Oracle® Product Data Quality

Governance Studio Reference Guide

Version 5.6

E20597-01

January 2011

ORACLE®

Oracle Product Data Quality Governance Studio Reference Guide, Version 5.6

E20597-01

Copyright © 2001, 2011 Oracle and/or its affiliates. All rights reserved.

Primary Author: Lorna Vallad

This software and related documentation are provided under a license agreement containing restrictions on use and disclosure and are protected by intellectual property laws. Except as expressly permitted in your license agreement or allowed by law, you may not use, copy, reproduce, translate, broadcast, modify, license, transmit, distribute, exhibit, perform, publish, or display any part, in any form, or by any means. Reverse engineering, disassembly, or decompilation of this software, unless required by law for interoperability, is prohibited.

The information contained herein is subject to change without notice and is not warranted to be error-free. If you find any errors, please report them to us in writing.

If this software or related documentation is delivered to the U.S. Government or anyone licensing it on behalf of the U.S. Government, the following notice is applicable:

U.S. GOVERNMENT RIGHTS Programs, software, databases, and related documentation and technical data delivered to U.S. Government customers are "commercial computer software" or "commercial technical data" pursuant to the applicable Federal Acquisition Regulation and agency-specific supplemental regulations. As such, the use, duplication, disclosure, modification, and adaptation shall be subject to the restrictions and license terms set forth in the applicable Government contract, and, to the extent applicable by the terms of the Government contract, the additional rights set forth in FAR 52.227-19, Commercial Computer Software License (December 2007). Oracle USA, Inc., 500 Oracle Parkway, Redwood City, CA 94065.

This software is developed for general use in a variety of information management applications. It is not developed or intended for use in any inherently dangerous applications, including applications which may create a risk of personal injury. If you use this software in dangerous applications, then you shall be responsible to take all appropriate fail-safe, backup, redundancy, and other measures to ensure the safe use of this software. Oracle Corporation and its affiliates disclaim any liability for any damages caused by use of this software in dangerous applications.

Oracle is a registered trademark of Oracle Corporation and/or its affiliates. Other names may be trademarks of their respective owners.

This software and documentation may provide access to or information on 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. 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.

Contents

Audience vii
Related Documents vii
Conventions viii

1 Introduction

Welcome to the Governance Studio 1-1
Convenience 1-2
Flexibility 1-2
Benefits 1-2
Improve Business Communication 1-2
Increase the Value of Your Reports 1-2
Maximize Efficiency & Scalability 1-2
Starting the Software 1-3
Understanding the Client Workspace 1-4
Frame Functionality 1-5
Menu Commands and the Toolbar 1-5
File Menu 1-6
Edit Menu 1-7
View Menu 1-8
Run Menu 1-8
Tools Menu 1-9
Help Menu 1-10
Keyboard Shortcuts 1-10
Tabs 1-10
Standard Tabs 1-10
Source Tab 1-11
Graph Summary Tab 1-11
Data Summary Tab 1-12
Output Tabs 1-12
Creating Output Tabs 1-13
Tab Scroll Arrows and Show Tab List Button 1-14
Filtering Data 1-14
Task Panes 1-15

2 Configuration and Administration

Configuring a DSA for Governance Studio 2-1
Accessing the Output Step in the DSA 2-1
Governance Studio Output Type Options 2-2
Governance Studio Graph Options 2-2
Starting the Governance Studio 2-4
Setting the Transform Server 2-4
Creating and Running a Governance Studio Project 2-5
Retrieve the Input Data 2-6
Opening an Excel Input File 2-7
Project Templates 2-9
Creating a Project Template 2-9
Opening a Project Template 2-9
Opening Projects Without Results 2-9
Configuring Options 2-10
Graphs Tab 2-10
Tabs Tab 2-11
General Tab 2-11
Memory Tab 2-12

3 Run and Review a Project

Running a Governance Studio Project 3-1
Synchronous Execution 3-1
Asynchronous Execution 3-1
Data Override Execution 3-1
Reviewing Project Output 3-2
General Processing 3-3
Applying a Secondary DSA 3-3
Re-Running a Secondary DSA 3-3
Completing Projects 3-4
Applying Quick Lookup 3-4
Other Features 3-5
Viewing Job Status 3-5
Getting Job Results 3-6
Emailing Checked Rows 3-7
Exporting and Importing Projects 3-7
Exporting a Project 3-7
Importing a Project 3-8

4 Basic Tabs

Source Tab 4-1
Graph Summary Tabs 4-2
Data Summary Tab 4-3

5 DSA Output Tabs

Review, Approve and Route Output Tabs 5-1

Output Tab Functions 5-1
Row Selection 5-1
Copying Data 5-2
Modifying Cell Contents 5-2
Review, Approve, and Route AutoSuggest Output Tabs 5-3
AutoSuggest Feature 5-3
Apply Augmentations Feature 5-4
Match Results Tabs 5-5
Top Pane Functions 5-7
Bottom Pane Functions 5-7
Working with a Match Set 5-8
Request for Quote 5-8
Functional Equivalents Cross Reference 5-9
Survivorship Process 5-9
Trend Analysis 5-10
Create Task from Checked Rows 5-11
View My Tasks 5-12
Changing the Task Status 5-13
Creating a Task 5-13

A Installing the Client Software

Preface

This reference guide is intended to explain the basic capabilities of the Oracle Product Data Quality Governance Studio.

To understand all of the advanced features presented, you must use this reference guide in conjunction with the Oracle Product Data Quality documents listed in "Related Documents" on page -vii.

You must have Oracle Product Data Quality client software installed on your computer including all of the sample files.

Review the following Oracle Product Data Quality documentation prior to the use of this manual is recommended:

- *Oracle Product Data Quality Knowledge Studio Reference Guide*
- *Oracle Product Data Quality Application Studio Reference Guide*

In addition, Oracle Product Data Quality Governance Studio training is encouraged.

Audience

You should have a basic understanding of the DataLens Technology, including the functionality of the Oracle Product Data Quality Knowledge Studio and the Oracle Product Data Quality Application Studio.

This document is intended for all users of the DataLens Technology, including:

- Customers
- Sales Consultants
- Implementation Personnel
- Software Engineers
- Knowledge Engineers

Related Documents

For more information, see the following documents in the documentation set:

- The *Oracle Product Data Quality Oracle DataLens Server Installation Guide* provides detailed Oracle Product Data Quality Oracle DataLens Server installation instructions.
- The *Oracle Product Data Quality Oracle DataLens Server Administration Guide* provides information about installing and managing an Oracle DataLens Server.

- The *Oracle Product Data Quality COM Interface Guide* provides information about installing and using the Oracle DataLens Server COM APIs.
- The *Oracle Product Data Quality Java Interface Guide* provides information about installing and using the Oracle DataLens Server Java APIs.
- The *Oracle Product Data Quality Application Studio Reference Guide* provides information about creating and maintaining Data Service Applications (DSAs).
- The *Oracle Product Data Quality AutoBuild Reference Guide* provides information about creating initial data lens based on existing product information and data lens knowledge.
- The *Oracle Product Data Quality Knowledge Studio Reference Guide* provides information about creating and maintaining data lenses.
- The *Oracle Product Data Quality Glossary* provides definitions to commonly used Oracle Product Data Quality technology terms.
- The *Oracle Product Data Quality Services for Excel Reference Guide* provides information about creating a DSA based on data contained in a Microsoft Excel spreadsheet.
- The *Oracle Product Data Quality Task Manager Reference Guide* provides information about managing tasks created with the Task Manager or Governance Studio applications.

See the latest version of this and all documents listed at the Oracle Product Data Quality Documentation Web site at:

`http://download.oracle.com/docs/cd/E20593_01/index.htm`

Conventions

The following text conventions are used in this document:

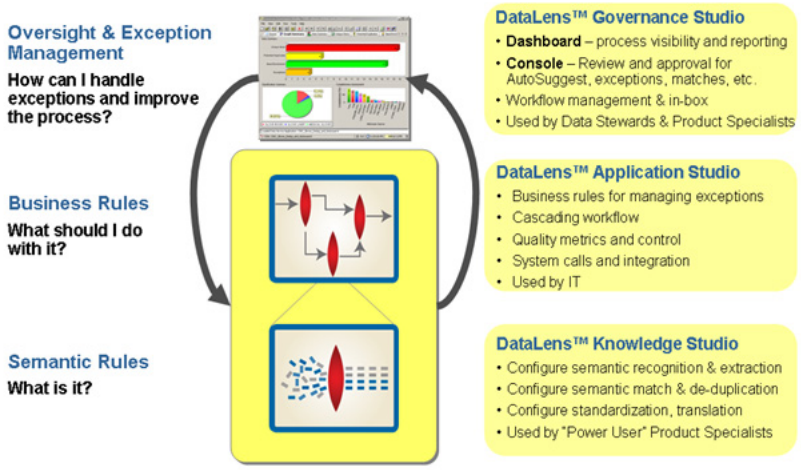
Convention	Meaning
boldface	Boldface type indicates graphical user interface elements associated with an action, or terms defined in text or the glossary.
italic	Italic type indicates book titles, emphasis, or placeholder variables for which you supply particular values.
`monospace`	Monospace type indicates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

1

Introduction

Oracle DataLens Server is built on industry-leading DataLens™ Technology to standardize, match, enrich, and correct product data from different sources and systems. The core DataLens Technology uses patented semantic technology designed from the ground up to tackle the extreme variability typical of product data.

Oracle Product Data Quality uses three core DataLens Technology modules: Governance Studio, Knowledge Studio, and Application Studio. The following figure illustrates the process flow of these modules.



The Governance Studio is a reporting and graphing application that streamlines the analysis of results from the Application Studio Data Service Application (DSA) job runs. You can create a Governance Studio project to create reports on data quality, analyze transformed data, email records to coworkers, and output the results to a file. The AutoSuggest feature in the Governance Studio assists you in identifying missing attributes with the click of a button.

Welcome to the Governance Studio

The Governance Studio provides a flexible framework for you to design an analysis tool that suits your data interpretation needs. This tool, or project, is entirely dependent on the DSA that you design to process your input data.

This guide explains the features and functionality delivered in the Governance Studio and provides examples of how you might use them to build your own projects. These

examples are based on a set of use cases to demonstrate Governance Studio features. Because Governance Studio jobs reflect the underlying DSA, it is possible to create many types of jobs to meet the needs of an individual project. For example, Governance Studio jobs that rely on database input and output or do not return results will behave very differently than those jobs that use text input and output.

Convenience

The Governance Studio offers a convenient and easy-to-use interface for launching jobs, managing processes, reporting results, and analyzing output from Application Studio jobs, as well as analyzing all output data created. You can schedule and run jobs synchronously or asynchronously, generate reports, do trend analysis, and identify missing attributes with ease and report discrepancies to coworkers.

Flexibility

The Governance Studio provides a flexible, user-friendly interface for:

- Running DSAs
- Outputting DSA results
- Manipulating DSA results
- Automatically suggesting attributes using AutoSuggest
- Saving output data

Benefits

The benefits of using the Governance Studio are described in the following sections.

Improve Business Communication

Data lens reports can help you meet your business goals. Use a data lens to:

- Demonstrate value and benefit (ROI) quickly and effectively.
- Enable you to visualize results and readily 'connect the dots'.

Increase the Value of Your Reports

The Governance Studio enables highly effective reporting in real time using high quality graphical images that reflect the quality of your data along a set of parameters you define to enable downstream processing effectively or communication about your data.

The Governance Studio graphically enhances results from the output of any DSA template and allows you to create and leverage DSA templates for one-to-one alignment with business processes, such as:

- General purpose product data assessment
- Attribute extraction
- Matching data records between existing and legacy data sources
- Proposing alternatives based on search criteria

Maximize Efficiency & Scalability

The Governance Studio leverages Java Web Start to enable you to:

- Launch the Governance Studio application from a link on a web page
- Update software to the current version automatically
- Seamlessly integrate with Windows or Linux operating system (OS)
- Leverage DSA database interfaces

Starting the Software

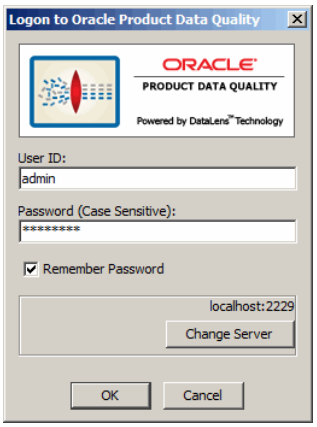
You can start Oracle Product Data Quality by using either the desktop shortcut or the Windows **Start** menu as follows:

> **Note:** If Oracle Product Data Quality is not installed, you can install it using the instructions in Installing the Client on page 72.

- Double-click the desktop shortcut.

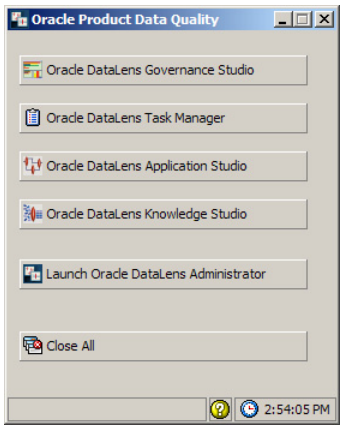
- Click **Start, Programs**, **Oracle Product Data Quality**, and select **Oracle Product Data Quality**.

The **Oracle Product Data Quality Login** dialog box appears.



Enter your user name and password and click **OK**. You can avoid entering your password every time you logon by selecting the **Remember Password** check box. If you want to change your Oracle DataLens Server, click **Change Server** to select a new server.

The **Oracle Product Data Quality Launch Pad** appears.

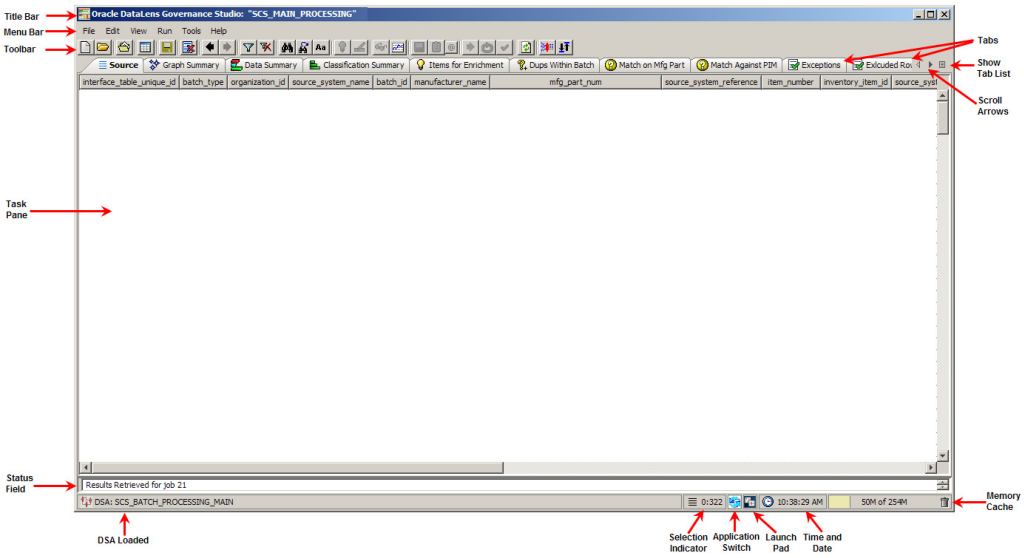


The **Oracle Product Data Quality Launch Pad** allows you to quickly start any of the Oracle DataLens Server applications by clicking on any of the buttons. You can close all open Oracle Product Data Quality applications using the **Close All** button.

Click the **Oracle DataLens Governance Studio** button to start the application.

Understanding the Client Workspace

The Governance Studio graphical user interface (GUI) provides the client workspace used to create and manage a data lens.



This section describes the following areas of the client workspace:

- "Frame Functionality" on page 1-5
- "Menu Commands and the Toolbar" on page 1-5
- "Tabs" on page 1-10
- "Task Panes" on page 1-15

Frame Functionality

The Governance Studio client workspace frame contains useful information and interactive functions including the following:

Title Bar
Indicates the current application and open project.

Status Field
Provides the status of the project one line at a time. Though this field cannot be resized, the scroll arrows on the right-hand side can be used to view all available status information. The status data does not change based on the selected tab; rather it is a compilation of all data.

Loaded DSA
Indicates the currently loaded DSA that is used for transforming the project data.

Selection Indicator
Indicates the number of the row currently selected and the total number of rows available for selection.

Application Switch
Returns you to the last Oracle Product Data Quality application used.

Oracle Product Data Quality Launch Pad
This button opens the Oracle Product Data Quality Launch Pad so that you can select other applications.

Time and Date
The time is displayed and when you hover over this field.

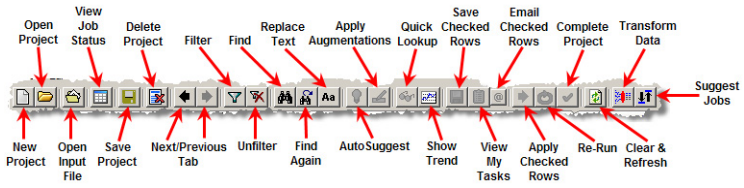
Memory Cache
Indicates the amount of memory cache currently used and the total amount allowed. You can dump the memory cache by clicking on the trash can icon in this interactive field.

Note: This feature is only used for system diagnosis and should not be used unless requested by the support team.

Menu Commands and the Toolbar

The Governance Studio toolbar allows easy access to the most frequently used Governance Studio functions. Though the set of toolbar buttons remains the same during UI operation the buttons are enabled or disabled based the current state of the interface and the options set. Buttons displayed with shades of gray are disabled. Full-color buttons are enabled. All toolbar buttons are standard push buttons, requiring a single click of the mouse to activate.

The following briefly describes the toolbar buttons from left to right.



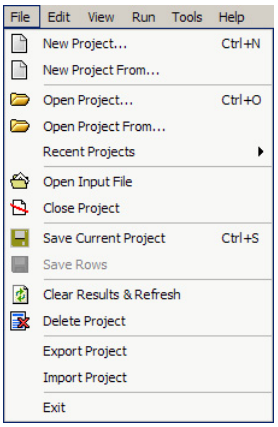
The Governance Studio GUI menus provide access to most functions. All of the buttons on the toolbar have a corresponding menu command, which are indicated on each menu with the button icon displaying adjacent to the command. The set of menu commands remains the same during the GUI operation.

Menu commands are enabled or disabled based on the current state of the data lens; commands that are dimmed are unavailable. Some menu commands perform functions that are more complex and are indicated by an ellipsis symbol (...). These commands open dialog boxes to collect information needed to complete the requested function. Menu commands that toggle functions are preceded by check mark (ü).

> **Tip:** The tooltips appear when you rest your mouse pointer on a menu item, button, tab, icon, or similar content.

The following section briefly describes each of the Governance Studio menu commands and corresponding buttons.

File Menu



New Project

Creates a new Governance Studio project for building reports about your enterprise data. These project files are stored in one of the following directories:

`C:\Documents and Settings\`*`Username`*`\Applications\DataLens\data\project`

or

`C:\Users\`*`Username`*`\AppData\DataLens\export`

New Project From…

Creates a new project using an existing project file as the basis and closes any open project.

Open Project

Opens an existing project file and closes any open project file.

Open Project From…

Opens an existing project file and allows you to rename it using the same base project name with an added unique suffix (for example, V2.)

Open Project - No Results

Opens an existing project file though does not display data processing results in any Output tabs. This feature allows you to open a project that previously exceeded the maximum memory allocation and so could not complete opening. For more information, see "Opening Projects Without Results" on page 2-9.

Recent Projects

Provides a list of recently opened projects so that you can quickly select a project to open.

Open Input File

Allows you to select the data file for the current project. For more information, see "Retrieve the Input Data" on page 2-6.

Close Project

Closes the open project file.

Save Current Project

Saves all project changes to disk.

Save Rows

Saves only rows that have been checked; does not save selected rows. The checked rows are automatically saved to a Microsoft Excel spreadsheet in the datalens directory where the project resides. The name of the file is based on the name of the tab and is stored in the datalens directory where the project resides. You can change the file name or save location in the dialog provided.

Clear Results & Refresh

Allows you to refresh, or remap, a modified DSA to the source data to ensure that all DSA changes are reflected in the current project.

If you open your project and do not refresh it, the Governance Studio will prompt you to do so to ensure consistency.

Delete Project

Deletes the open project. You are no longer able to edit this project although you are no longer able to edit this project, it can be modeled for use as a new project using the **New Project From…** option.

Export Project

Exports the project. For more information, see "Exporting and Importing Projects" on page 3-7.

Import Project

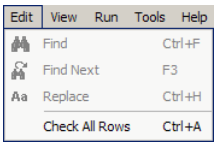
Import the project. For more information, see "Exporting and Importing Projects" on page 3-7.

Exit

Exits the Governance Studio application; a prompt is given to save any changes that were made.

Edit Menu

Find



Allows you to specify a search string using a regular expression and attempts to find it.

Find Next

Repeats the last search defined by a **Find** operation.

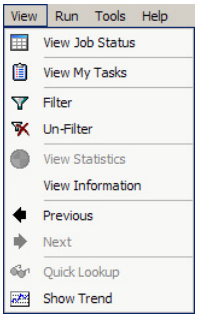
Replace

Allows you to specify a search string and a replacement string, and then attempts to replace the text from the **Source** tab.

Check All Rows

Selects all rows in the table on the active tab. Only active when the selected tab contains a table with the Check mark column.

View Menu



View Job Status

Allows you to view the status of all jobs.

View My Tasks

Allows you to view any tasks that are scheduled or have run. For more information, see "View My Tasks" on page 5-12.

Filter

Allows you to filter the displayed data based on text or a text pattern. The filter operation applies only to the currently selected tab. Only the rows that match the text entered in the **Filter** dialog are displayed in the table.

Un-Filter

Removes the filter applied to the selected tab, thus displaying all of the data.

View Statistics

Provides statistical information about the selection. For more information, see "Match Results Tabs" on page 5-5.

View Information

Provides an informational message containing the project name, Transform Server, DSA name, and the location of the data file.

Previous / Next

The arrow buttons allow you to view alternative definitions for a line item. They operate differently depending on the selected tab.

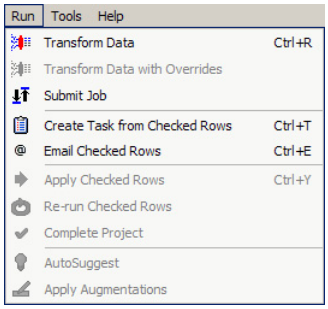
Quick Lookup

Allows you to quickly locate and view Ngram matching results. When the associated DSA does not contain Ngram matching, this option is inactive. For more information, see "Applying Quick Lookup" on page 3-4.

Show Trend

Provides a graphical trend analysis of the output data. For more information, see "Trend Analysis" on page 5-10.

Run Menu



Transform Data

Processes the output data in real-time. For more information, see "Running a Governance Studio Project" on page 3-19.

Transform Data with Overrides

Allows you to process the output data with settings that you provide at runtime that override those configured in the DSA. You can override the threshold and fuzzy matching settings for Ngram Match or an Attribute Match 2 (associated with Semantic Key 2) processing. When the associated DSA does not contain these matching processes, this option is inactive.

Submit Job

Processes the output data in the background. For more information, see "Running a Governance Studio Project" on page 3-1.

Create Task from Checked Rows

Creates a new job using the selection. For more information, see "Create Task from Checked Rows" on page 5-117.

Email Checked Rows

E-mails the selected rows. For more information, see "Emailing Checked Rows" on page 3-7

Apply Checked Rows

Processes the output data in the rows that you have selected. For more information, see "General Processing" on page 3-3.

Re-Run Checked Rows

Reprocesses the output data in the rows that you have selected. For more information, see "General Processing" on page 3-3.

Complete Project

Closes out a project. For more information, see "Completing Projects" on page 3-4.

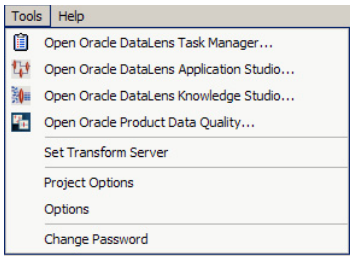
AutoSuggest

Obtain data processing suggestions from the Governance Studio. For more information, see "AutoSuggest Feature" on page 5-3.

Apply Augmentations

Add or augment knowledge in the loaded DSA. For more information, see "Apply Augmentations Feature" on page 5-4.

Tools Menu



Open Oracle DataLens Task Manager…

Starts the Oracle Product Data Quality Task Manager. See the *Oracle Product Data Quality Task Manager Reference Guide*.

Open Oracle DataLens Application Studio…

Starts the Oracle Product Data Quality Application Studio. See the *Oracle Product Data Quality Application Studio Reference Guide*.

Open Oracle DataLens Knowledge Studio…

Starts the Oracle Product Data Quality Knowledge Studio. See the *Oracle Product Data Quality Knowledge Studio Reference Guide*.

Open Oracle Product Data Quality…

Starts the Oracle Product Data Quality Launch Pad.

Set Transform Server

Identifies the Oracle DataLens Server that will be used to transform data. For more information, see "Setting the Transform Server" on page 2-4.

Project Options

Allows you to set options for the open project. For more information, see "Configuring Options" on page 2-10.

Options

Allows you to set options for global use in the Governance Studio. For more information, see "Configuring Options" on page 2-10.

Change Password

Enables you to change your Oracle Product Data Quality password.

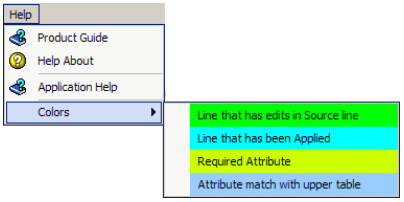
Help Menu

Product Guide

Opens a list of Oracle Product Data Quality documents for your selection in a browser.

Help About

Provides information regarding the product including the version number and a link to view third party product licenses.



Application Help

Opens the Governance Studio online help in a browser.

Colors

Describes the colors used in the data tables used in the Governance Studio.

Keyboard Shortcuts

The following table contains keyboard shortcuts that can help make the Governance Studio easier to use.

Function	Shortcut Key
New Project	Ctrl+N
Open Project	Ctrl+O
Save Project	Ctrl+S
Undo	Ctrl+Z
Find	Ctrl+F
Find Next	F3
Replace	Ctrl+H
Check All Rows	Ctrl+A
Transform Data	Ctrl+R
Create Task from Checked Rows	Ctrl+T
Email Rows	Ctrl+E
Apply Checked Rows	Ctrl+Y

Tabs

A tab groups like information into easy to read and access areas that include graphs, panes, and text entry boxes. Tabs are displayed in the client workspace directly under the toolbar and can be activated in any order. Not all tabs are available at all times. For example, the Output tabs are not active when you create a new Governance Studio project because the output has not yet been generated.

Standard Tabs

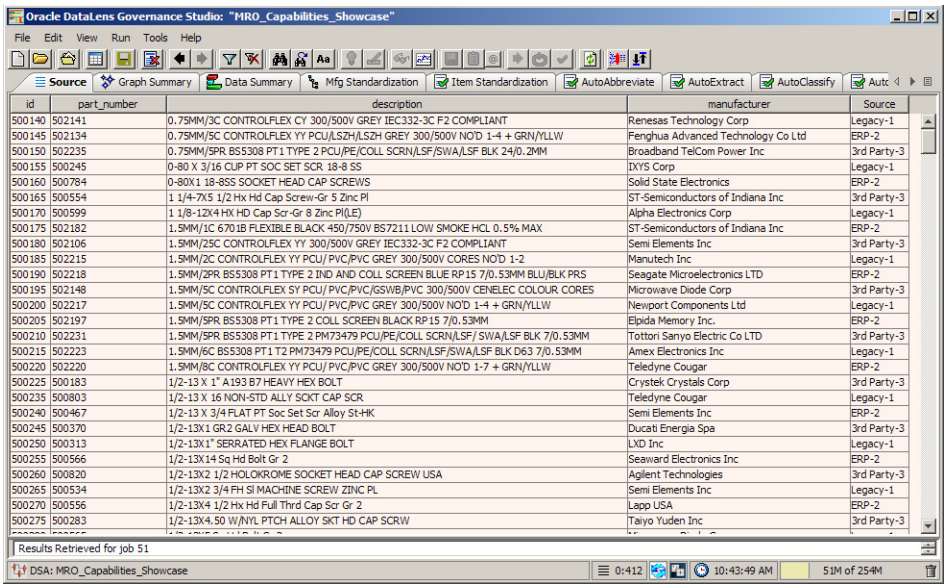
The standard data tabs provide representations of the aggregate quality and productivity metrics for the project. The graphical output tabs are highly configurable, and double-clicking a bar takes you to the underlying output. You can use these tabs to achieve business objectives in a variety of areas:

- Manage data quality by source:

- Determine quality of the data source
- Allow closed-loop governance of sources

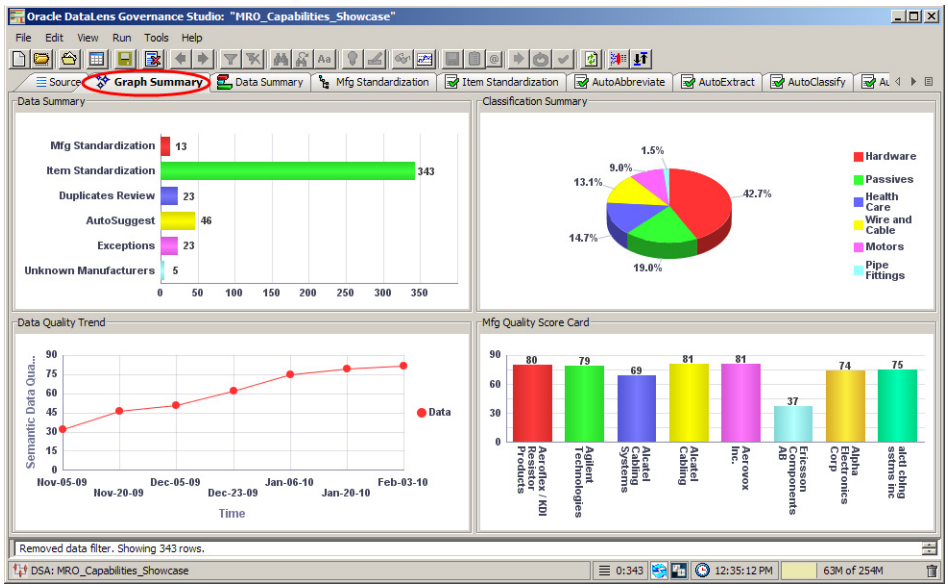
- Monitor process effectiveness:
  - Determine how often a process is invoked, how often it is successful
  - Identify process improvement opportunities
- Assess work-queue visibility and productivity metrics by user

Source Tab The **Source** tab is created when the Governance Studio project is based on an input file load or on manual input. If the project specifies an external file, or if the underlying DSA is based on database input, or if the project connects to a scheduled job, there is no **Source** tab.

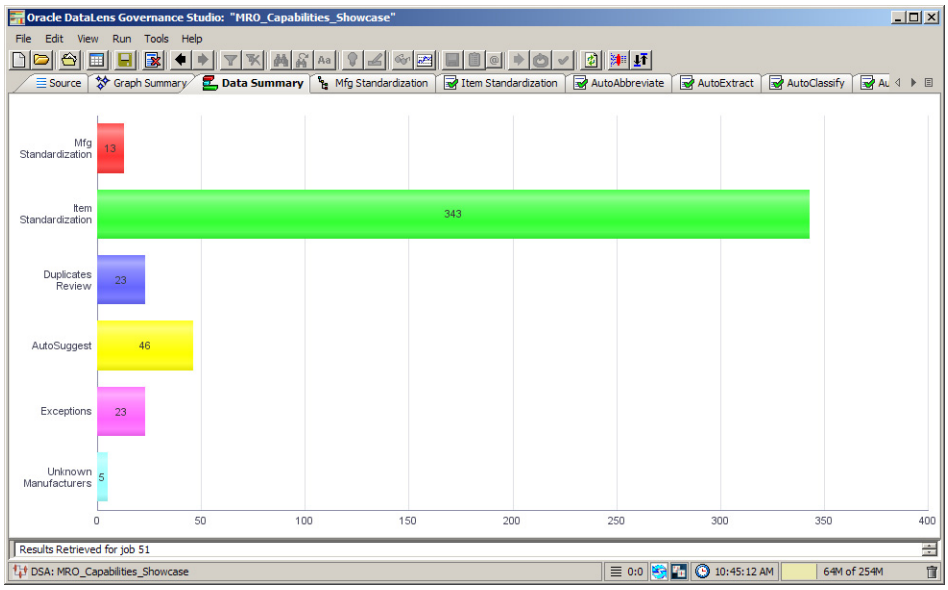


id	part_number	description	manufacturer	Source
500140	502141	0.75MM/3C CONTROLFLEX CY 300/500V GREY IEC332-3C F2 COMPLIANT	Renesas Technology Corp	Legacy-1
500145	502134	0.75MM/5C CONTROLFLEX YY PCU/LSZH/LSZH GREY 300/500V NO'D 1-4 + GRN/YLLW	Fenghua Advanced Technology Co Ltd	ERP-2
500150	502235	0.75MM/5PR BS5308 PT1 TYPE 2 PCU/PE/COLL SCRN/LSF/SWA/LSF BLK 24/0.2MM	Broadband TelCom Power Inc	3rd Party-3
500155	500245	0-80 X 3/16 CUP PT SOC SET SCR 18-8 SS	IXYS Corp	Legacy-1
500160	500784	0-80X1 18-8SS SOCKET HEAD CAP SCREWS	Solid State Electronics	ERP-2
500165	500554	1 1/4-7X5 1/2 Hx Hd Cap Screw-Gr 5 Zinc Pl	ST-Semiconductors of Indiana Inc	3rd Party-3
500170	500599	1 1/8-12X4 HX HD Cap Scr-Gr 8 Zinc Pl(LE)	Alpha Electronics Corp	Legacy-1
500175	502182	1.5MM/1C 6701B FLEXIBLE BLACK 450/750V BS7211 LOW SMOKE HCL 0.5% MAX	ST-Semiconductors of Indiana Inc	ERP-2
500180	502106	1.5MM/25C CONTROLFLEX YY 300/500V GREY IEC332-3C F2 COMPLIANT	Semi Elements Inc	3rd Party-3
500185	502215	1.5MM/2C CONTROLFLEX YY PCU/ PVC/PVC GREY 300/500V CORES NO'D 1-2	Manutech Inc	Legacy-1
500190	502218	1.5MM/2PR BS5308 PT1 TYPE 2 IND AND COLL SCREEN BLUE RP15 7/0.53MM BLU/BLK PRS	Seagate Microelectronics LTD	ERP-2
500195	502148	1.5MM/5C CONTROLFLEX SY PCU/ PVC/PVC/GSWB/PVC 300/500V CENELEC COLOUR CORES	Microwave Diode Corp	3rd Party-3
500200	502217	1.5MM/5C CONTROLFLEX YY PCU/ PVC/PVC GREY 300/500V NO'D 1-4 + GRN/YLLW	Newport Components Ltd	Legacy-1
500205	502197	1.5MM/5PR BS5308 PT1 TYPE 2 COLL SCREEN BLACK RP15 7/0.53MM	Elpida Memory Inc.	ERP-2
500210	502231	1.5MM/5PR BS5308 PT1 TYPE 2 PM73479 PCU/PE/COLL SCRN/LSF/ SWA/LSF BLK 7/0.53MM	Tottori Sanyo Electric Co LTD	3rd Party-3
500215	502223	1.5MM/6C BS5308 PT1 T2 PM73479 PCU/PE/COLL SCRN/LSF/SWA/LSF BLK D63 7/0.53MM	Amex Electronics Inc	Legacy-1
500220	502220	1.5MM/8C CONTROLFLEX YY PCU/ PVC/PVC GREY 300/500V NO'D 1-7 + GRN/YLLW	Teledyne Cougar	ERP-2
500225	500183	1/2-13 X 1" A193 B7 HEAVY HEX BOLT	Crystek Crystals Corp	3rd Party-3
500235	500803	1/2-13 X 16 NON-STD ALLY SCKT CAP SCR	Teledyne Cougar	Legacy-1
500240	500467	1/2-13 X 3/4 FLAT PT Soc Set Scr Alloy St-HK	Semi Elements Inc	ERP-2
500245	500370	1/2-13X1 GR2 GALV HEX HEAD BOLT	Ducati Energia Spa	3rd Party-3
500250	500313	1/2-13X1" SERRATED HEX FLANGE BOLT	LXD Inc	Legacy-1
500255	500566	1/2-13X14 Sq Hd Bolt Gr 2	Seaward Electronics Inc	ERP-2
500260	500820	1/2-13X2 1/2 HOLOKROME SOCKET HEAD CAP SCREW USA	Agilent Technologies	3rd Party-3
500265	500534	1/2-13X2 3/4 FH SI MACHINE SCREW ZINC PL	Semi Elements Inc	Legacy-1
500270	500556	1/2-13X4 1/2 Hx Hd Full Thrd Cap Scr Gr 2	Lapp USA	ERP-2
500275	500283	1/2-13X4.50 W/NYL PTCH ALLOY SKT HD CAP SCRW	Taiyo Yuden Inc	3rd Party-3

Graph Summary Tab The **Graph Summary** tab provides summaries of the **Transformed Data** tabs that are configured in the DSA for graphical output. This tab is created if there are at least two graphs to display.



Data Summary Tab The **Data Summary** tab shows a graph of the number of transformed rows associated with each of the **Transformed Data** tabs.



Output Tabs

The Output tabs offer item-level validation, approval, edit, and routing. The output is displayed in a single tabular view in the top pane.

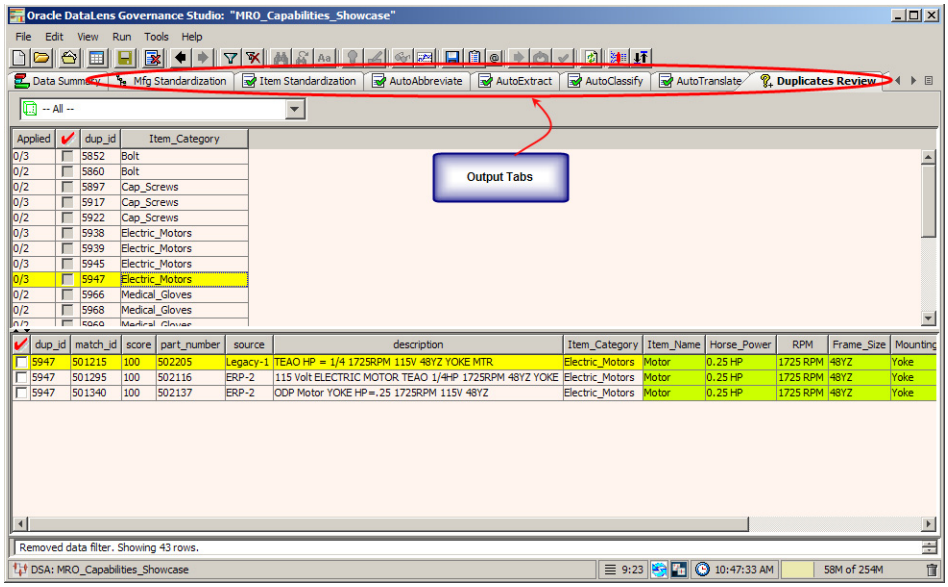
There are several types of Governance Studio Output tabs. Each type has different rules governing its creation and different properties that permit you to perform actions on the output each tab displays. Output tabs in the Governance Studio correspond to output steps in your Data Service Application.

The Output tabs provide task-specific User Interfaces (UIs) for any exception process including the following:

- Review and approve parsing results

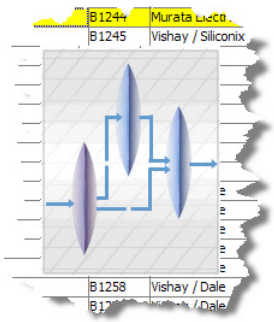
- Review and approve potential matches
- Reassign low quality items to different workflow/process
- Direct edit/update of low quality items
- Enables manual intervention while preserving process integrity and efficiency

The following is an example of a match selection process, upon the selection of a row, a parent-child relationship view appears in the bottom pane for match results. These split-screen task panes provide you with an easy to use interactive console to perform tasks associated with reviewing and approval or rejecting duplicates, matches, and functional equivalents



Creating Output Tabs For a new project, Output tabs are not initially displayed in the Governance Studio project. After transforming the **Source** tab data using the **Transform** button, the Output tabs are displayed. While transforming data or obtaining source data from a database query, the **Job-In-Progress** icon in the center of the tab shows that the run is in progress.

For an existing job, the Output tab data is recreated with each job run.

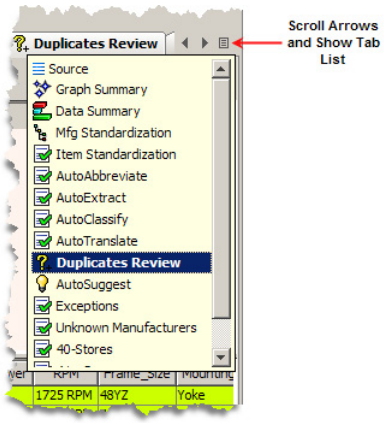


If Output tabs already exist, the Governance Studio compares the column headings of the new data with the column headings of the data in all of the existing Output tabs. If there is a data match (the columns are identical), the existing Output tabs are used.

Otherwise, the Governance Studio indicates the mismatch with an error message though it automatically creates the associated Output tab.

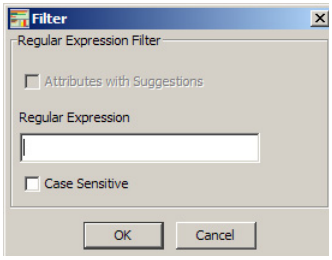
Tab Scroll Arrows and Show Tab List Button

After transforming your project data, one or more Output tabs will appear depending on whether the output steps in your DSA specified Governance Studio output. If there are more Output tabs in the Governance Studio console than can display on one page, you can click the left and right **Scroll Arrows** at the right to scroll to the right or back to the left. Also, you can click the **Show Tab List** button as shown in following figure to see a list of Output tabs to select them individually.



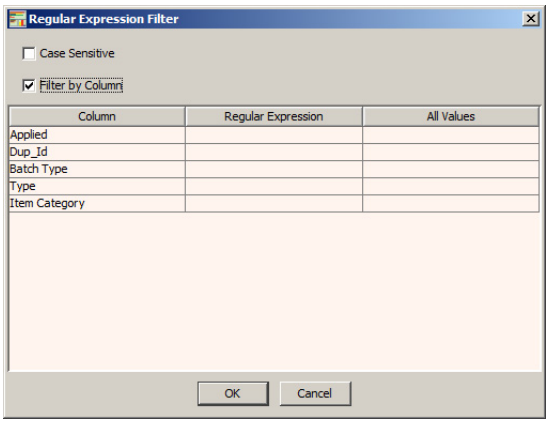
Filtering Data

The tabs that contain tables in the Governance Studio can be filtered to view specific information thereby reducing the number of rows displayed in a table. To filter your table click the **Filter** button on the toolbar or from the **View** menu, select **Filter**.



If the data that you enter is case sensitive, select the **Case Sensitive** check box to ensure that proper filtering. Enter a regular expression or a text string that you want to use to filter your data.

You can filter each column of data by selecting the **Filter by Column** check box.



Each column of the selected tab is displayed. You can enter a regular expression for one or more columns or you click in the **All Values** column to select a filter value from the auto-populated lists. The **Regular Expression** and **All Values** columns are mutually exclusive so an entry or selection in one column resets the other column. The values listed in the **All Values** column are derived from the values currently in the table so repeated filtering will produce fewer selections. This allows you to refine your filtering based on what is visible. If the values you want to filter on do not appear, unfilter your data, and then filter again.

The Attributes with Suggestions is active on those tabs that the AutoSuggest output (see "AutoSuggest Feature" on page 5-3) and allows you to search suggested data as well.

The **Filter** button is colored green when a filter is active so that you can easily identify that all data is not displayed. When you remove the filter, using the **Un-Filter** button or from the **View** menu, select **Un-Filter**, all of the data is displayed.

Task Panes

The Governance Studio includes two interactive task panes: the Governance Console where data is displayed in tabular form similar to a Microsoft Excel spreadsheet and the Governance Dashboard where graphical representations of the output data are displayed.

The small up/down arrows between the panes on the left-hand side, allow you to resize the panes. In addition, you can fully expand either pane to see more data by clicking on an arrow, which makes the pane inactive. To redisplay the inactive pane, click the opposite arrow and the pane reappears.

There are various context-sensitive (shortcut) menus that appear in the Governance Studio panes when you right-click on data within a pane. The contents of these menus are described throughout this reference.

2

Configuration and Administration

How to configure and administer the Governance Studio is explained in this chapter.

Configuring a DSA for Governance Studio

In order to use the Governance Studio, you must use a DSA that has been properly configured. The configuration requires selection of appropriate output types. The functions of the Governance Studio options for output types are described in this section.

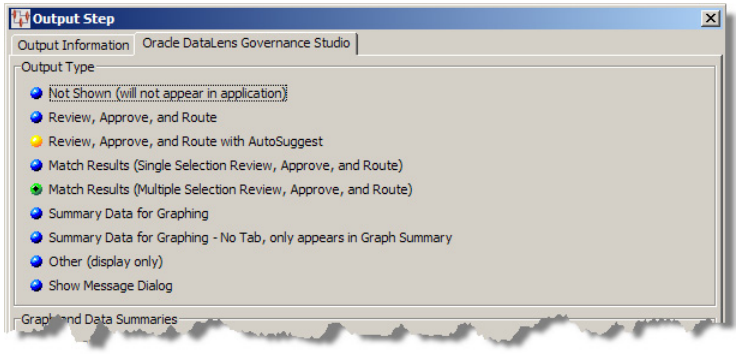
Additionally, this section is a summary of the information required to create a DSA template for a Governance Studio project. For more information on DSA design, see *Oracle Product Data Quality Application Studio Reference Guide* or consult with your Oracle Product Data Quality Administrator.

Accessing the Output Step in the DSA

In the Application Studio, open a DSA, and then double-click any **Text** output node.

Note: Governance Studio projects report on text-based output only. The Governance Studio will ignore other types of outputs such as database outputs provided you set the output type to *Not Shown (will not appear in application)* as described in this section.

Click the **Oracle DataLens Governance Studio** tab to view the Governance Studio output configuration options.



Governance Studio Output Type Options

As shown in previous figure, you may choose one of the following output types to determine how data is displayed in the Governance Studio using the **Output Step** dialog in the Application Studio:

Not shown (will not appear in application)
Results will not appear in the Governance Studio console Output tabs. Use this output type for results that do not conform to Governance Studio requirements, such as database output steps.

Review, Approve and Route
Results are the records of data that exactly meet the output criteria specified are displayed. Use this output type for displaying the results of a DSA process that requires a person to continue processing.

Review, Approve and Route with AutoSuggest
Results are the records of data that exactly meet the output criteria specified are displayed. Use this output type for displaying the results of a DSA process where the records fall below the quality threshold and require enrichment. A person would have the opportunity to invoke the AutoSuggest feature to provide a set of suggestions to enrich the records that have missing attribute values.

Match Results (Single Selection Review, Approve, and Route)
Possible data records meeting match or duplicate criteria are displayed below a split-screen in the output spreadsheet. Only a single data record from the list of matched items can be selected; the record is available for downstream processing. Use this output type to display the results of a DSA template that runs a match process.

Match Results (Multiple Selection Review, Approve, and Route)
Records meeting match or duplicate criteria are displayed below a split-screen in the output spreadsheet. Multiple records from the list of matched items can be selected; the records are available for downstream processing. Use this type for displaying the results of a DSA template that runs a match process.

Summary Data for Graphing
Summary information is displayed in the graph; no selection of records is available with this option. The bars in the graph operate like buttons and when clicked jump to the Output tab for the data represented in the selected bar.

Summary Data for Graphing - No Tab, only appears in Graph Summary
An individual tab is not created; the data appears in the **Graph Summary** tab only.

Other (Display Only)
Result information is displayed in a spreadsheet form that does not allow user interaction.

Show Message Dialog
Informational messages, including errors and warnings, are displayed to the user.

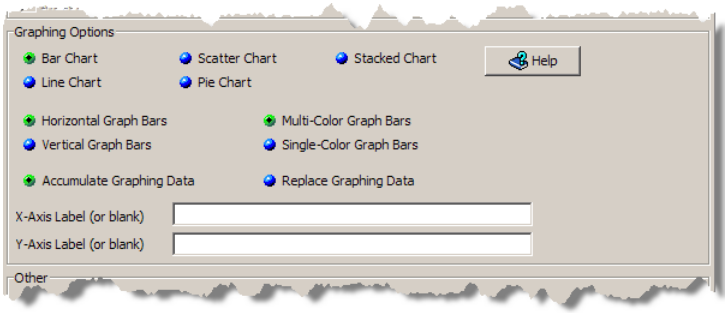
Governance Studio Graph Options

This section of the **Output Step** dialog in the Application Studio is active only when one of the **Summary Data for Graphing** options are selected as described in the previous section.

There are four graphing options available for the output types. You can mix and match between two sets of options for the graphs and create a label for the X and Y axes as follows:

- Bar direction
  - Horizontal
  - Vertical bars
- Bar color
  - Multi-color bars
  - Single-color bars

As shown in following figure, you may choose from the following graphing options to configure how graphs appear in the Governance Studio:



The data values across the defined Output tabs can be compared and displayed in the following types of charts:

Bar Chart
A 3-D, colorized visual effect that shows the data in a bar format. The orientation of the bars is determined by the selection of **Horizontal Graph Bars** or **Vertical Graph Bars** options. The use of color in this type of chart is defined using the **Multi-Color Graph Bars** (a different color for each Output tab) or **Single-Color Graph Bars** (one color for all Output tabs.)

Line Chart
A line connecting the various output data points in a simplified manner.

Scatter Chart
A series of connected markers that show the data relationship.

Pie Chart
A 3-D, colorized representation where each tab is depicted by its percentage of contribution to the total.

Stacked Chart
A 3-D, colorized representation where each tab is depicted by its percentage of contribution to the total.

Accumulate Graphing Data
The original data is retained when new data is added and all are rendered in the graphs.

Replace Graphing Data
All data is replaced by new data prior to rendering in a graph. The DSA is responsible for recalculating the results each time to ensure that the graph has the correct values.

X and Y-Axis Labels
You can add meaningful labels to indicate the data is graphed on the X and Y axes or you can leave it blank to use the default labeling.

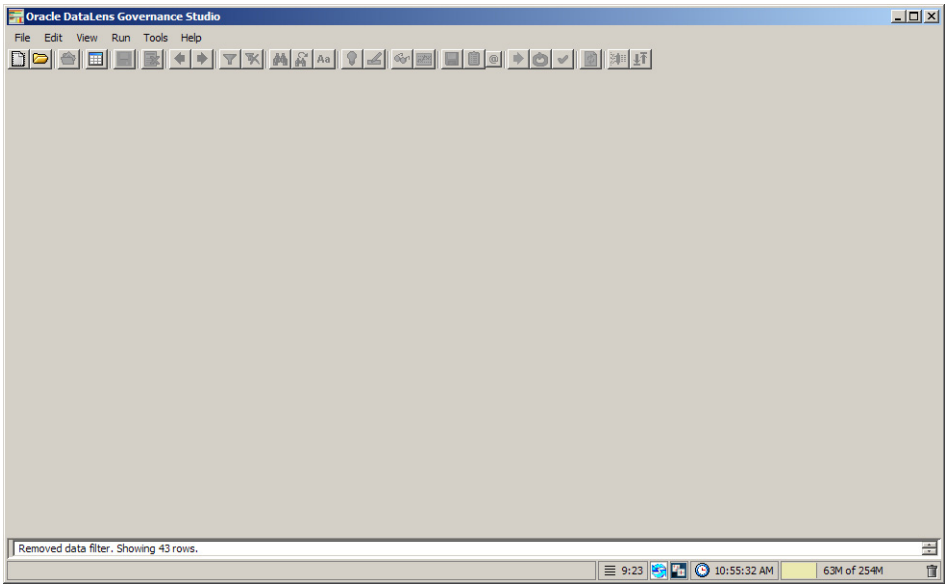
Help Button
Use the **Help** button to review information about the various graphing options.

Note: If you modify an output step in the DSA that underlies your Governance Studio project, you will need to check in your DSA changes and then refresh the Governance Studio project.

Starting the Governance Studio

Once you have added an output step in your DSA and configured it for Governance Studio output, you are ready to create a project in the Governance Studio. This section describes the steps to follow to create and run a Governance Studio project from a DSA created with the Application Studio.

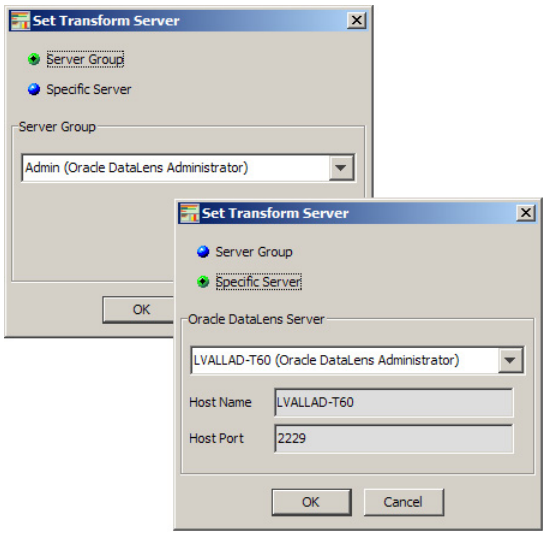
If this is the first time you have started the Governance Studio, the client workspace appears blank as in the following figure; otherwise the results from the last job run are displayed.



Setting the Transform Server

On first use of the Governance Studio GUI, a message is also displayed, requesting that you set the Transform Server to the location of the Oracle DataLens Server that will be running your Governance Studio projects.

1. 1. Click **OK**.



You can choose to select a specific server or a group of servers to use for data transformation.

2. Select either the **Server Group** or **Specific Server** option.
3. Click the down arrow to view your server options, and select the appropriate server.

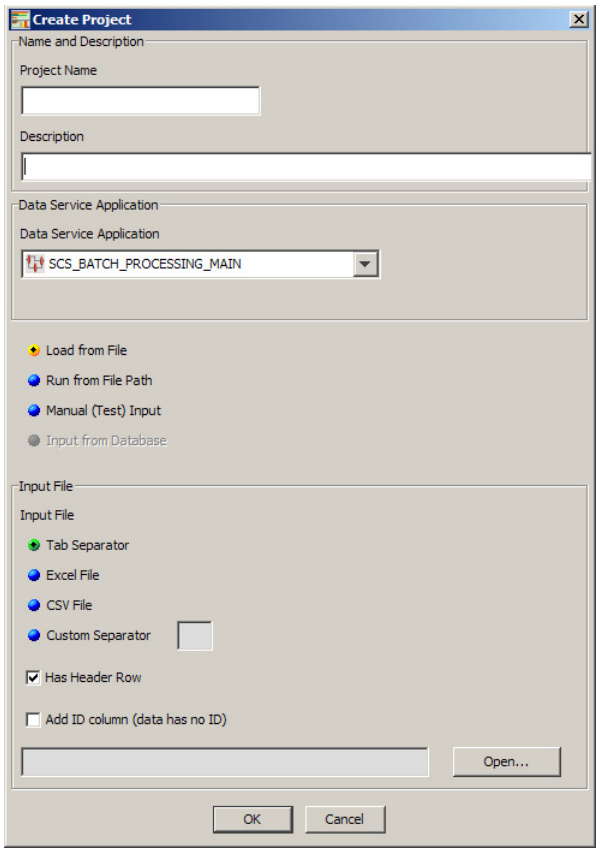
   If selecting a group of servers, you can choose from the Administrator, Development, Production, or Quality Assurance server groups. The Administrator server group is the default.
4. Click **OK**.

   The transform server is contacted and this server is used for all data transformations. You can use the **Set Transform Server** option from the **Tools** menu to change your Transform Server and host port at any time.

   The Governance Studio application start-screen is displayed.

Creating and Running a Governance Studio Project

From the **File** menu, select **New Project** to create a new Governance Studio project. The **Create Project** dialog box is displayed as shown in the following figure.



The **Create Project** dialog creates a new Governance Studio project. Following are the Governance Studio project parameters that you can configure:

Project Name
Name of the Governance Studio project.

Description
A description of the project. You must specify a description.

Data Service Application
Identifies the DSA that is run for the Governance Studio project and how the input data is received.

Input File
Specifies the type of input data.

Retrieve the Input Data

After you specify the project name and supply a description, you must select a DSA from the list of DSAs that have been checked in to your Oracle DataLens Server (for more information on creating DSAs, see *Oracle Product Data Quality Application Studio Reference Guide*). The input data that you specify must be compatible with the specifications of the DSA you created earlier. For more information, see Configuring a DSA for Governance Studio *on page25*.

To load the input data into the project, you must first select one of the data modes listed, and then the corresponding input data file type using one of the colorized

options. As you hover over a option, the color changes and when you click on it the selected option becomes green.

The options for receiving input data for the DSA are as follows:

- **Load from File**
- **Run from File Path**
- **Manual (Test) Input**
- **Input from Database**

The **Create Project** dialog box input data types options change based on your data mode selection. For example, if the input data is from a database, only the **Input from Database** selection is active and is selected by default.

The input data types supported are:

- **Tab Separator**
- **Excel File**
- **CSV File**
- **Custom Separator**

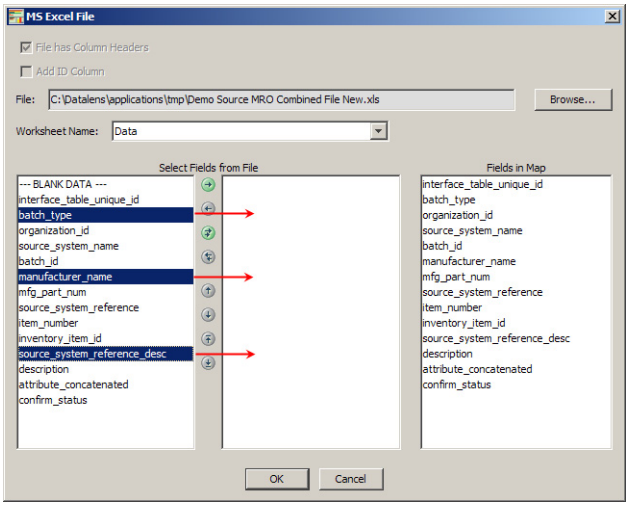
> **Note:** Count data must be in integer format to avoid generating an error.

The following section details an example of opening an input data file to retrieve and load the data.

Opening an Excel Input File

1. Select the **Load from File** option.
2. Select the **Excel File** option to specify an Excel file.
3. Click **Open…**.

   The **MS Excel File** dialog box is displayed.



4. Click **Browse**, locate the data file, and then double-click it.

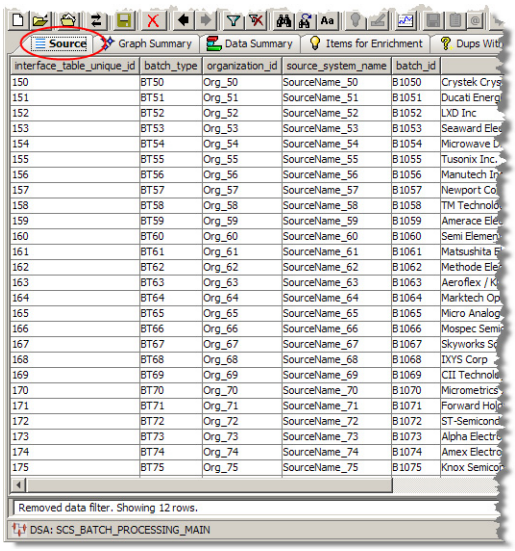
5. If the worksheet containing your data is not selected, select if from the **Worksheet Name** list.
6. Select the fields from the input file shown in the left pane that match the fields in the DSAs transform map as shown in the right pane and move them using the arrow buttons.

   The single arrow buttons move one field and the double arrow buttons move all fields to the left or to the right. You can reorder how the fields are processed by the DSA using the up and down arrows. The arrows that contain a line above or below it move the selected field to the top or bottom position respectively.

   The **-BLANK DATA-** selection may be used if there are no output fields associated with the input data in the DSA.
7. Click **OK** to accept your changes.

   You are returned to the Create Project dialog box.
8. Click **OK**.
9. Click the **Transform Data** button to process the data, which populates the **Summary** and Output tabs upon a successful run.

   Your new Governance Studio project is created and the **Source** tab of the Governance Studio console displays your input data as shown in following figure.



Note: The Governance Studio input data row limit is 64K rows for Excel 2003 and 1M for Excel 2007. Attempts to load input data files greater than this will result in a file input error. You should create several smaller input data files and process them separately.

The data file loading information, including the path to the file, is displayed in the status bar. Below the file information is the name of the DSA that will be used to process the input data.

Project Templates

The ability to create and subsequently open a project from a particular template can be very useful in ensuring consistent use of DSAs. Additionally, it allows privileged users to process multiple batches reusing the same template.

Creating a Project Template

1. To create a new project from an existing DSA template, from the **File** menu, click **New Project From…**. You are prompted for a project ID as shown in the following figure.
2. Enter a unique ID text string to identify this project and click **OK**.

   This unique text string is appended to the current DSA name so that it is clear which DSA was used to create the template.

   For example, if the DSA name is 'SCS_MAIN_PROCESSING' and you enter '1000999', then the template name would be 'SCS_MAIN_PROCESSING_1000999'.

Opening a Project Template

You can retrieve any template created using the **New Project From…** feature.

1. From the **File** menu, click **Open Project From…**.
2. Locate the template you want to use, select it, and then click **OK**.

   The project is opened, the data transformed, and the project is ready for use.

Opening Projects Without Results

You can open an existing project file *without* processing and displaying data results in any Output tabs. To open a project without data results, from the **File** menu, click **Open Project - No Results**, select the project you want to open, and then click **OK**.

> **Note:** If you created your project with version 5.5 or later, the **Data Summary** tab is populated as this data is saved with the project.

This feature is useful in opening a project that previously could not be opened because it exceeded the maximum memory allocation (you can monitor the **Memory Cache**, located in the client workspace frame, usage as it is dynamically updated). When that is the case, use this feature to open the project as previously described and then set the **Conserve Memory** and **Limits on Rows Retrieved** project options as appropriate so that you can process your data. For more information, see "Memory Tab" on page 2-12.

> **Note:** To avoid exceeding the allocated memory, you may need to adjust your client's OPDQ Java Web Start memory settings. For more information, see *Oracle Product Data Quality Oracle DataLens Server Administration Guide*.

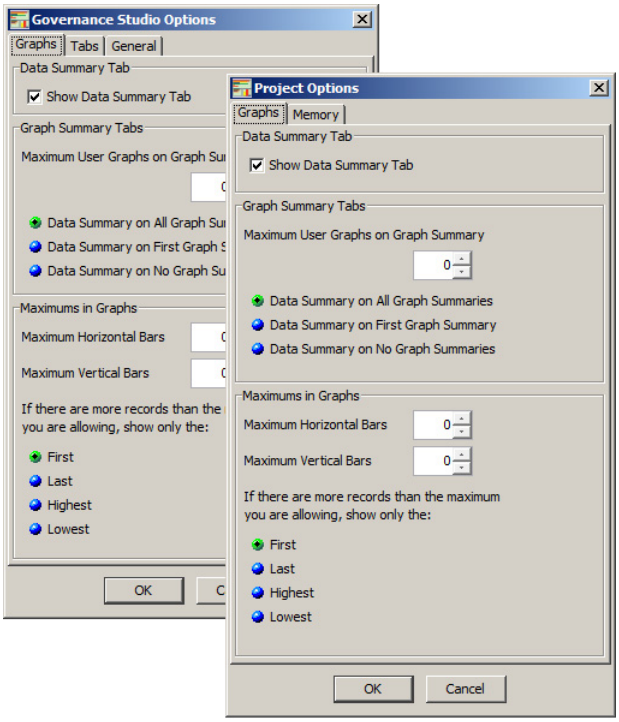
This option can not be used if the project was last run asynchronously though it can be used after you retrieve results using the **Open Project** command on the **File** menu.

Configuring Options

You can configure project options including how graphs and tabs are displayed with the two options menu selections on the **Tools** menu.

To set the options for the open project only, from the **Tools** menu, select **Project Options**.

To set the global options for all projects in the Governance Studio, from the **Tools** menu, select **Options**.



Graphs Tab

This tab is available for both **Governance Studio Options** and **Project Options**. It allows you to specify how graphs are displayed on the **Graph Summary** and **Data Summary** tabs.

The **Governance Studio Options** graphing settings are global and are the default settings when a new project is created. However, the **Project Options** graphing settings that you define for a project are used for that project *only* and are exported with the project so that they remain the same when imported by another user.

The following options can be configured:

Data Summary Tab Section

Select the **Show Data Summary Tab** check box to show the **Data Summary** tab.

Graph Summary Tab Section

Set the maximum number of graphs that will be displayed on the **Graph Summary** tab using the list.

Set the placement of the tab summary by selecting one of the following options:

- **it appears on all tabs**
- **it appears only on the first tab**

- **no graph summaries appear on any tabs**

Maximum in Graphs Section

Set the maximum number of horizontal and/or vertical bars using the lists.

Set which records are shown if the total number of records exceeds the maximum number of bars allowed.

Tabs Tab

This is an **Options** tab. It allows you to specify one or the other or both of the following to change the way that tabs display as follows:

Show all Tabs

All of the tabs are displayed rather than being hidden as described in "Task Panes" on page 1-15.

Show Tab Icons

The various icons for the types of tabs are shown on the tab adjacent to the name so that you can easily identify the function of the tab. For example, all Output tabs will display a green check mark or yellow question mark icon while all **Data Summary** tabs display a bar graph icon.

General Tab

This is an **Options** tab. Use the general options as follows:

Finish Job before Retrieval

When set, the time required to complete a job then retrieve and display the data in each Output tab is displayed in the **Status Field** for your review as in the following example:

```
Results Retrieved for job 212. Run=57 seconds. Retrieval=5 seconds.
```

The normal Governance Studio behavior of retrieving and displaying results before the job has completed is suspended when this option is selected.

Validate local Data Service Application

Allows you to validate the DSA that resides on a local drive rather than the same DSA that resides on the Oracle DataLens Server. The local DSA is the one you have checked-out to modify so this option ensures its validity before checking the DSA into the Oracle DataLens Server.

Additionally, the Governance Studio ensures that the local DSA mirrors the functionality of the Oracle DataLens Server DSA that is used to process the data.

Default Row Warning Count

Allows you to set a default number of source data rows beyond which the Governance Studio will display a warning message to allow the user to choose to retrieve all rows or the default number of rows specified. A warning message is displayed for *each* tab that has a resulting number of transformed rows (result count) higher than the specified warning count.

The default number of rows is set to 5000. Setting it to zero ensures that no warning messages are displayed. Setting a reasonable default can avoid potential memory issues when your project returns large sets of data.

This overall Governance Studio warning count option can be overridden for each Output tab by setting individual limits using the **Memory** tab of the **Project Options** as described in the next section.

Memory Tab

This is an **Project Options** tab. Use the options to set how the Governance Studio will use the available memory as follows:

Conserve Memory

Set this option to populate Output tabs only when the tab is selected thereby saving memory. When you run a job (or load it from disk), the data is retrieved and saved for each Output tab though it is not displayed until it is selected. With this option selected, your project will open and no results are displayed until you click an Output tab.

When using this option, the **Keep Lines Cutoff** option becomes active so that you can specify if data is to be retained in the selected Output tab when another one is selected.

> **Note:** You can not use the Re-run button (as described in "Re-Running a Secondary DSA" on page 3-3) to reprocess data when this option is set because all Output tabs must be fully populated so that the re-run operation can add data to tabs as necessary.

Keep Lines Cutoff

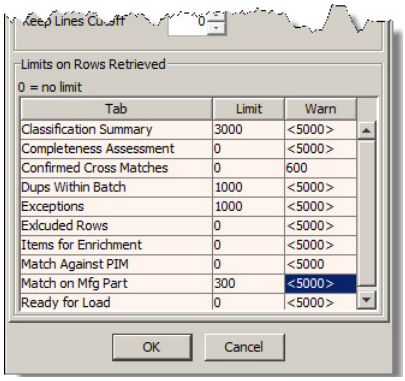
Set the amount of data to be retained in an Output tab when another tab is selected; this option is active after the **Conserve Memory** check box is selected.

Setting this value to:

- **zero (the default) -** ensures that all of the data is retained in the tab when another one is selected.
- **> zero -** removes the data from the tab if the rows of data exceeds the number specified. For example, if you set the value to 1000 rows and 1200 rows exist, then the data is removed when you select a different tab. Conversely, if only 900 rows exist, then the data remains because it did not exceed the 1000 row limit set.

Limits on Rows Retrieved

Set an individual value for the maximum number of rows that will be retrieved for each of the Output tabs. For example, 50,000 rows may be generated for an Output tab and you can choose to retrieve a maximum of 5000 to decrease the amount of time required to process the job.



Use the table as follows:

- **Tab column -** Each tab is listed.
- **Limit column -** Enter the maximum number of rows you want to retrieve for one or more tabs. Setting this value overrides the **Warn** column, locks it from editing, and no warning messages are displayed when the maximum is reached.
- **Warn column -** Use the default warning message value displayed or enter the number of source data rows beyond which the Governance Studio will display a warning message to allow the user to choose to retrieve all rows or the default number of rows specified. If you set this value then enter a **Limit** value for the tab, the **Warn** value is reset to the Governance Studio default as it is ignored by the override **Limit** value. You can reset this value by entering the Governance Studio default warning count surrounded by less than and greater than symbols. For example, if the default warning count was not changed you would enter'<5000>'.

These **Limit** and **Warn** values specified override any **Default Row Warning Count** value set as described in "General Tab" on page 2-11.

Tip: You can find out what the total number of generated rows by viewing the tooltip for a given tab without selecting it.

Note: To avoid exceeding the allocated memory, you may need to adjust your client's OPDQ Java Web Start memory settings. For more information, see *Oracle Product Data Quality Oracle DataLens Server Administration Guide*.

3

Run and Review a Project

This chapter explains how to run a Governance Studio project and use it to review your source data.

Running a Governance Studio Project

A Governance Studio job may be run synchronously or asynchronously. Synchronous execution runs in real time the foreground of the Governance Studio client workspace and is started by clicking on the job execution button. Program interaction is suspended until execution is completed when running a synchronous job. A synchronous execution runs in the background and releases the Governance Studio application to continue work. Large jobs should be run asynchronously though smaller jobs may be run using Synchronous execution.

Synchronous Execution

To run your Governance Studio job in real time (synchronously), click the **Run Synchronously** button on the toolbar or from the **Run** menu, select **Transform Data**.

Asynchronous Execution

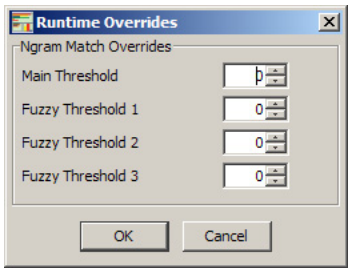
To run your Governance Studio job in the background (asynchronously), you would click the **Submit Job** button on the toolbar or from the **Run** menu, select **Submit Job**. This submits the job and closes your project.

Data Override Execution

If the DSA for your Governance Studio project contains an Ngram Match or an Attribute Match 2 (associated with Semantic Key 2), you can change the matching thresholds set in the DSA to use settings you provide for data transformation.

> **Note:** When the associated DSA does not contain these matching processes, this option is inactive.

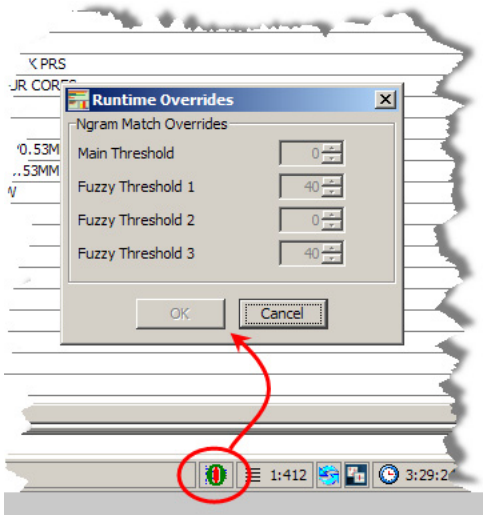
To run your Governance Studio with matching overrides, from the **Run** menu, select **Transform Data with Overrides**.



Use the **Runtime Overrides** dialog box to configure runtime only settings for the overall matching threshold and/or the fuzzy matching settings, and then click **OK**. Your data is immediately transformed using the runtime settings and the results are displayed.

For more information about the threshold settings, see *Oracle Product Data Quality Application Studio Reference Guide*.

Once your data transformation is complete, the **Override** button appears in the Governance Studio frame as in the following:



If you click **Override**, the override matching threshold settings are displayed for you to review and cannot be changed.

Reviewing Project Output

After running a Governance Studio project you can review the output and prepare some or all of the output for downstream processing. This includes sending alternatives to a web search interface to enable cross-sell or up-sell, or sending exceptions to knowledge engineers for appropriate modifications to data lenses to improve recognition and classification of the data.

You may also use the AutoSuggest feature to have the Governance Studio automatically suggest attributes that may have been missed due to misspellings or abbreviations. Autosuggested data may then be transferred to an Output tab table for further processing.

This section provides some examples of the Governance Studio output that extracts attributes and provides classification statistics and suggestions.

Depending on the type of output configured in the underlying DSA, you can take further action based upon your results. Each output type is indicated by a tab in the Governance Studio project.

General Processing

Rows that have been selected in an Output tab may be 'checked' (selected) to continue processing based on the following actions designated in the underlying DSA output step:

Action	Toolbar Button
Process (Apply DSA)	
Reprocess (Re-run DSA)	
Complete (Completion DSA)	
Quick Lookup (Lookup DSA)	

The primary DSA that is used to process the source data is displayed in the **Status Field**. When a secondary DSA will be used to process the data, it is displayed in parenthesis in the tooltip of that processing button so that you can view it before processing your data.

Applying a Secondary DSA

Once data has been reviewed, the set of data can be sent to another DSA to continue the processing and to generate text output. This text output can be inserted into a database, emailed, or saved as a file.

1. To apply a secondary DSA to a set of data, select a set of rows on an Output tab using the check mark column.
2. From the **Run** menu, select **Apply Checked Rows** or click the **Apply** button on the toolbar.

Re-Running a Secondary DSA

Similar to applying a secondary DSA to process a set of reviewed data, the Re-Run feature not only applies the DSA it automatically returns the results to the current project and merges it into existing output data. Once the results are merged, the data selected for processing is then deleted from the main DSA output.

1. To apply a secondary DSA to a set of data, select a set of rows on an Output tab using the check mark column.

> **Note:** All data selected for re-run is included in a re-run operation so all of the source data in the top task pane and the matching results in the bottom task pane are reprocessed.

2. From the **Run** menu, click **Re-run Checked Rows** or click the **Re-run** button on the toolbar.

Note: Re-run is not available (dimmed) if you have selected the **Conserve Memory** option. For more information, see "Configuring Options" on page 2-10.

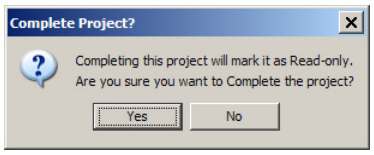
Completing Projects

You can associate a current DSA with a completion DSA to set a specific project run to a completed status. The completion DSA can include project closing tasks and functions necessary a final closeout of the project.

Once the project completion (closing) has occurred, the project is set to a read-only state that does not allow any further changes. The project can be used to create a new project using the **New Project From…** option though all other application functionality used to modify or manipulate data is made inactive (dimmed).

1. To close a project, click the **Release Batch** tab.
2. Locate the Release Batch ID that you want to close and click the adjacent check box.
3. From the **Run** menu, click **Complete Project** or the **Complete Project** button on the toolbar.

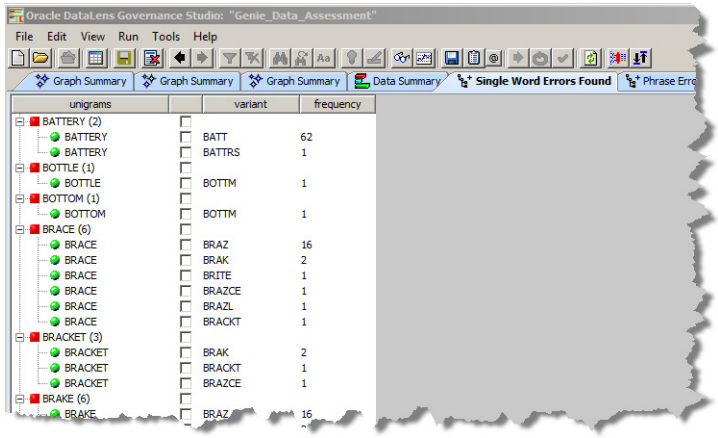
   A verification dialog is displayed as shown in the following figure.



4. If you are sure that you want to close this project, click **Yes**; otherwise, click **No** to cancel the process.

Applying Quick Lookup

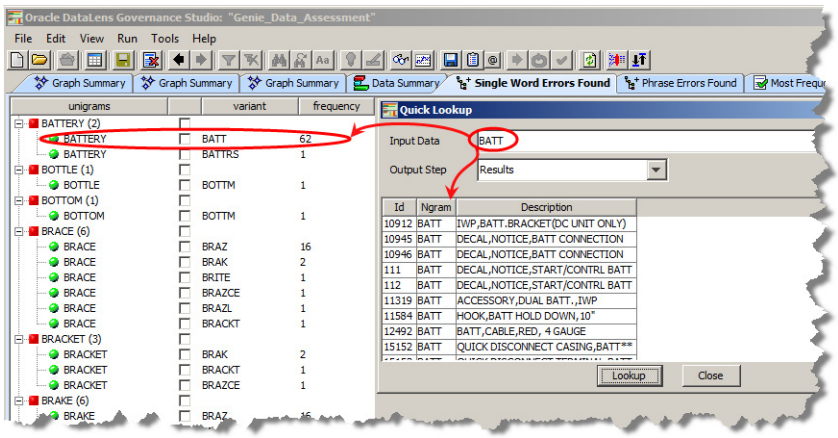
A Quick Lookup DSA is assigned to the output step of a DSA that uses Ngram matching to allow you to show a variant in the context of the original description. The results are displayed in the Governance Studio on the Output tab for the Ngram output step as in the following sample tab:



An output tab can contain a set of variants for either unigrams, bigrams, or trigrams along with the number of times (frequency) the variant is found in the input data.

1. To lookup the associated input data information for a given Ngram, select it and click the **Quick Lookup** button on the toolbar.

   In the following example, the 'BATT' variant of the 'BATTERY' unigram is selected for a quick lookup of this variant in the input data:



   The 62 instances of the 'BATT' variant are displayed in the **Quick Lookup** dialog box.

2. You can enter a different variant to lookup for the selected Ngram in the **Input Data** field and click **Lookup**. Additionally, you can select a different DSA **Output Step** to use the Quick Lookup functionality on. Either of these options can be used repeatedly to locate result data.

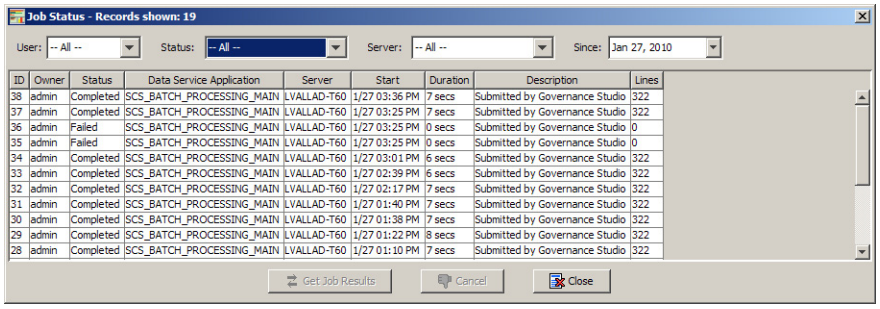
3. Click **Close** to close the **Quick Lookup** dialog box.

Other Features

This section describes other features that you can use to review data in your project.

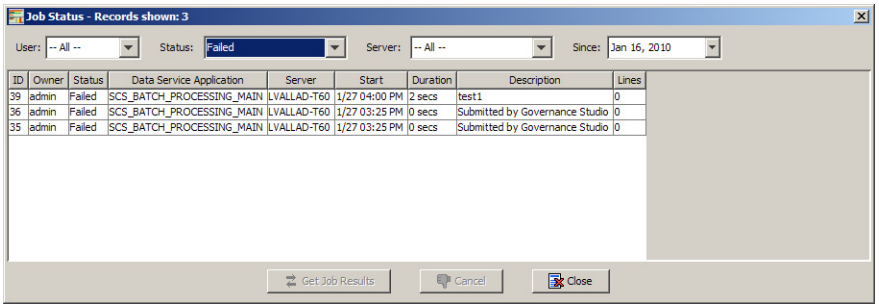
Viewing Job Status

You can view the status of all jobs associated with a DSA at one time using the **View Job Status** option on the **View** menu.



Any job that was created within the Governance Studio or the Oracle DataLens Server is displayed including those that are pending or have failed. By default, only the current user's jobs are listed.

The view can be changed using the **User**, **Status**, **Server**, and **Since** lists. These controls filter the displayed data based on your selections. For example, you could use the **Since** list to change the date that you want to set as the starting point for data retrieval. Alternatively, selecting **Failed** from **Status** filters all failed jobs from the data retrieved.



If a job is running, you can select it and click **Cancel** to stop it. For those jobs that output a result file, you can obtain those results by selecting the job from the list and clicking **Get Job Results**, which is described in the following section.

Getting Job Results

Any job that is run from the Oracle DataLens Server whose output is designated for display in the Governance Studio can be used to create a new project.

Caution: When a DSA returns the results of a processing job (synchronous or asynchronous), you *must* retrieve or delete the data immediately, particularly for those jobs that return large amounts of data. Failure to do so can result in an Oracle DataLens Server crash once the server memory is exhausted. The memory can be exhausted from many small jobs that are not retrieved because each DSA job that is under the DSA maximum memory or chunk size set for the server is held in memory until it is retrieved or the server is restarted. For information about setting these values, see *Oracle Product Data Quality Oracle DataLens Server Administration Guide*.

To view job run results and create a new project based on one, use these steps:

1. From the **View** menu, select **View Job Status** as described in the previous section.

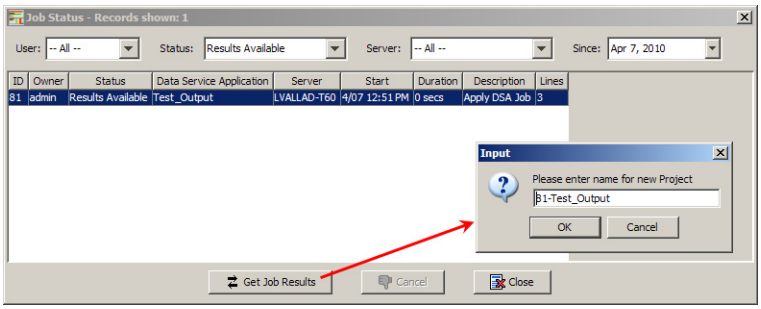
   The list of available jobs is displayed.

2. From the **Status** list, select **Results Available**. The Results Available status is only for completed jobs.

3. Select the job whose results you want.

   Once you select a job that has results available, the **Get Job Results** button is activated.

4. Click **Get Job Results**.



You are prompted for a name for the project. You can either accept the current name or enter your own name. The **Input** dialog box is populated with a name for the new project based on the job ID and the DSA name by default.

5. Enter a different name for the new project or use the default name provided, and then click **OK**.

The new project is created and the data transformed.

Note: If you view job results while a project is open, it is closed before the new project is created and you are prompted to save any changes.

Emailing Checked Rows

This feature allows you to email only rows that have been selected by check marks; it does not save selected rows. For example, those rows that are not applicable to the **Source** tab. The checked rows are automatically saved to an Excel spreadsheet, with the file name created using the name of the tab, in the DataLens directory where the project resides. You can change the file name or save location in the dialog provided.

Exporting and Importing Projects

The ability to export a project that contains all of the associated source input data, DSA, and modifications is a powerful tool. It allows you to share a complete and cohesive project with others when they import it into the Governance Studio so that identical viewing and reporting can be achieved by multiple people.

Exporting a Project

1. Open the project you want to export.
2. Ensure that you have completed all project refinements and have transformed all data as appropriate.
3. From the **File** menu, select **Export Project**.

   A file save dialog appears with the current project name populated as the file name with a `.ams` file extension.
4. Use the file name and directory provided to save the file in or enter them, and click **OK**.

The project with its data is exported into identified file name and directory for you to share with others.

Importing a Project

1. From the **File** menu, select **Import Project**.

   A file open dialog appears.

2. Locate and select the project file that you want to import. The file extension must be `.ams`.

3. Click **OK**.

Any open projects are closed and the selected project is opened.

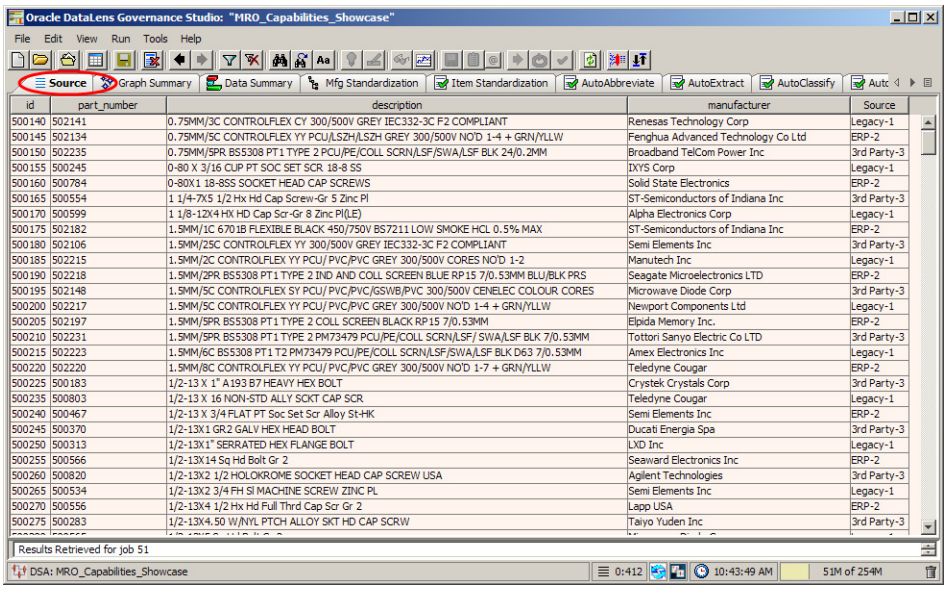
4

Basic Tabs

This chapter describes the basic types of tabs that display in Governance Studio.

Source Tab

The **Source** tab in any existing Governance Studio project displays the input data that can be processed.



id	part_number	description	manufacturer	Source
500140	502141	0.75MM/3C CONTROLFLEX CY 300/500V GREY IEC332-3C F2 COMPLIANT	Renesas Technology Corp	Legacy-1
500145	502134	0.75MM/5C CONTROLFLEX YY PCU/LSZH/LSZH GREY 300/500V NO'D 1-4 + GRN/YLLW	Fenghua Advanced Technology Co Ltd	ERP-2
500150	502235	0.75MM/5PR BS5308 PT1 TYPE 2 PCU/PE/COLL SCRN/LSF/SWA/LSF BLK 24/0.2MM	Broadband TelCom Power Inc	3rd Party-3
500155	500245	0-80 X 3/16 CUP PT SOC SET SCR 18-8 SS	IXYS Corp	Legacy-1
500160	500784	0-80X1 18-8SS SOCKET HEAD CAP SCREWS	Solid State Electronics	ERP-2
500165	500554	1 1/4-7X5 1/2 Hx Hd Cap Screw-Gr 5 Zinc Pl	ST-Semiconductors of Indiana Inc	3rd Party-3
500170	500599	1 1/8-12X4 HX HD Cap Scr-Gr 8 Zinc Pl(LE)	Alpha Electronics Corp	Legacy-1
500175	502182	1.5MM/1C 6701B FLEXIBLE BLACK 450/750V BS7211 LOW SMOKE HCL 0.5% MAX	ST-Semiconductors of Indiana Inc	ERP-2
500180	502106	1.5MM/25C CONTROLFLEX YY 300/500V GREY IEC332-3C F2 COMPLIANT	Semi Elements Inc	3rd Party-3
500185	502215	1.5MM/2C CONTROLFLEX YY PCU/ PVC/PVC GREY 300/500V CORES NO'D 1-2	Manutech Inc	Legacy-1
500190	502218	1.5MM/2PR BS5308 PT1 TYPE 2 IND AND COLL SCREEN BLUE RP15 7/0.53MM BLU/BLK PRS	Seagate Microelectronics LTD	ERP-2
500195	502148	1.5MM/5C CONTROLFLEX SY PCU/ PVC/PVC/GSWB/PVC 300/500V CENELEC COLOUR CORES	Microwave Diode Corp	3rd Party-3
500200	502217	1.5MM/5C CONTROLFLEX YY PCU/ PVC/PVC GREY 300/500V NO'D 1-4 + GRN/YLLW	Newport Components Ltd	Legacy-1
500205	502197	1.5MM/5PR BS5308 PT1 TYPE 2 COLL SCREEN BLACK RP15 7/0.53MM	Elpida Memory Inc.	ERP-2
500210	502231	1.5MM/5PR BS5308 PT1 TYPE 2 PM73479 PCU/PE/COLL SCRN/LSF/ SWA/LSF BLK 7/0.53MM	Tottori Sanyo Electric Co LTD	3rd Party-3
500215	502223	1.5MM/6C BS5308 PT1 T2 PM73479 PCU/PE/COLL SCRN/LSF/SWA/LSF BLK D63 7/0.53MM	Amex Electronics Inc	Legacy-1
500220	502220	1.5MM/8C CONTROLFLEX YY PCU/ PVC/PVC GREY 300/500V NO'D 1-7 + GRN/YLLW	Teledyne Cougar	ERP-2
500225	500183	1/2-13 X 1" A193 B7 HEAVY HEX BOLT	Crystek Crystals Corp	3rd Party-3
500235	500803	1/2-13 X 16 NON-STD ALLY SCKT CAP SCR	Teledyne Cougar	Legacy-1
500240	500467	1/2-13 X 3/4 FLAT PT Soc Set Scr Alloy St-HK	Semi Elements Inc	ERP-2
500245	500370	1/2-13X1 GR.2 GALV HEX HEAD BOLT	Ducati Energia Spa	3rd Party-3
500250	500313	1/2-13X1" SERRATED HEX FLANGE BOLT	LXD Inc	Legacy-1
500255	500566	1/2-13X14 Sq Hd Bolt Gr 2	Seaward Electronics Inc	ERP-2
500260	500820	1/2-13X2 1/2 HOLOKROME SOCKET HEAD CAP SCREW USA	Agilent Technologies	3rd Party-3
500265	500534	1/2-13X2 3/4 FH SI MACHINE SCREW ZINC PL	Semi Elements Inc	Legacy-1
500270	500556	1/2-13X4 1/2 Hx Hd Full Thrd Cap Scr Gr 2	Lapp USA	ERP-2
500275	500283	1/2-13X4.50 W/NYL PTCH ALLOY SKT HD CAP SCRW	Taiyo Yuden Inc	3rd Party-3

From the **Source** tab you can select one or multiple rows to process using the DSA. Rows are selected using the following methods:

All rows in the table

Use **Ctrl+a**.

Discontinuously rows

Press and hold **Ctrl** and click on each row to select.

Group of adjacent rows

Press and hold **Shift** then click the first row and then the last in the group you want to select.

For more information, see "Running a Governance Studio Project" on page 3-1.

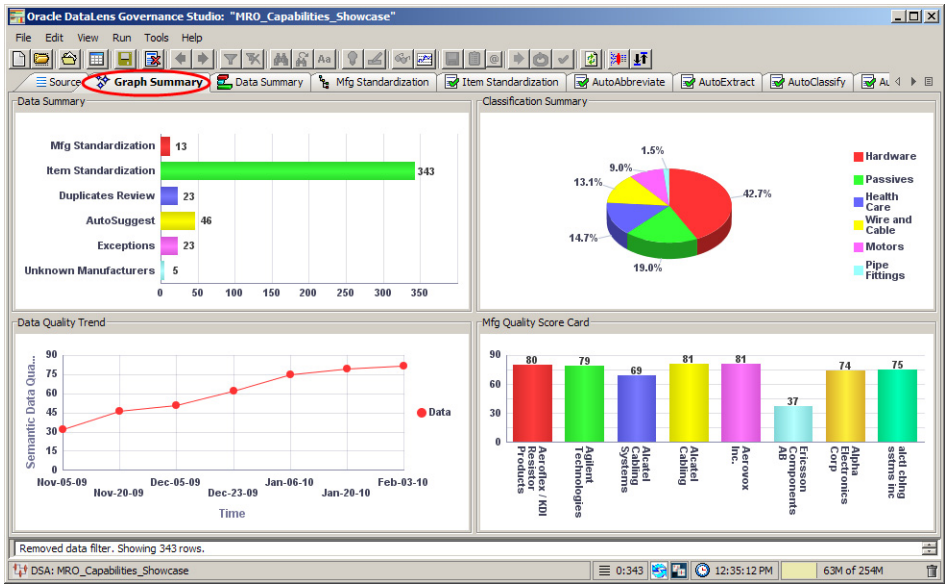
The fields in the table, with the exception of those in the unique id column, can be edited. You can double-click in a field, edit it, and then either press **Enter** or click outside the edited field. Changed fields are highlighted green so that you can easily identify them before processing. The edited data is available for further processing.

> **Tip:** When you select rows, you can copy them to the clipboard using Ctrl+c. In another Windows application, such as Excel, you can then paste the data into the open file using Ctrl-v.

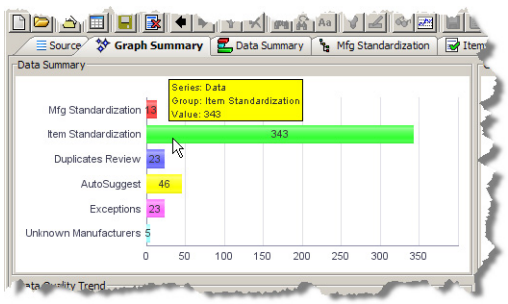
Graph Summary Tabs

The **Graph Summary** tabs appear in the list of tabs only if an output step in your DSA has a DGS output type of **Summary Data for Graphing** defined. Depending on your DSA, it is possible that there are too many graphs to appear on one page in the Governance Studio.

You can define the number of graphs you want to appear (see "Configuring a DSA for Governance Studio" on page 2-1.) Additional Graph Summary tabs are created for additional graphs until all of the graphs have been accounted for.

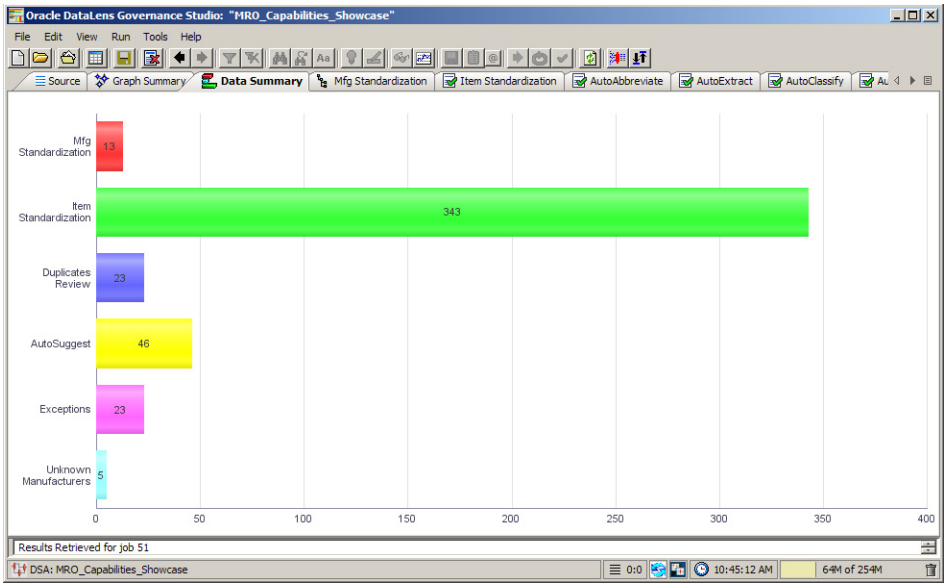


When you hover your cursor over any of the bars on this tab, a tooltip describes the information shown in that bar. You can double-click on the bar in the graph to activate the Output tab (if there is a tab associated with that graph data) that corresponds to the graphed data.



Data Summary Tab

The **Data Summary** tab displays a bar chart of the number of rows contained in each of the output data tabs. Unlike the **Graph Summary** tabs, there can be only one **Data Summary** tab. It appears by default unless you disable the option (see "Configuring a DSA for Governance Studio" on page 2-1.) The **Data Summary** tab graph shows you at a glance the number of rows of data you have in each of the Output tab tables. For example, if an Output tab table contains exception data and the number is very large relative to the other output steps, you would be able to identify this problem easily. Double-clicking on any bar in the graph immediately activates the Output tab table as identified to the left of the bar.



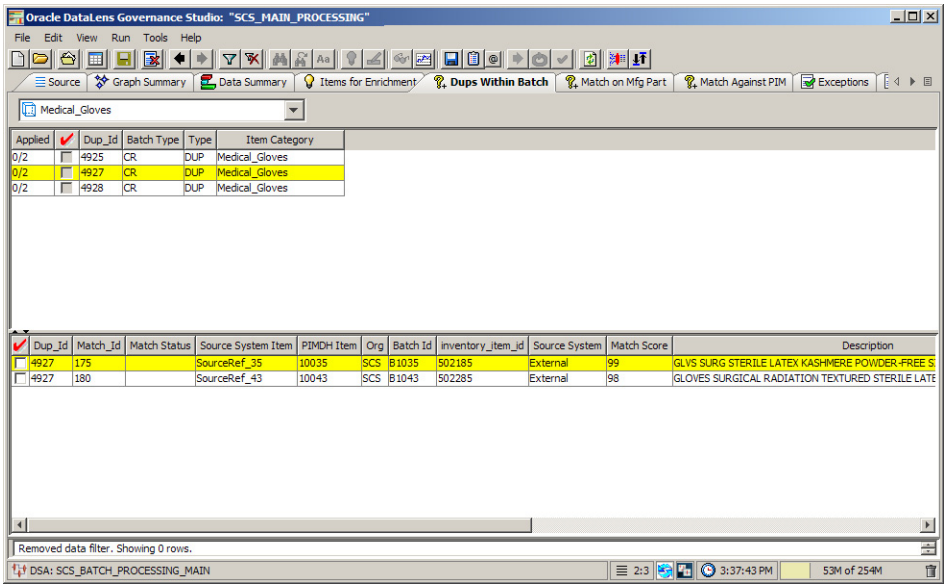
5

DSA Output Tabs

This chapter explains how to use the Output tabs created by your DSA so that you can review your data.

Review, Approve and Route Output Tabs

The “Review, Approve and Route” Output tabs are identified by the green check mark icons. These types of tabs contain tables of output data and correspond to a DSA output step with the output type of 'Review, Approve, and Route'. Data in these tabs may be selected for output to a file, edited, emailed to others, or run as input to a synchronous job. The following figure shows an example of a “Review, Approve and Route” Output tab.



Output Tab Functions

The following sections describe how to use Output tabs.

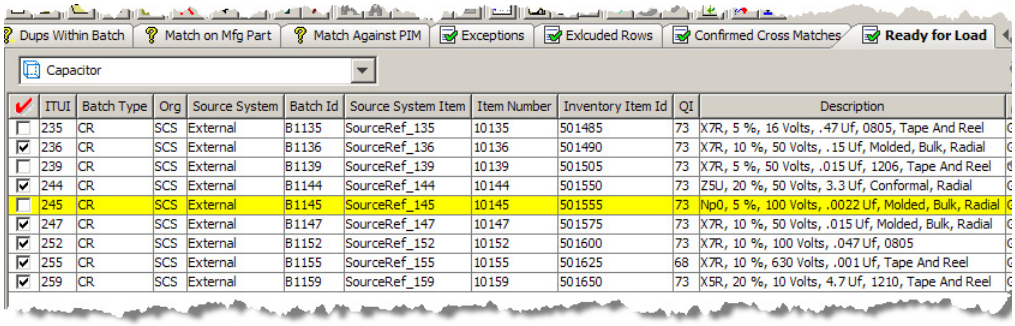
Row Selection

You can select a single row or multiple rows for additional processing using the following methods:

All rows in the table
Use **Ctrl+a**.

Discontinuously rows
Press and hold **Ctrl** and click on each row to select.

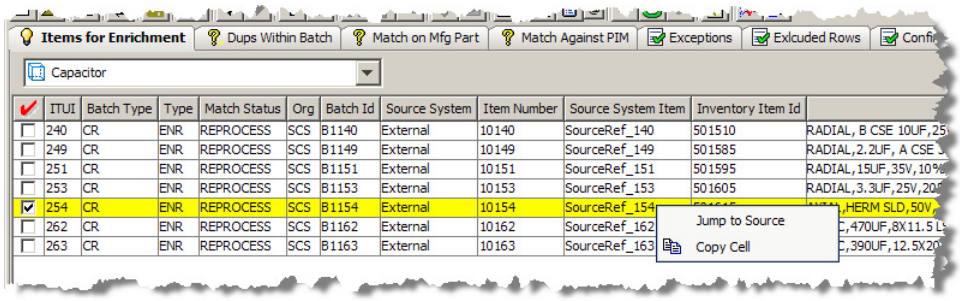
Group of adjacent rows
Press and hold **Shift** then click the first row and then the last in the group you want to select.



	ITUI	Batch Type	Org	Source System	Batch Id	Source System Item	Item Number	Inventory Item Id	QI	Description
☐	235	CR	SCS	External	B1135	SourceRef_135	10135	501485	73	X7R, 5 %, 16 Volts, .47 Uf, 0805, Tape And Reel
☑	236	CR	SCS	External	B1136	SourceRef_136	10136	501490	73	X7R, 10 %, 50 Volts, .15 Uf, Molded, Bulk, Radial
☐	239	CR	SCS	External	B1139	SourceRef_139	10139	501505	73	X7R, 5 %, 50 Volts, .015 Uf, 1206, Tape And Reel
☑	244	CR	SCS	External	B1144	SourceRef_144	10144	501550	73	Z5U, 20 %, 50 Volts, 3.3 Uf, Conformal, Radial
☐	245	CR	SCS	External	B1145	SourceRef_145	10145	501555	73	Np0, 5 %, 100 Volts, .0022 Uf, Molded, Bulk, Radial
☑	247	CR	SCS	External	B1147	SourceRef_147	10147	501575	73	X7R, 10 %, 50 Volts, .015 Uf, Molded, Bulk, Radial
☑	252	CR	SCS	External	B1152	SourceRef_152	10152	501600	73	X7R, 10 %, 100 Volts, .047 Uf, 0805
☑	255	CR	SCS	External	B1155	SourceRef_155	10155	501625	68	X7R, 10 %, 630 Volts, .001 Uf, Tape And Reel
☑	259	CR	SCS	External	B1159	SourceRef_159	10159	501650	73	X5R, 20 %, 10 Volts, 4.7 Uf, 1210, Tape And Reel

Copying Data

You may right-click on any row in a Governance Studio table and copy the data in the cell that you are hovered over to the clipboard. Data that has been copied into the clipboard may then be pasted into another application or into a different Output tab table in the Governance Studio.



	ITUI	Batch Type	Type	Match Status	Org	Batch Id	Source System	Item Number	Source System Item	Inventory Item Id	
☐	240	CR	ENR	REPROCESS	SCS	B1140	External	10140	SourceRef_140	501510	RADIAL, B CSE 10UF,25
☐	249	CR	ENR	REPROCESS	SCS	B1149	External	10149	SourceRef_149	501585	RADIAL,2.2UF, A CSE
☐	251	CR	ENR	REPROCESS	SCS	B1151	External	10151	SourceRef_151	501595	RADIAL,15UF,35V,10%
☐	253	CR	ENR	REPROCESS	SCS	B1153	External	10153	SourceRef_153	501605	RADIAL,3.3UF,25V,20
☑	254	CR	ENR	REPROCESS	SCS	B1154	External	10154	SourceRef_154		,HERM SLD,50V,
☐	262	CR	ENR	REPROCESS	SCS	B1162	External	10162	SourceRef_162		C,470UF,8X11.5 L
☐	263	CR	ENR	REPROCESS	SCS	B1163	External	10163	SourceRef_163		C,390UF,12.5X20

For more information about context-sensitive menus in task panes, see "Top Pane Functions" on page 5-7.

Modifying Cell Contents

If the cell has not been designated as a read only cell in the DSA Output tab, double-clicking a cell in any of the Output tab tables enables one of the following actions:

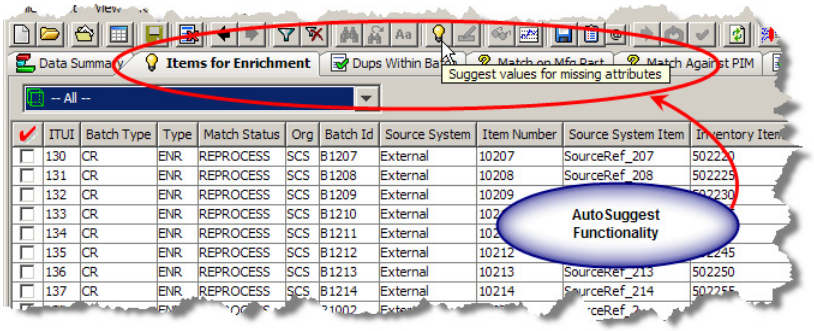
- The cell can be edited. A text insertion cursor appears, the background of the cells turns to white, and you can enter or change the text in the cell.

> **Note:** The ability to edit a field is configured in the DSA in the Application Studio. If a field cannot be edited and you want that ability, you can change the behavior by adding two exclamation points (!!) to the **Edit Value** output column of that output step.

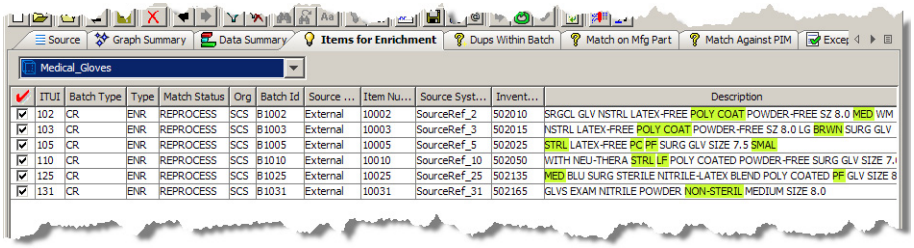
- A list appears and you can select one of the values predetermined by the DSA. The value, "<empty>", allows you to invalidate (or null) the selected value from the list.
- Disables a suggestion when in the AutoSuggest table, which changes the color to pink. Double-clicking it again re-enables the suggestion, which turns the highlight back to green. You can choose the suggestions you want to keep and those to discard. For more information, see "Review, Approve and Route Output Tabs" on page 5-1.

Review, Approve, and Route AutoSuggest Output Tabs

The "Review, Approve and Route" tab with AutoSuggest output is identified by the light bulb icon. Data in these tabs may access the Governance Studio AutoSuggest feature; this feature automatically predicts attributes that may have been missed due to misspellings and abbreviations or missing grammar. When this tab is not empty and there are suggested values for missing attributes available, the light bulb button is active. If no suggestions were found, this button is inactive (dimmed). The following figure shows an AutoSuggest table where the suggestion button is available.



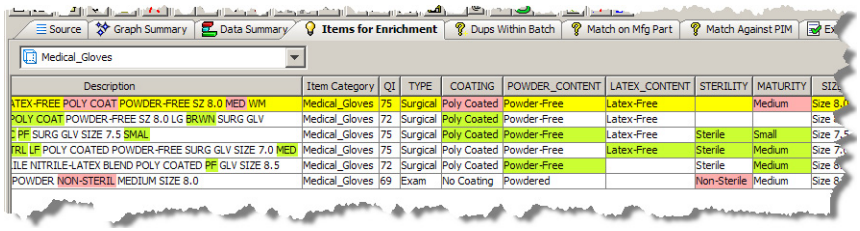
After clicking on an AutoSuggest Output tab, a table of data designated for output by the respective output step in the DSA is displayed. To view suggestions, click on the **AutoSuggest light bulb** button or from the **Run** menu, select **AutoSuggest**. Suggestions are now highlighted in green as shown in the following figure.



AutoSuggest Feature

There are several operations that are allowed in the AutoSuggest table to process your data further and make it easier to edit, move, copy or accept suggestions. In addition to the normal table functions available to Output tab tables, the AutoSuggest table allows you to accept or reject suggestions by double-clicking the cell you want to reject.

The following figure shows the results of double-clicking a suggested cell in the **Maturity** column. The cell is now highlighted in pink showing that the attribute has been rejected; you can return the cell to accept by double-clicking the cell again.

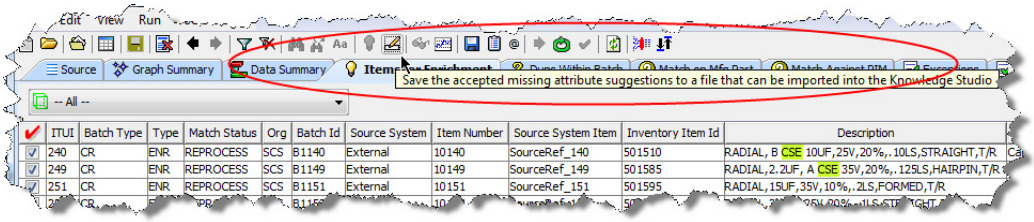


When you have completed analyzing suggestions by rejecting the bad ones and keeping the good ones you may select rows you want to complete the enrichment process by clicking the check box to the left of the row to select it. Once all of the rows you to be enriched are selected, click the **Re-run** button to send the checked rows to a secondary DSA process to complete the process. The checked rows move from the AutoSuggest tab if the enriched attributes and/or additions to the description increase the **QI** to the point where the row(s) no longer require enrichment.

Apply Augmentations Feature

You can add or augment the knowledge in your data lens using this feature. Once you have used the AutoSuggest feature to identify matching text in your data further, you can add these additional matches to your data lens so that they are automatically recognized in subsequent processing. For example, 'STRL' is a variation of 'Sterile' and adding this variant form would increase the knowledge in your data lens therefore increasing the quality index.

Using the Apply Augmentations feature creates missing attribute suggestions and this knowledge is then written to an Excel file. The Apply Augmentations button, which is a pencil on paper, is active only after using AutoSuggest otherwise this button is inactive (dimmed).



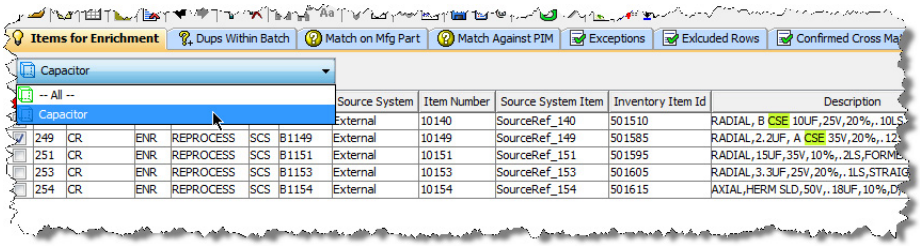
You can review and modify this knowledge before importing the data in the Excel file into your data lens in the Knowledge Studio.

Caution: If you intend on both applying Augmentations and using a Re-Run secondary DSA to continue processing the rows, you must Apply Augmentations **first**.

When you process the Re-Run first, the rows and the included Augmentations are removed from the tab and no longer available to process the Augmentations.

Use the following steps to augment attribute variants in your data lens:

1. Select an AutoSuggest Output tab.
2. Click the **AutoSuggest** button to identify variants.
3. Select **All** or a single Item Definition for which you want to add attribute variant knowledge. If you select a single Item Definition, only knowledge related to that Item Definition is created; selecting **All** creates knowledge for all of the Item Definitions that have suggestions.



4. Select the rows that contain the variants you want to add.
5. Click the **Apply Augmentations** button or from the **Run** menu, select **Apply Augmentations**.
6. Locate the directory you in which to save the Excel file and click **Save**. The file is automatically named using the following convention:

   `Autosuggest-JobID-DataLensName-ItemDefinition.xls`
7. Open the Excel file to review and modify (if necessary) the new knowledge.
8. Open the corresponding data lens that was identified as part of the file name created in the previous step in the Knowledge Studio.
9. From the **File** menu, select **Import Phrases and Terms**.
10. Locate the AutoSuggest file, select it, and then click **Open**.

Review the effected attributes; the augmentations are denoted with a light bulb icon in the hierarchical tree.

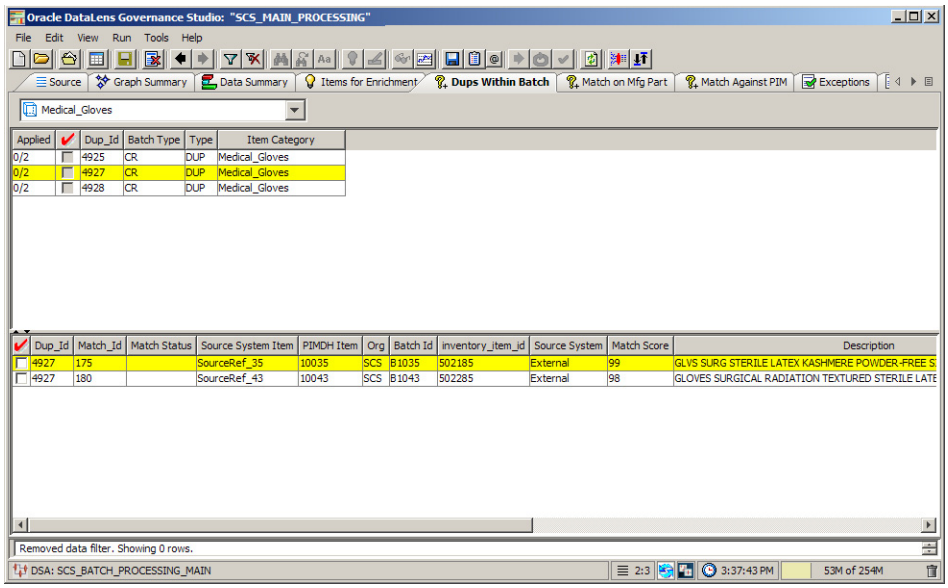
Match Results Tabs

A matching application is a set of two DSAs that you use to find appropriate data records that match pre-specified criteria. A matching process is built upon a data lens (or set of data lenses), used to recognize items by their attributes and to rank attributes in order of importance.

The output of the matching process is used in a DSA output step. There are two match result output types variations available for you to create in your DSA:

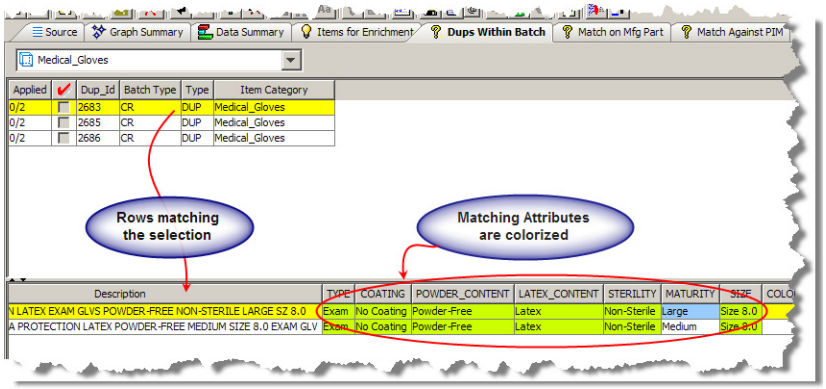
- Single Selection Review, Approve, and Route
- Multiple Selection Review, Approve and Route

If you have created a match results output step in your DSA, then the corresponding tab appears in your Governance Studio project. The following figure shows a project with a sample match result set.



The Output tabs that contain data duplicates display a parent-child view of the match results. In a New Item Introduction use case, the system can match the new item against any existing items in your Item Master. The inbound row will be classified, its key attributes extracted and standardized, and a Semantic Key is created. This Semantic Key is used to find any potential duplicate rows in the Items Master. If matches for the incoming row are found in the item master, the system can route the row to a potential duplicates Output tab and give you a parent child view of the new record on the top pane, and the potential matches in the bottom pane.

The split screen function is only used for match result Output tabs. Once a row is selected from the top pane, the corresponding matched candidates for that row on the bottom pane are highlighted as in the following figure.



The Governance Studio output is color coded in order to highlight key attributes participating in the match. Additionally, the attributes are colorized based on the type of match rule that has been triggered on the attribute. Attributes are colorized as follows:

Green Attributes

Match the parent row and were set to required in the item specific match rules.

Blue Attributes
Match the parent row and were set to participating, but not required in the item-specific match rules.

Blank and un-highlighted
Indicates that the matched row did not contain a value for that particular attribute so that attribute did not participate in the match score.

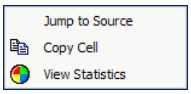
Populated and un-highlighted
Indicates that the matched row contains a value for that attribute, but that the attribute value for the matched row did not match the attribute value for the parent row.

Top Pane Functions

You can use the context-sensitive menu by right-clicking a data record and use the options as follows:

Jump to Source

Activate the **Source** tab with the same selected output data selected. For example, this allows you to view the ID if it were not included in the output.

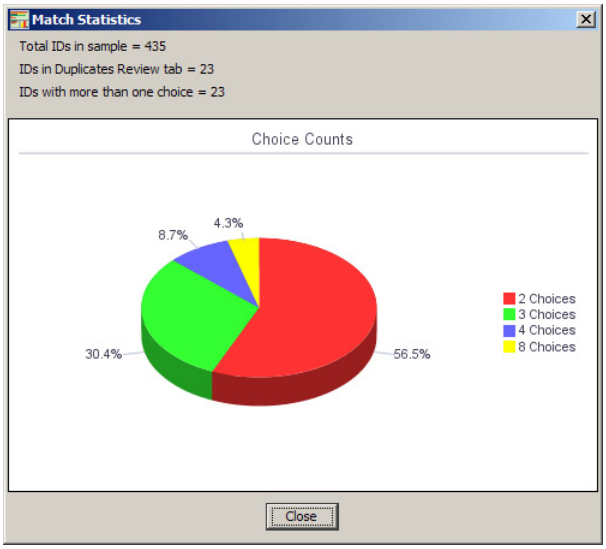


Copy Cell

Copy the cell for downstream processing in a data lens or to another worksheet or file.

View Statistics

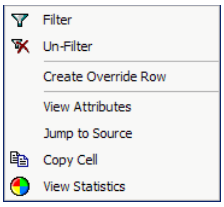
Displays number of choices per record; use this option for Match Process results to produce a graph of the breakout of rows with matching records as in the following figure:



Bottom Pane Functions

Similar to the top pane functionality, you can right-click on a matched row to use the following additional context-sensitive menu options.

Filter



Allows you to filter the displayed data based on text or a text pattern. The filter operation applies only to the currently selected tab. Only the rows that match the text entered in the **Filter** dialog are displayed in the table. For more information about filtering, see "Filtering Data" on page 1-14.

Un-Filter

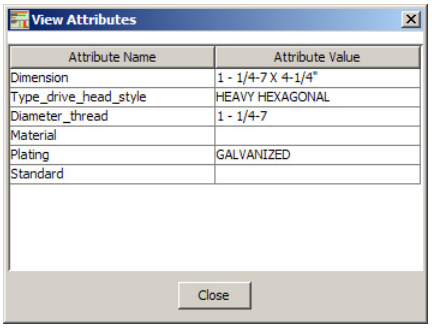
Removes the filter applied to the selected tab so that all of the data is displayed.

Create Override Row

This functionality is detailed in the Survivorship Process section.

View Attributes

Provides a table view of the attributes for the selected row as in the following figure.



Working with a Match Set

The Governance Studio interprets a set of results and populates match results tabs based on:

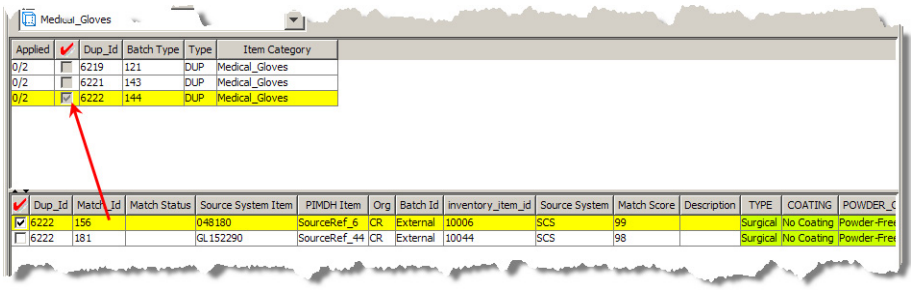
- the DSA options that were set indicating a matched set, and
- the accuracy of the attribute naming, order and structure.

You review the rows and determine the matching rows based on the defined use case as in the examples in the following sections.

Request for Quote

You select a single match item for each requested item to be quoted and the Governance Studio identifies the best record to for use in the quote based on the information in the data lens and other industry related criteria.

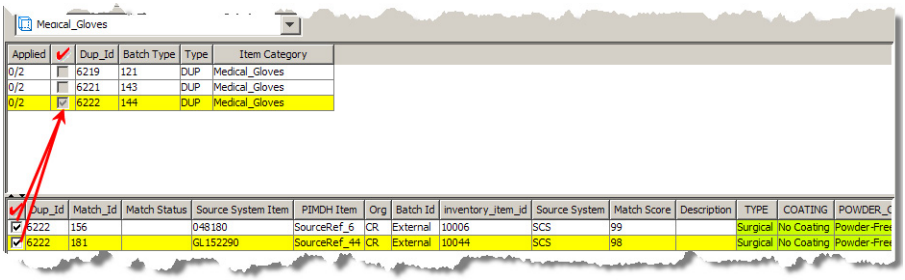
In the following figure, the selected **Dup_Id** is 6222 for which there are two items in the item master that can be quoted. To process the request for the quote process, you would select the best match based on the information provided by the system and possibly other criteria, such as the supplier or margin, and then select a single row to continue to be quoted. In this case, the Governance Studio determined that record 156 was the best quotable item for the **Dup_Id** 6222.



Once the record for quote has been selected, you can move the record to a downstream process using the options previously described.

Functional Equivalents Cross Reference

A second popular use case is the creation of a functional equivalents cross-reference to be used by web-based merchandising applications. In this case, a cross reference of all the functional equivalents found in the item master by the Oracle DataLens Server corresponding to **Dup_Id** 6222 is created. In the following figure, **Match_Id** 156 and 181 are selected as the two functional equivalent records for record 12.



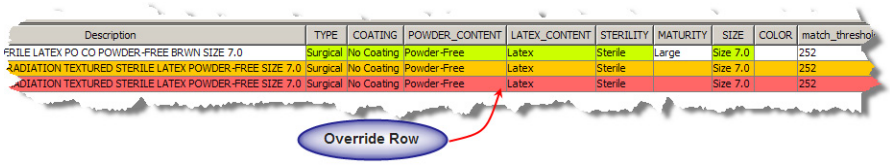
Once a record has been selected, you can process the checked rows with the options previously described.

The **Dup_Id**/**Match_Id** relationship can be used as a cross-reference, and by a down stream process, such as a PIM to create the cross-reference.

Survivorship Process

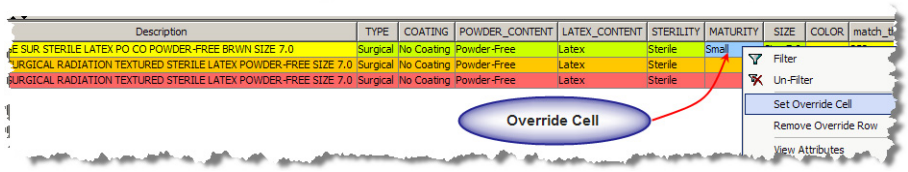
This advanced option assists in an item survivorship process by creating an override row that can then be manipulated to create a blended record of all possible duplicates. This is used during a system consolidation process.

1. Begin creating an override row by selecting a duplicates tab.
2. In the upper-pane, locate a duplicate to be merged and click the adjacent check box.
3. In the lower-pane, right-click on the duplicate row you want to merge and select **Create Override Row**.



The selected override row is highlighted in orange and is indicated in the Match Status column by **MO** (manual override). Once an override row is selected, you can set an override cell from the context-sensitive menu or remove the override row selection.

In the following example, the third row has been selected as the override row, and the contents of the **Maturity** column (attribute) in the first row has been selected as an override cell.



The resulting blended record is that all the information from the override row with the exception of the size attribute, which is extracted from the override cells. You can select an unlimited number of override cells to create the blended record. Once the rules for the composite record have been created, the rows are selected.

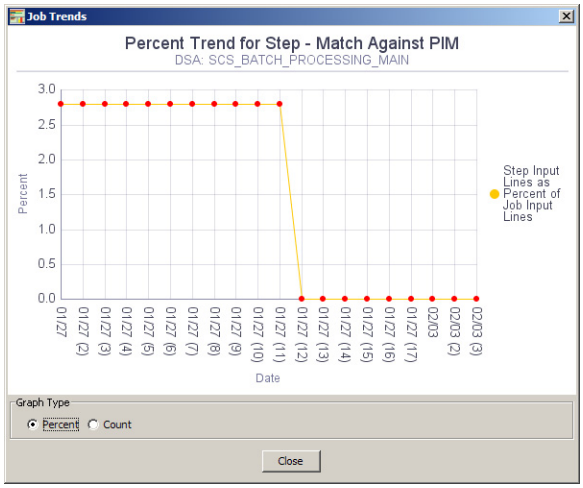
All of the additional options create a new row for the blended record tagged with an **MO** for Manual Override in the matched **Match Status** field. This informs the downstream survivorship process that this row is the override, which contains the blended record information.

> **Tip:** You can use a regular expression to filter your data to make it easier to work with a subset of the data and find the row that you want to use as the override row. For example, if your data contains an ellipsis (…) then you could filter your data using this as part of the regular expression with the text you are searching for, such as '…abc'. The filtered results would return only those rows that meet that condition. As a result, only checked rows of the filtered data are processed rather than the entire set so that you can easily create your override row. You first must remove the filter to remove the override row from the set of data. For more information about filtering, see "Filtering Data" on page 1-14.

Trend Analysis

Trend analysis is a Governance Studio feature that allows you to see a graph of the output data row count for various job runs you have made over time for a given

output step. You can use this feature on any tab by clicking the **Trend Analysis** button or from the **View** menu, select **Show Trend**.



You can toggle the view of the graph as a percentage or count of the rows in the selected Output tab using the **Graph Type** options provided.

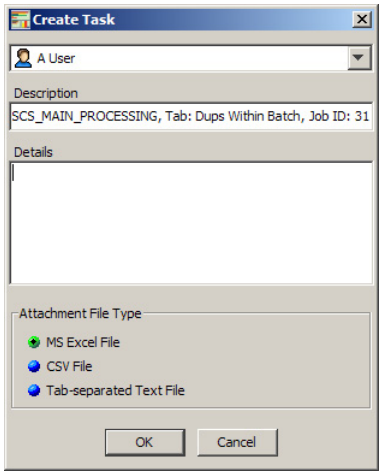
Though this functionality is available on all tabs, if trend data is not available, a message indicating this is displayed.

Create Task from Checked Rows

You can select one or more rows to create a task for another user to update the data lens. This capability is available on any Output tab.

1. From the **Run** menu, select **Create Task from Checked Rows**.

> **Note:** If the Output tab is a split-screen, you must choose whether you want to use total number of rows that include both the detail and total number of checked rows.



2. Select the user to whom this task will be assigned.

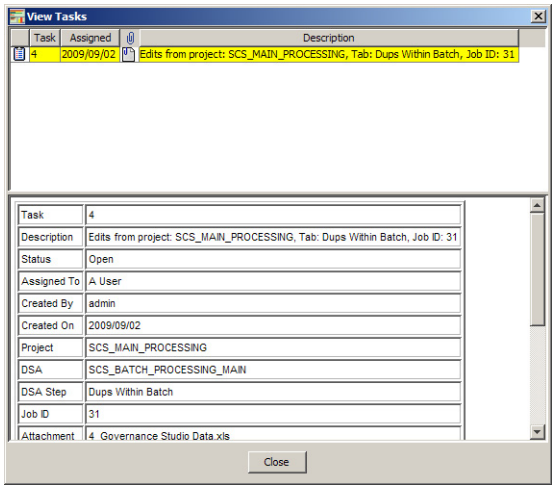
3. Edit the automatically populated **Description** field if you desired. It is populated with the DSA name, Output tab that the task is being created from, and a unique Job ID for the task (though this is not recommended.)
4. Add all instructions for user so that the task can be accomplished in the **Details** field.
5. Select the type of file to contain the selected to attach to the task using the options.
6. Click **OK**.

The task is created and an email containing the task details is sent to the assigned user. Attachments can be downloaded by the assigned user as described in the View My Tasks section.

View My Tasks

You can see if you have any tasks assigned to use with this feature.

From the **View** menu, select **View My Tasks**.



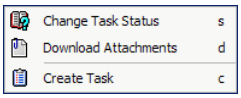
All assigned tasks are displayed in the top pane, while the bottom pane provides the details for the selected task.

> **Note:** Though the fields in the bottom pane appear to be editable, the changes are not saved.

The context-sensitive menu in the top pane is activated by right-clicking the attachment icon and is used as follows:

Change Task Status

See "Changing the Task Status" on page 5-13.



Download Attachments

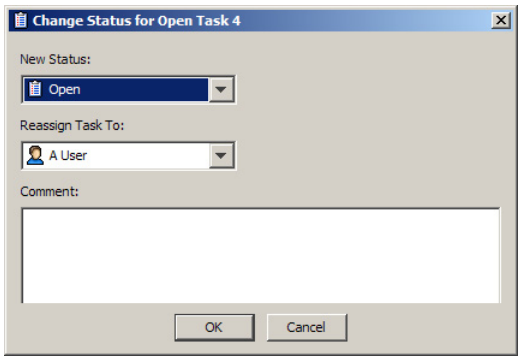
You can download the file that was saved when the task was created for use in completing the task. A file save dialog appears for you to select the directory in which to save the file.

Create Tasks

See "Creating a Task" on page 5-13.

Changing the Task Status

Selecting this option allows you to change the status of the task and/or reassign the task to another user.



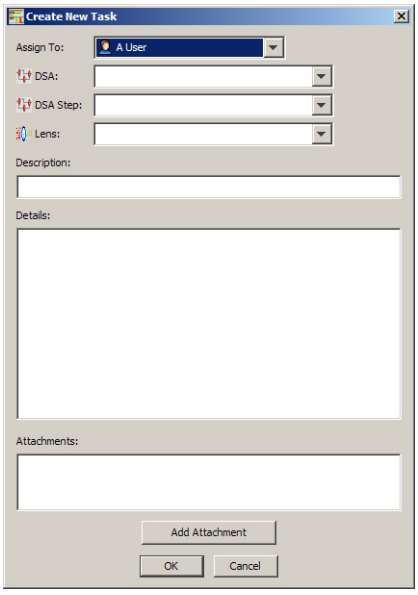
1. Select a new status and/or a user to reassign the task to from the lists.

   Tip: You can use the Unassigned Tasks user if you are unsure who you want to review this task and intend to assign it the proper person later.

2. Enter a comment that reflects why you have affected the change for future reference or to alert the new recipient of the task and why they are now responsible for it.
3. Click **OK**.

Creating a Task

Selecting this option allows you to create an entirely new review task.



1. Select a user to complete this task.
2. Select the DSA and the DSA step that you want to change.
3. Select the data lens to which the change is to be applied.

4. Enter a description and specific instructions on how to perform the task.
5. If you have a data file or other information that you want to attach, click **Add Attachment**, locate the file, and then click **OK**.

   Repeat this step until all necessary files are attached.
6. Click **OK**.

The task is created and an email containing the task details is sent to the assigned user.

A

Installing the Client Software

Oracle Product Data Quality uses a concept called Java Web Start to initially install and maintain the current version of the software on your client desktop. The process requires you to access the Oracle DataLens Server to initiate the connection and download the software.

The Oracle Product Data Quality client applications downloaded and installed using Java Web Start by browsing to the installation page for your Oracle DataLens Server as follows:

1. Ensure that you have the Java SE Runtime Environment (JRE) 6 Update 21 installed. You can download the JRE and obtain the installation instructions by browsing to:

   `http://www.oracle.com/technetwork/java/javase/downloads/index.html`

2. Start Microsoft Internet Explorer.

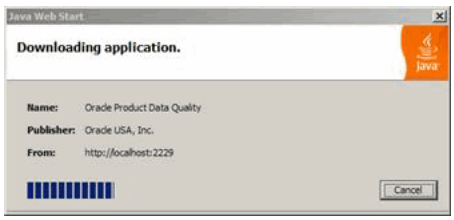
3. Initiate a connection and download the client software by browsing to:

   `http://`*`server`*`:2229/datalens/datalens.`

   Where *server* is the hostname of the Oracle DataLens Server.

   Note: If you have setup a different port number for your application server other than 2229, you must use that port number in the following URL when browsing to the Oracle DataLens Server to download the client applications.

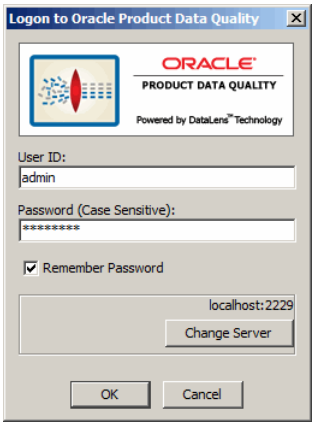
   The application download and verification begins.



   Note: If you receive a **File Download** message indicating that the . file is not associated with a program, you do not have the supported JRE installed. Click **Cancel** and return to Step

After the verification completes, the installation begins. Oracle Product Data Quality files are digitally signed by a trusted source so the following security warning is displayed:

Tip: To avoid this security dialog in the future, select the **Always trust content from this publisher** check box.

4. Click **Run** to continue and complete the installation.

   The Oracle Product Data Quality log on dialog is displayed.