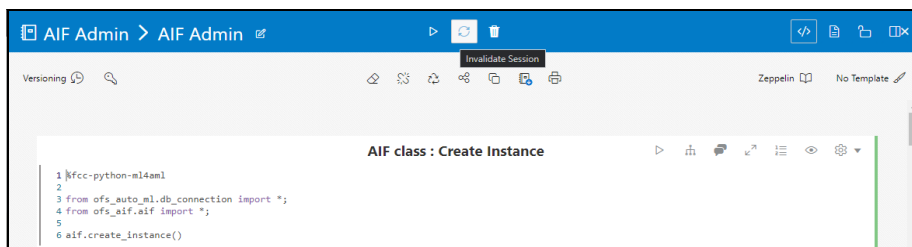
13.4   Dynamic Forms

The following are the supported input fields in Notebook Server:

- Textbox
- Select
- Slider

- Checkbox
- Date Picker
- Time Picker
- DateTime Picker
- Setting Values through a Notebook URL

The following sections in this chapter provide details about how you can specify the dynamic parameters for the existing fields.

13.4.1   Textbox

The textbox offers a simple user input that accepts any characters.

- 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}`

Figure 253: Textbox

```
%md

My name is: ${name(Label)=Default Value}

Label

Default Value

My name is: Default Value
```
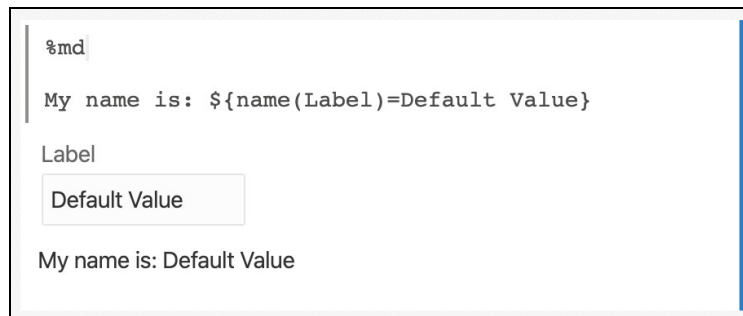
13.4.2   Select

A select field allows you to choose from a given selection.

- 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}`

```
%md

My name is: ${name(Label)=default,default(Default Option)|other1(Other Option 1)|other2}

Label

Default Option                    ▼

My name is: default
```

13.4.3 Slider

A slider allows you to choose from a given range.

- 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}`

Figure 255: Slider

```
%md

My age is: ${slider(0,30,1.5):name(Label)=3}

Label

0  ──◯────────────────────  30

My age is: 3
```
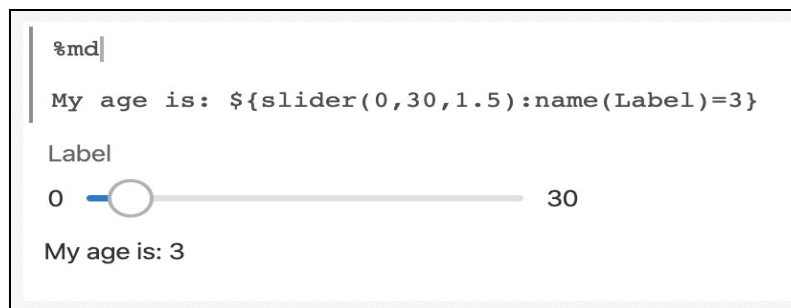
13.4.4 Checkbox

A checkbox allows you to choose from a given selection.

- 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}`

Figure 256: Checkbox

```
%md
My favorites are: ${checkbox( or ):name(Label)=default,default(Default Option)|other1(Other Option 1)|other2}
Label
☑ Default Option   ☑ Other Option 1
☐ other2
My favorites are: default or other1
```

13.4.5   Date Picker

A date-picker allows you to choose from a given date.

- 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}`

Figure 257: Date Picker

```
%md
My birthday is on ${date(yyyy-MM-dd):name(Label)=2019-02-27}
Label
2019-02-27              🗓
My birthday is on 2019-02-27
```

13.4.6   Time Picker

A time picker allows you to choose from a given time.

- 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}`.

Figure 258: Time PickerDateTime Picker

```
%md
Lunch will be served at ${time(HH:mm):name(Label)=T14:30}
Label
14:30              🕐
Lunch will be served at 14:30
```

13.4.7   DateTime Picker

A datetime picker allows you to choose from a given date and time.

- 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}`.

Figure 259: DateTime Picker

13.4.8   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.

13.5   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 figure shows that you can run the risk report query in the ECM notebook using the java interpreter.

Figure 260: Run Risk Report query using java interpreter
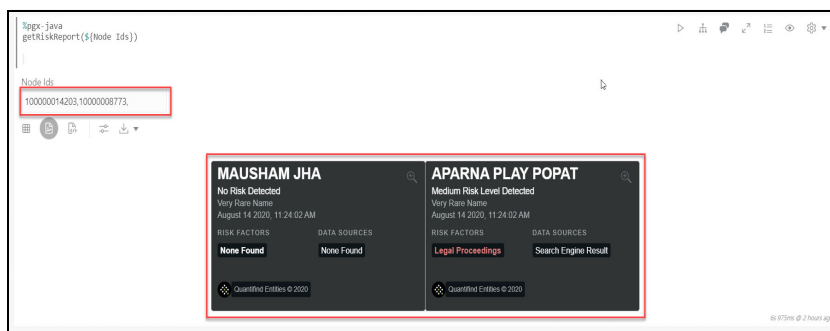
The results are displayed as a card, which displays the risk status of the identified node details. Based on the risk, you can perform the required measures for the risk analysis.

13.6   Using Scenario Conversion Utility

13.6.1   Introduction

Scenario Conversion Utility converts BD scenarios to Compliance Studio notebooks for the scenarios and respective thresholds.

> **NOTE**   The utility converts only BD AML scenarios to Compliance Studio notebooks based on the rule matcher algorithm.

13.6.2   Prerequisites

- Ensure that the OFS Compliance Studio v8.1.2.3.0 is installed and running.
- Ensure that the OFS BD v8.1.2.2.0 is installed and running.
- Ensure that the following tables and sequence exist in the Compliance Studio Schema:
    - Tables:
        — ds_notebook
        — ds_paragraph
    - Sequence:
        — seq_paragraph

13.6.3   Conversion Steps

The following terms are used in the scenario:

- AC – Account
- CU – Customer
- ML – Money Laundering
- ML/AC – Money Laundering Account focus Scenario
- ML/CU – Money Laundering Customer focus Scenario

The following scenarios are supported:

- (ML/AC) Transactions in Round Amounts
- (ML/CU) Early Payoff or Paydown of a Credit Product
- (ML/AC) CIB: High Risk Geography Activity
- (ML/AC) Anticipatory Profile - Expected Activity
- (ML/AC) CIB: Significant Change from Previous Peak Activity
- (ML/AC) CIB: Significant Change from Previous Average Activity
- (ML/AC) CIB: Product Utilization Shift
- (ML/CU) Anomalies in ATM, Bank Card: Foreign Transactions
- (ML/CU) Anomalies in ATM, Bank Card: Excessive Withdrawals
- (ML/CU) Rapid Movement of Funds - All Activity

- (ML/AC) Rapid Movement of Funds - All Activity
- (ML/CU) Single or Multiple Cash Transactions: Large Significant Transactions
- (ML/AC) Deposits / Withdrawals in Same or Similar Amounts
- (ML/AC) Patterns of Funds Transfers Between Customers and External Entities
- (IML/CU) Frequent Changes to Instructions
- (IML/CU) High Risk Electronic Transfers
- (IML/EN) High Risk Electronic Transfers
- (ML/AC) Patterns of Funds Transfers Between Internal Accounts and Customers
- (ML/CU) High Risk Transactions: Focal High Risk Entity
- (ML/AC) High Risk Transactions: Focal High Risk Entity
- (ML/AC) High Risk Transactions: High Risk Counter Party
- (ML/HH) High Risk Transactions - High Risk Counter Party
- (ML/AC) High Risk Transactions: High Risk Geography
- (ML/CU) Single or Multiple Cash Transactions: Possible CTR
- (ML/HH) Single or Multiple Cash Transactions: Possible CTR
- (ML/AC) Anomalies in ATM, Bank Card: Excessive Withdrawals
- (ML/AC) Deposits/Withdrawals in Same or Similar Amounts
- (ML/EN) Hub and Spoke
- (ML/AC) Terrorist Financing
- (ML/EN) Terrorist Financing
- (ML/EN) Transactions in Round Amounts (MI)
- (ML/CU) Large Reportable Transactions
- (ML/CU) Patterns of Funds Transfers Between Customers and External Entities

> **NOTE**   Other BD scenarios are under the verification process.

To convert the BD AML scenario, perform the following:
- Setting up Database
- Importing the Notebook (Draft)
- Generating Threshold and Scenario Notebook
- Running the Scenario

13.6.3.1   Setting up Database

1. Log in to the server where Compliance Studio is installed.
2. Navigate to

   ```
   <OFS_COMPLIANCE_STUDIO_INSTALLATION_PATH>/deployed/ficdb/Scenario-
   Conversion-Utility.
   ```

3. Download the `BDtoStudioScenarioMigrationUtility.dsnb` Notebook file from the above path and save it to the local folder.

4. Copy and unzip the `BD2S_Utility.zip` file to the local machine.

5. Navigate to the **BD** folder inside the unzipped folder.

6. Open the **ATOMIC** folder inside the **BD** folder. The following folders are available within the **ATOMIC** folder in sequential order:

   a. DB Link

   b. Synonyms

   c. sequences

   d. Tables

   e. Views

   f. Package

Open all the individual folders and execute the SQL scripts in each folder in BD atomic schema in the following sequence:

> **NOTE**
>
> Before running the first script, the following place holders inside the script must be modified.
>
> - HOST IP address of database where Studio schema is available
> - Service Name of database
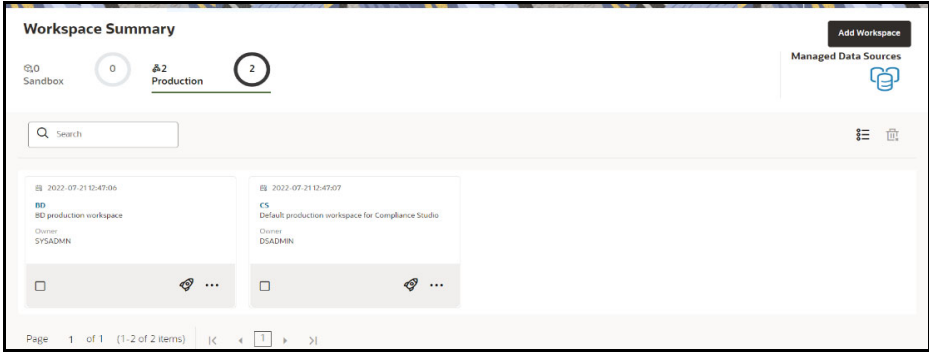> - Username for Studio schema
> - Password for Studio schema

   i.   `@<Fully Qualified path of scenario conversion utility>/db_link.sql`

   ii. `@<Fully Qualified path of scenario conversion utility>/synonym.sql`

   iii. `@<Fully Qualified path of scenario conversion utility>/Sequence.sql`

   iv. `@<Fully Qualified path of scenario conversion utility>/Table.sql`

   v.   `@<Fully Qualified path of scenario conversion utility>/View.sql`

   vi. `@<Fully Qualified path of scenario conversion utility>/pkg_fcc_b-d_nb.sql`

13.6.3.2 Importing the Notebook (Draft)

To import the Notebook, follow these steps:

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

Figure 261: Workspace Summary

2. Navigate to any Sandbox workspace where BD atomic schema is the data source.

> **NOTE** Ensure that the BD atomic schema is added as a data source for the workspace.

3. Create an **Objective**. To create an Objective, see the Creating Objective (Folders) section.

4. Import the following Notebook as a Draft. To import, see the Import a Workspace Model Data into a New Model section.

   `BDtoStudioScenarioMigrationUtility.dsnb`

   A new model is created by importing the model data dump of another model.

5. Open the Model in **Pipeline Designer**.

6. Click the **Notebook** tab, and the following page is displayed.
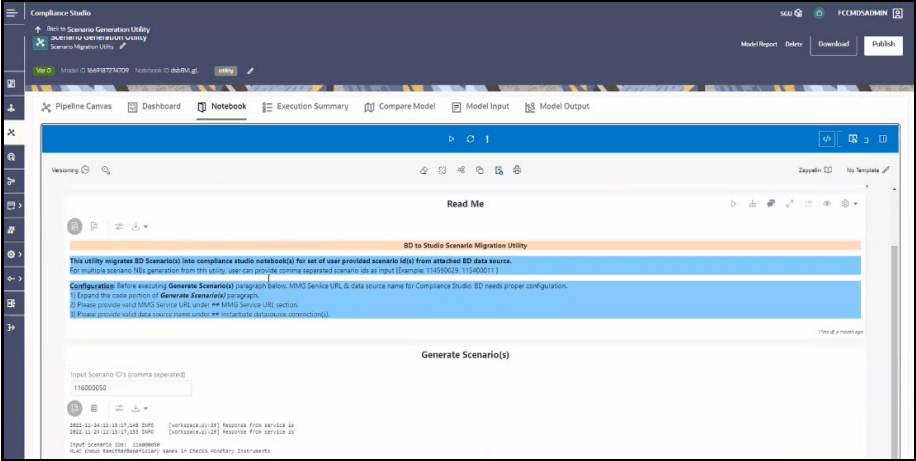
Figure 262: Notebook Tab
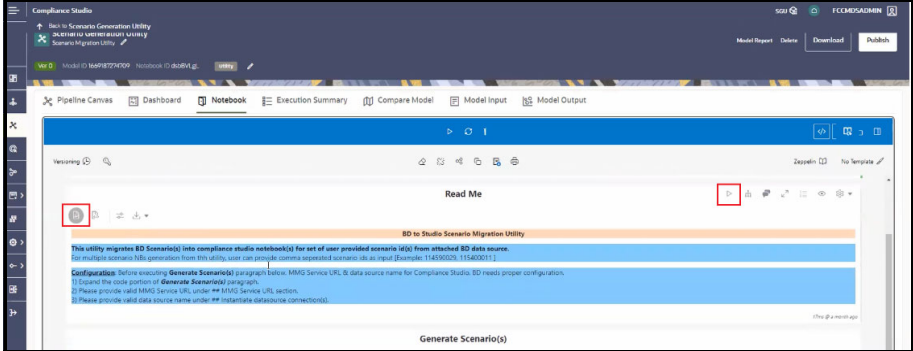
7. Click the **Run Paragraph** icon in the **Read Me** paragraph to open further instructions.

13.6.3.3 Generating Threshold and Scenario Notebook

To generate a threshold and Scenario Notebook, follow these steps:

1. Navigate to **Notebook** tab, and the following page is displayed.

Figure 263: Generate Scenario(s)

2. In the **Generate Scenario(s)** paragraph, click the **Visibility** icon and select the **Code** option.

3. Replace the Hostname for `obj_url` and `url` variables with your server Hostname under `##MMG Service URL` section in the **Generate Scenario(s)** paragraph.

4. Provide valid data source name for `cs_datasource` and `bd_datasource` variables under `## Instantiate datasource connection(s)` section in the **Generate Scenario(s)** paragraph.

Figure 264: Editing MMG Server URL and Data Source Name
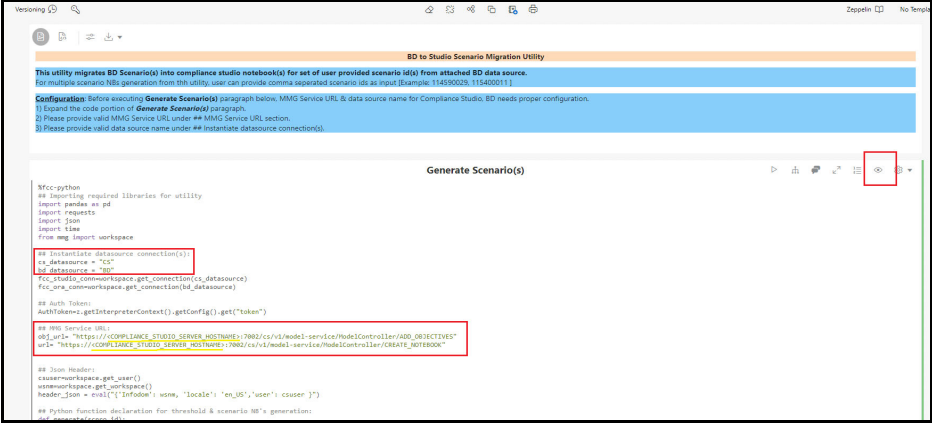
5. Enter the required Scenario ID(s) in the **Input Scenario ID's (comma-separated)** text box.

   You can enter multiple IDs with commas separated.

6. Click **Run Paragraph** to generate threshold and Scenario Notebook. Once it is executed successfully, you can view the success message and scenario notebook details in **Generate Scenario(s)** paragraph.
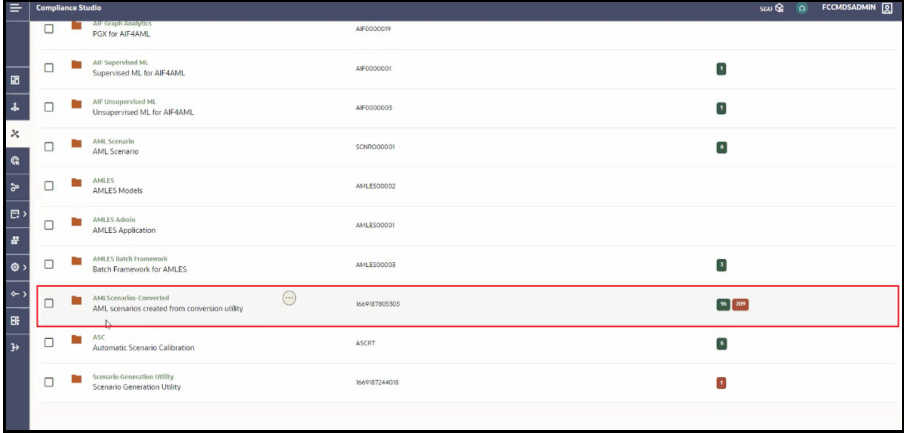
NOTE	Threshold and Scenario Notebook IDs are displayed under the `Threshold NB Name` and `Scenario NB name` values within the success message in **Generate Scenario(s)** paragraph. The Scenarios are generated inside the **AML Scenarios-Converted** folder with the above Notebook ID.

Figure 265: Successful Output Message
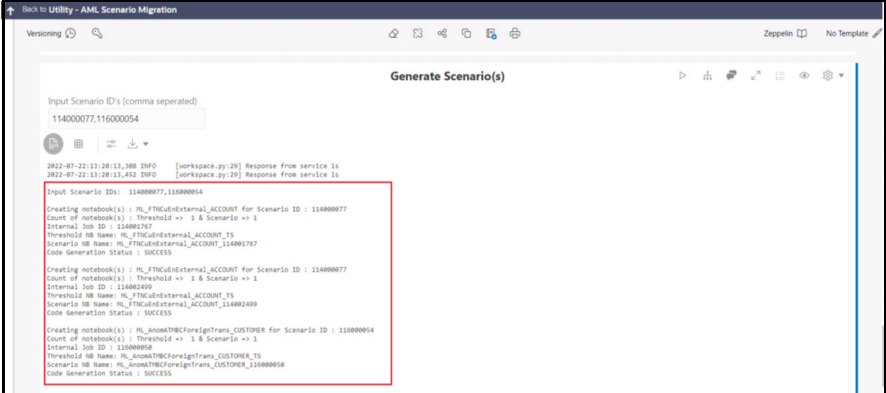
7. After successful execution, navigate to the **Model Pipelines** and click the **AML Scenarios-Converted** folder to display the list of scenarios created from conversion utility.
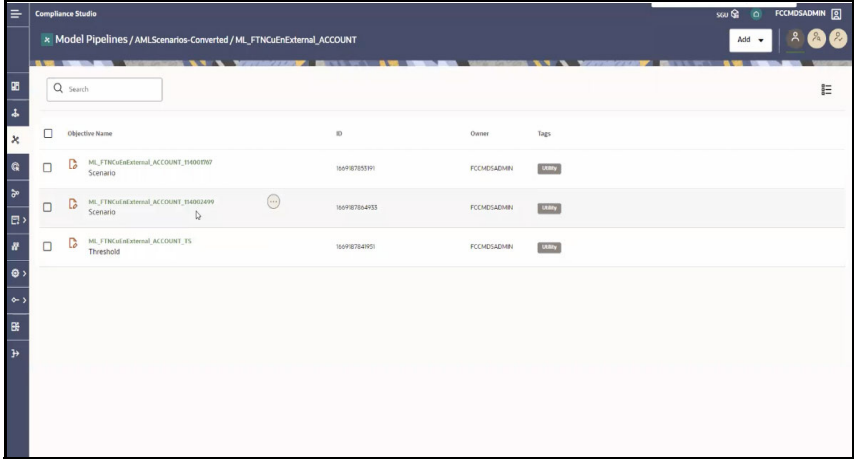
Figure 266: AML Scenarios-Converted Folder

8. Select the Scenario Notebook folder from the **AML Scenarios-Converted** folder.

Use the threshold and Scenario Notebook ID generated during the execution in step 5 to identify the Scenario Notebook folder.

| NOTE | For **n** number of jobs **n** number of Notebook IDs are created for scenario. |

Figure 267: Objective Name

9. Click **Scenario** Objective Name and select **Notebook tab** to run the paragraph and generate the events.

13.6.3.4 Running the Scenario

To run the scenario, follow these steps:

1. In Threshold Notebook, configure the following paragraphs:

   - Metadata
   - Base threshold set
   - Custom threshold set

Follow the markdown provided under **(1) About** and **(5) Custom Threshold Builder – Instructions** paragraphs for instructions.
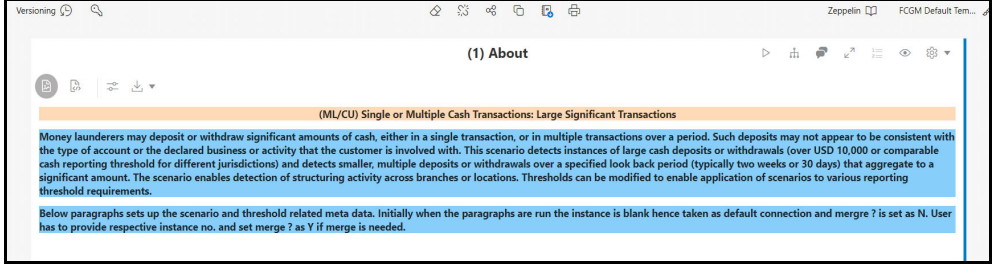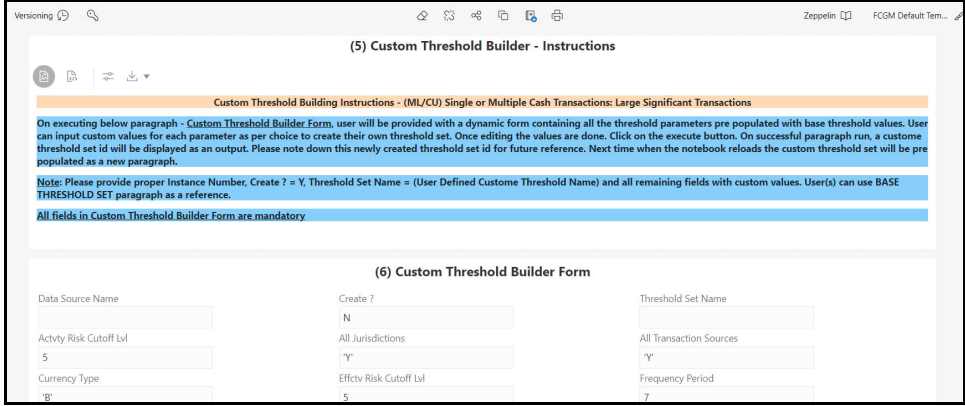
Figure 268: (1) About Paragraph

**Figure 269: (5) Custom Threshold Builder – Instructions Paragraph**

> **NOTE**   Threshold Notebook naming is created in the following format:
>
> **ML_<Scenario Catalogue Name>_<Focus>_TS**

2. Once the threshold configuration is completed, navigate to the respective scenario Notebook and click **Run Paragraphs** to generate events.

   Enter the custom value in the **Custom Threshold Set ID** field to run the scenario with a custom threshold set id.

> **NOTE**   Scenario Notebook naming is created in the following format:
>
> **ML_<Scenario Catalogue Name>_<Focus>_<Job ID>**

Figure 270: Run Paragraph
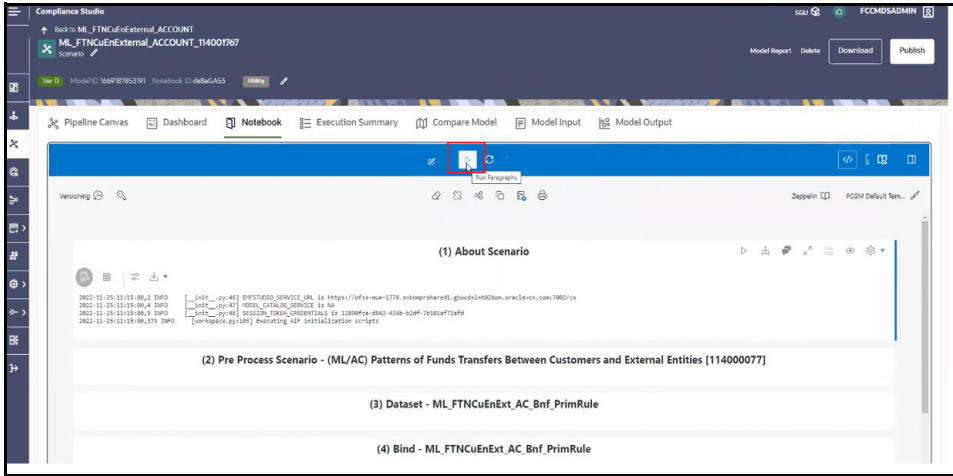
> **NOTE**   You can use the **Run ID** from the **(15) Create Event** paragraph to verify the data for whom the events are generated with the event details created in the BD schema.
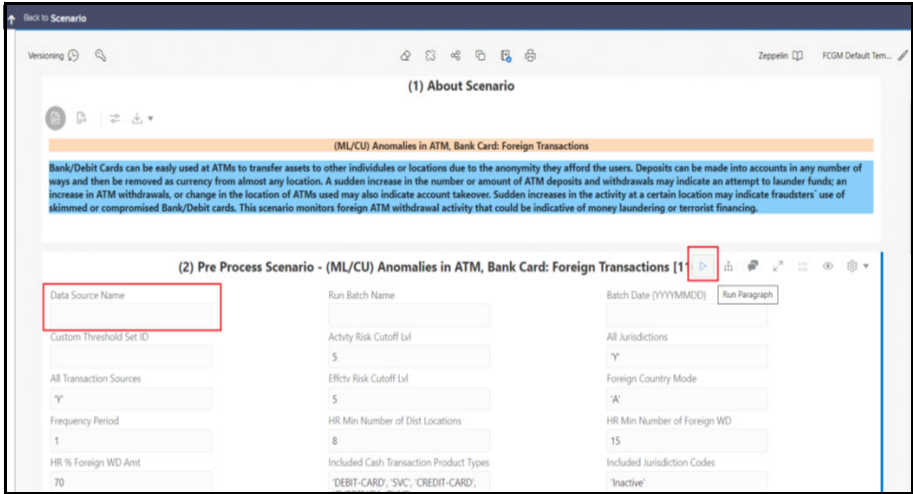
13.6.3.4.1 Using Custom Data Source

To use a custom data source (example: sandbox) for running the scenario(s), follow these steps:

1. To create a data source, see the Configuring Workspace Schema section.

2. Navigate to **Workspace summary** > **Managed Data Sources**. Select the custom data source name and enter the same name in the **Data Source Name** text box of the **(2) - Pre Process Scenario -<Scenario notebook>** paragraph in the respective scenario notebooks.

Figure 271: Custom Data Source

3. Click **Run Paragraph** to execute the scenario using a custom data source.

Figure 272: Run Paragraphs Action
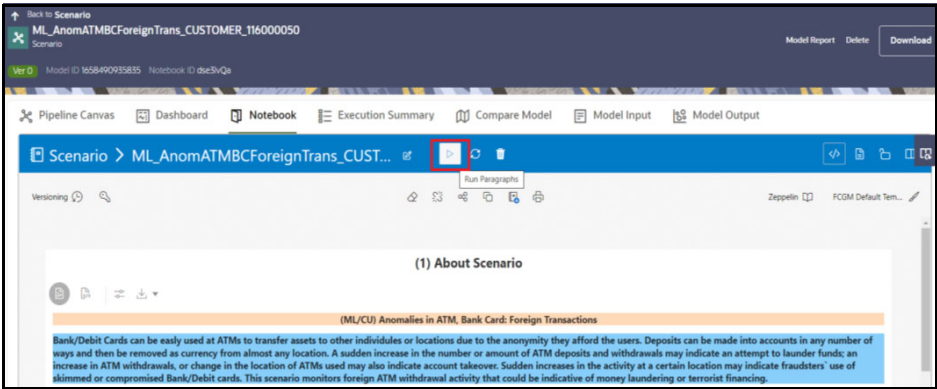
4. Once it is executed, click **Run Paragraphs** at the top of the Notebook to run the scenario notebook entirely rather than executing each paragraph in sequence.

> **NOTE**   By default, scenario notebooks use data sources as BD.

OFSAA Support

Raise a Service Request (SR) in My Oracle Support (MOS) for queries related to OFSAA applications.

Send Us Your Comments

Oracle welcomes your comments and suggestions on the quality and usefulness of this publication. Your input is an important part of the information used for revision.

- Did you find any errors?
- Is the information clearly presented?
- Do you need more information? If so, where?
- Are the examples correct? Do you need more examples?
- What features did you like most about this manual?

If you find any errors or have any other suggestions for improvement, indicate the title and part number of the documentation along with the chapter/section/page number (if available) and contact the My Oracle Support.

Before sending us your comments, you might like to ensure that you have the latest version of the document wherein any of your concerns have already been addressed. You can access My Oracle Support site which has all the revised/recently released documents.