



Informatica® Cloud Data Profiling
July 2024

Introduction

Informatica Cloud Data Profiling Introduction

July 2024

July 2024

© Copyright Informatica LLC 2019, 2024

This software and documentation are provided only under a separate license agreement containing restrictions on use and disclosure. No part of this document may be reproduced or transmitted in any form, by any means (electronic, photocopying, recording or otherwise) without prior consent of Informatica LLC.

Informatica and the Informatica logo are trademarks or registered trademarks of Informatica LLC in the United States and many jurisdictions throughout the world. A current list of Informatica trademarks is available on the web at <https://www.informatica.com/trademarks.html>. Other company and product names may be trade names or trademarks of their respective owners.

Subject to your opt-out rights, the software will automatically transmit to Informatica in the USA information about the computing and network environment in which the Software is deployed and the data usage and system statistics of the deployment. This transmission is deemed part of the Services under the Informatica privacy policy and Informatica will use and otherwise process this information in accordance with the Informatica privacy policy available at <https://www.informatica.com/in/privacy-policy.html>. You may disable usage collection in Administrator tool.

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 is 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.

Portions of this software and/or documentation are subject to copyright held by third parties. Required third party notices are included with the product.

See patents at <https://www.informatica.com/legal/patents.html>.

The information in this documentation is subject to change without notice. If you find any problems in this documentation, report them to us at infa_documentation@informatica.com.

Informatica products are warranted according to the terms and conditions of the agreements under which they are provided. INFORMATICA PROVIDES THE INFORMATION IN THIS DOCUMENT "AS IS" WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING WITHOUT ANY WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND ANY WARRANTY OR CONDITION OF NON-INFRINGEMENT.

Publication Date: 2024-07-19

Table of Contents

Preface	4
Informatica Resources.	4
Informatica Network.	4
Informatica Knowledge Base.	4
Informatica Documentation.	4
Informatica Product Availability Matrices.	5
Informatica Velocity.	5
Informatica Marketplace.	5
Informatica Global Customer Support.	5
Chapter 1: Introduction to Informatica Cloud Data Profiling	6
Data Profiling.	7
Administrator.	8
Monitor.	9
Operational Insights.	10
Informatica Intelligent Cloud Services security.	10
Chapter 2: Data Profiling tools	11
Projects.	11
Assets.	12
Explore Page.	12
Chapter 3: Informatica resources	15
Informatica Intelligent Cloud Services web site.	15
Informatica Intelligent Cloud Services Communities.	15
Informatica Intelligent Cloud Services Marketplace.	15
Informatica Intelligent Cloud Services Trust Center.	16
Informatica Product Availability Matrices.	16
Informatica Documentation.	16
Informatica Knowledge Base.	16
Informatica Global Customer Support.	17
Index	18

Preface

Read *Introduction* to learn about the tasks that you can perform using Informatica Intelligent Cloud Services Data Profiling.

Informatica Resources

Informatica provides you with a range of product resources through the Informatica Network and other online portals. Use the resources to get the most from your Informatica products and solutions and to learn from other Informatica users and subject matter experts.

Informatica Network

The Informatica Network is the gateway to many resources, including the Informatica Knowledge Base and Informatica Global Customer Support. To enter the Informatica Network, visit <https://network.informatica.com>.

As an Informatica Network member, you have the following options:

- Search the Knowledge Base for product resources.
- View product availability information.
- Create and review your support cases.
- Find your local Informatica User Group Network and collaborate with your peers.

Informatica Knowledge Base

Use the Informatica Knowledge Base to find product resources such as how-to articles, best practices, video tutorials, and answers to frequently asked questions.

To search the Knowledge Base, visit <https://search.informatica.com>. If you have questions, comments, or ideas about the Knowledge Base, contact the Informatica Knowledge Base team at KB_Feedback@informatica.com.

Informatica Documentation

Use the Informatica Documentation Portal to explore an extensive library of documentation for current and recent product releases. To explore the Documentation Portal, visit <https://docs.informatica.com>.

If you have questions, comments, or ideas about the product documentation, contact the Informatica Documentation team at infa_documentation@informatica.com.

Informatica Product Availability Matrices

Product Availability Matrices (PAMs) indicate the versions of the operating systems, databases, and types of data sources and targets that a product release supports. You can browse the Informatica PAMs at <https://network.informatica.com/community/informatica-network/product-availability-matrices>.

Informatica Velocity

Informatica Velocity is a collection of tips and best practices developed by Informatica Professional Services and based on real-world experiences from hundreds of data management projects. Informatica Velocity represents the collective knowledge of Informatica consultants who work with organizations around the world to plan, develop, deploy, and maintain successful data management solutions.

You can find Informatica Velocity resources at <http://velocity.informatica.com>. If you have questions, comments, or ideas about Informatica Velocity, contact Informatica Professional Services at ips@informatica.com.

Informatica Marketplace

The Informatica Marketplace is a forum where you can find solutions that extend and enhance your Informatica implementations. Leverage any of the hundreds of solutions from Informatica developers and partners on the Marketplace to improve your productivity and speed up time to implementation on your projects. You can find the Informatica Marketplace at <https://marketplace.informatica.com>.

Informatica Global Customer Support

You can contact a Global Support Center through the Informatica Network or by telephone.

To find online support resources on the Informatica Network, click **Contact Support** in the Informatica Intelligent Cloud Services Help menu to go to the **Cloud Support** page. The **Cloud Support** page includes system status information and community discussions. Log in to Informatica Network and click **Need Help** to find additional resources and to contact Informatica Global Customer Support through email.

The telephone numbers for Informatica Global Customer Support are available from the Informatica web site at <https://www.informatica.com/services-and-training/support-services/contact-us.html>.

CHAPTER 1

Introduction to Informatica Cloud Data Profiling

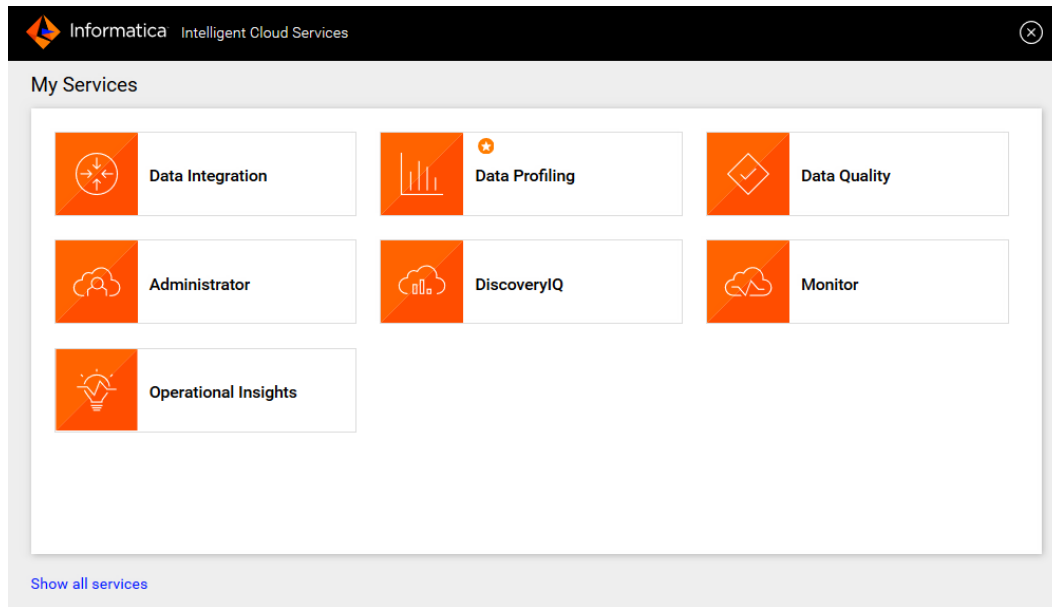
Informatica Cloud Data Profiling is one of several services available in Informatica Intelligent Cloud Services. Data Profiling is an on-demand subscription service that you can use to create data profiling tasks. You can create and run a data profiling task on a data source to determine the quality of data and understand the completeness, conformity, and consistency of data.

When you log in to Informatica Intelligent Cloud Services, the **My Services** page displays the services that apply to Data Profiling. The **My Services** page might also include other services that you subscribe to and trial versions of other services.

The following services apply to Data Profiling:

- **Administrator.** Perform administrative tasks such as creating users and managing privileges, downloading Secure Agents and configuring runtime environments, configuring connections, managing licenses, and creating schedules. For more information about the administrative tasks, see *Administrator* in the Administrator help.
- **Data Quality.** Create rule specifications and dictionaries. You can add one or more rule specifications to a data profiling task in Data Profiling. For information about creating rule specifications, see *Rule Specifications* in the Data Quality help.
- **Monitor.** Check the status of tasks that are in progress and tasks that have completed. For more information about monitoring jobs, see *Monitor* in the Monitor help.
- **Operational Insights.** View the performance and operational efficiency of data profiling tasks across your enterprise. View enterprise health, domain-level analytics, job and workflow statistics, and node-level resource consumption.

The following image shows the **My Services** page with some of the services that apply to Data Profiling:



Data Profiling

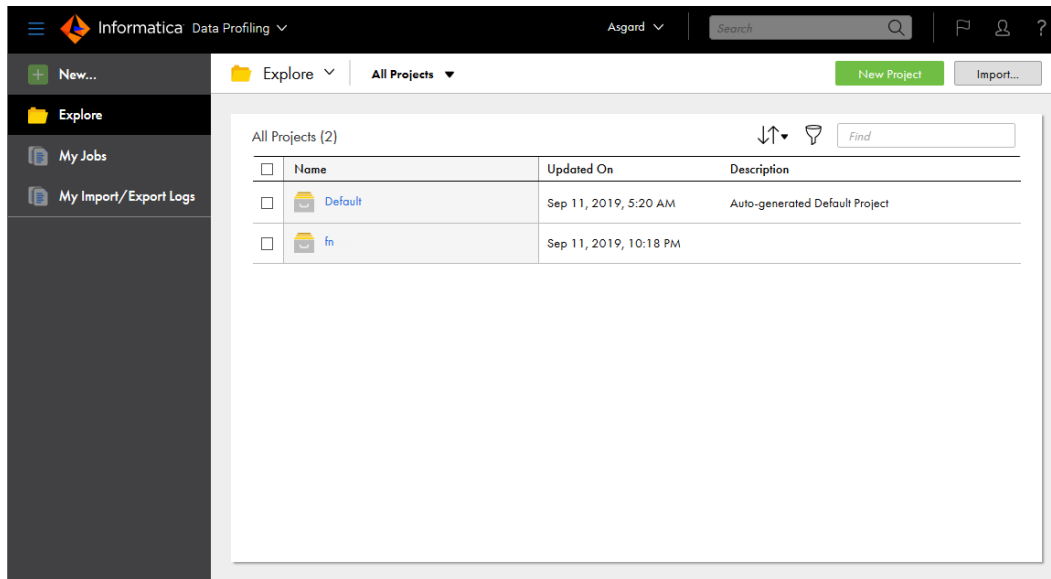
Use Data Profiling to create and run data profiling tasks. In Data Profiling, data profiling tasks are also called profiles.

When you run a profile on a source object, the results include the following column statistics:

- Number of distinct, non-distinct, and null values
- Percentage of distinct, non-distinct, and null values
- Documented and inferred data types
- Number of patterns
- Percentage of top pattern
- Maximum and minimum length of values
- Maximum and minimum values
- Average, sum, and standard deviation for numeric data types
- Value frequencies
- Outliers

You can run data profiling tasks on relational data sources and flat file data sources.

When you select Data Profiling from the **My Services** page, the **Explore** page appears as shown in the following image:



You can access the following pages from the navigation bar for Data Profiling:

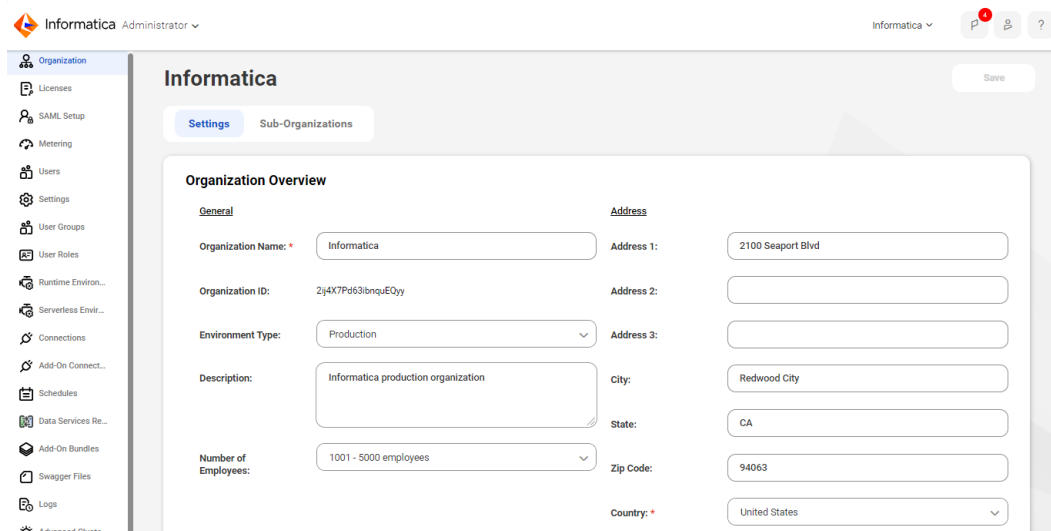
- New. Create data profiling tasks.
- Explore. Create and manage projects, folders, and assets.
- My Jobs. View the status of your data profiling and exception jobs.
- My Import/Export Logs. View the status of your imports and exports.

Note: You can open a profiling task and an exception task from the Explore page in Data Quality.

Administrator

Administrative tasks for Informatica Intelligent Cloud Services are performed in the Administrator service.

When you select Administrator from the **My Services** page, the navigation bar provides options for Informatica Intelligent Cloud Services administration, as shown in the following image:



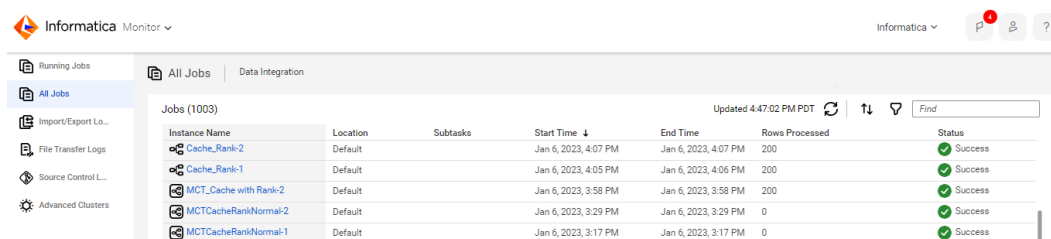
The following options are available based on your licenses:

- Organization. Manage your organization details and your sub-organizations.
- Licenses. View licenses for your organization and manage your sub-organization licenses.
- Ecosystem single sign-on. When you create your organization through Microsoft Azure, configure some single sign-on properties for Azure users.
- SAML Setup. Configure SAML single sign-on settings.
- Settings. Configure source control settings and upgrade settings for Secure Agent services.
- Users. Add new users. Configure user groups and roles for individual users.
- User Groups. Create user groups. Assign users to user groups.
- User Roles. Manage user roles for your organization. Configure the privileges associated with each role.
- Runtime Environments. Configure Secure Agents and Secure Agent groups.
- Connections. Create and manage connections.
- Add-On Connectors. Install add-on connectors.
- Schedules. Create and manage schedules for tasks.
- Add-On Bundles. Install, uninstall, upgrade, and copy bundles for your organization.
- Swagger Files. Generate swagger files for REST V2 connections.
- Logs. View asset and security logs.
- File servers. Configure file transfer servers and file transfer users to allow remote partners to send files to your organization.

Monitor

Use Monitor to monitor jobs, imports, and exports in your organization. A job is an instance of a mapping, task, or taskflow.

When you select Monitor from the **My Services** page, the navigation bar provides options for monitoring activity, as shown in the following image:



The screenshot shows the Informatica Monitor interface. The navigation bar includes 'Running Jobs', 'All Jobs', 'Data Integration', 'Import/Export Logs', 'File Transfer Logs', 'Source Control', and 'Advanced Clusters'. The main content area displays a table of jobs with the following columns: Instance Name, Location, Subtasks, Start Time, End Time, Rows Processed, and Status. The table shows five jobs, all with a 'Success' status.

Instance Name	Location	Subtasks	Start Time	End Time	Rows Processed	Status
Cache_Rank-2	Default		Jan 6, 2023, 4:07 PM	Jan 6, 2023, 4:07 PM	200	Success
Cache_Rank-1	Default		Jan 6, 2023, 4:05 PM	Jan 6, 2023, 4:06 PM	200	Success
MCT_Cache with Rank-2	Default		Jan 6, 2023, 3:58 PM	Jan 6, 2023, 3:58 PM	200	Success
MCTCacheRankNormal-2	Default		Jan 6, 2023, 3:29 PM	Jan 6, 2023, 3:29 PM	0	Success
MCTCacheRankNormal-1	Default		Jan 6, 2023, 3:17 PM	Jan 6, 2023, 3:17 PM	0	Success

The navigation bar provides the following options:

- Running Jobs. Provides run-time details about the Data Integration jobs that are running or have completed within the last five minutes.
- All Jobs. Provides details about all Data Integration jobs in the organization.
- Data Ingestion. Provides details about the Mass Ingestion jobs in the organization.
- Import/Export Logs. Provides details about imports and exports that are running and that have completed.
- File Transfer Logs. Provides details about the file transfers in the organization.

- Source Control Logs. Provides a log of actions on source-controlled objects in the organization. To view details about a specific job, import instance, or export instance, click the instance name.

Operational Insights

Use Operational Insights to view the performance and operational efficiency of data profiling assets across your enterprise. In Operational Insights, you can also view enterprise health, domain-level analytics, job and workflow statistics, and node-level resource consumption.

When you select Operational Insights from the **My Services** page, the **Home** page appears, as shown in the following image:

The screenshot shows the Informatica Operational Insights interface. The top navigation bar includes the Informatica logo, the text 'Operational Insights', and a user profile 'Asgard'. The main content area is titled 'Data Profiling' and displays a table of jobs. The table has columns for Instance Name, Location, Subtasks, Start Time, and End Time. The jobs listed include 'Outlier - x1 - run - 4-10', 'x1 - run - 4-9', 'Outlier - x1 - run - 3-8', 'x1 - run - 3-7', 'Outlier - x1 - run - 2-6', 'x1 - run - 2-5', 'qq2-4', and 'qq1-3'. The interface also shows a search bar, a refresh button, and pagination controls indicating '1 - 25 of 327' items, with '1 of 14' pages and 'Items Per Page: 25'.

Instance Name	Location	Subtasks	Start Time	End Time
Outlier - x1 - run - 4-10	Default	1	Feb 17, 2020, 10:09:53 ...	Feb 17, 2020, 1...
x1 - run - 4-9	Default	4	Feb 17, 2020, 10:08:35 P...	Feb 17, 2020, 1...
Outlier - x1 - run - 3-8	Default	1	Feb 17, 2020, 10:00:20 P...	Feb 17, 2020, 1...
x1 - run - 3-7	Default	4	Feb 17, 2020, 09:59:00 ...	Feb 17, 2020, 1...
Outlier - x1 - run - 2-6	Default	1	Feb 17, 2020, 09:58:06 ...	Feb 17, 2020, 0...
x1 - run - 2-5	Default	4	Feb 17, 2020, 09:56:47 ...	Feb 17, 2020, 0...
qq2-4	Default	1	Feb 17, 2020, 09:45:24 ...	Feb 17, 2020, 0...
qq1-3	Default	1	Feb 17, 2020, 09:45:07 ...	Feb 17, 2020, 0...

The navigation bar shows Data Profiling. To view details about a specific job, click the instance name.

Informatica Intelligent Cloud Services security

Informatica Intelligent Cloud Services uses authentication and encryption to ensure that data is secure and available only to users within the organization. Informatica Intelligent Cloud Services uses industry approved algorithms to encrypt sensitive information. All customer data is encrypted at rest using an AES-256 key. Data in transit is encrypted using the TLS 1.2 protocol.

For more information, see the [Informatica Intelligent Cloud Services Trust Center](#).

CHAPTER 2

Data Profiling tools

Informatica Intelligent Cloud Services provides the following tools that you can use for data profiling:

- Projects
- Assets
- Explore page

Projects

Informatica Intelligent Cloud Services assets are organized in projects. You can create multiple projects. A project can contain one or more folders or assets such as data profiling and rule specifications.

For example, you might have a business need to provide monthly analyses of your company's sales activity. You create a project called Monthly Sales Analysis. You create folders that contain the assets that you use on a monthly basis to integrate and run Data profiling tasks on the data from various systems to perform your analysis.

Assets

In Data Profiling, you can create data profiling tasks. You can create rule specifications in Data Quality and add them as rules to data profiling tasks. Rule specifications are Data Quality assets that define the data requirements of a business rule in logical form.

Explore Page

Use the **Explore** page to work with your data profiling projects, folders, and assets. The **Explore** page shows the assets created in the organization, which includes data profiling tasks and rule specifications.

Finding projects, folders, and assets on the Explore page

Use any of the following methods to find your projects, folders, and assets on the **Explore** page:

- Explore by projects and folders. View all projects or select a particular project.
- Explore by asset types. View all assets or view assets of a particular type.
- Explore by tags. View assets associated with a particular tag.
- Search for projects or assets. To search all projects, folders, and assets in the organization, view the **Explore** page by **All Projects**, and then enter a name or description in the Find box. Or, to narrow your search, view the **Explore** page by **All Assets**, select a specific asset type, project, or folder, and then enter a name or description in the Find box.
- Sort the search results. Sort the **Explore** page by name, last update date, description, or type. When you sort by type, the **Explore** page groups assets by asset type. It does not list the asset types in alphabetical order.
- Filter the objects on the page. To filter objects, click the **Filter** icon. To apply a filter, click **Add Field**, select the property to filter by, and then enter the property value. The filters available depend on how you view the page. You can specify multiple filters.
For example, to find all the assets in your organization with the tag SW Region, view the **Explore** page by **All Assets** and then click the **Filter** icon. Add the **Tags** filter and enter "SW Region."

You can see projects, folders, and assets for Data Profiling. If you select an asset to open it or perform an action, and the asset is created in a different service than the one you have open, the service opens in a new browser tab.

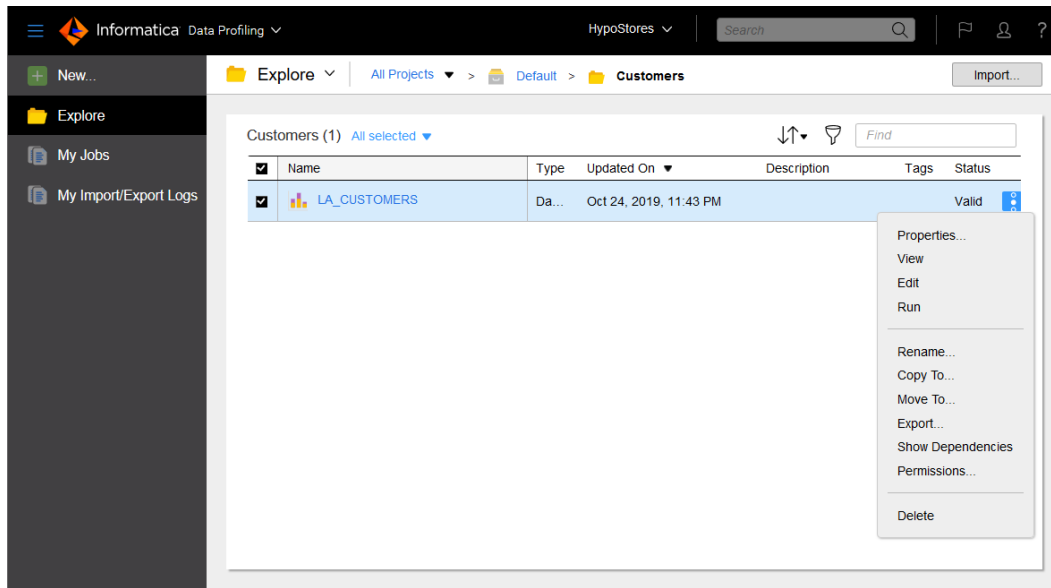
The **Explore** page does not support the following characters:

? ` | { } " ^ & [] / \

Do not use these characters in project, folder, asset, or tag names.

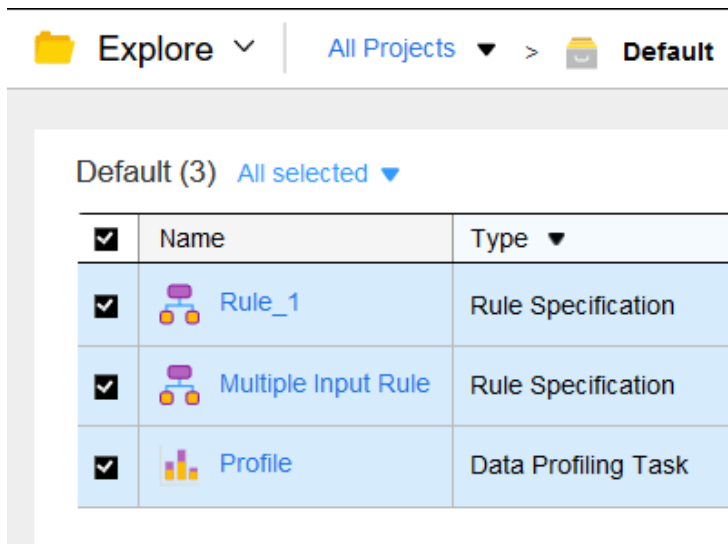
Working with projects and assets on the Explore page

Perform actions on projects, folders, and assets on the **Explore** page. To see what actions you can perform on an object, in the row that contains the object, click **Actions**, as shown in the following image:



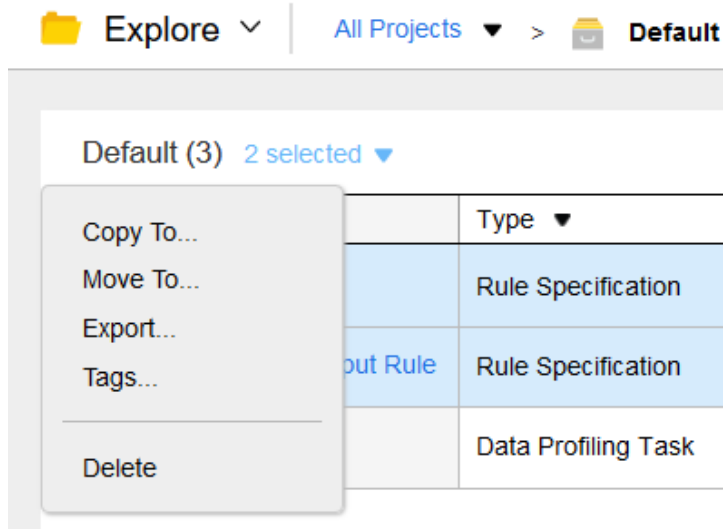
The Actions menu lists the actions you can perform based on your user role privileges and the permissions specified for the selected object. For example, your user role might have privileges to view and run tasks but not to delete tasks. You might not have permission to run a task because of the permissions set by the creator of the task.

You can also perform an action on multiple objects at one time. Select the check box to the left of each object, or select the **Select All** check box to select all of the objects that are displayed on the current page. The following image shows the **Select All** check box in use:



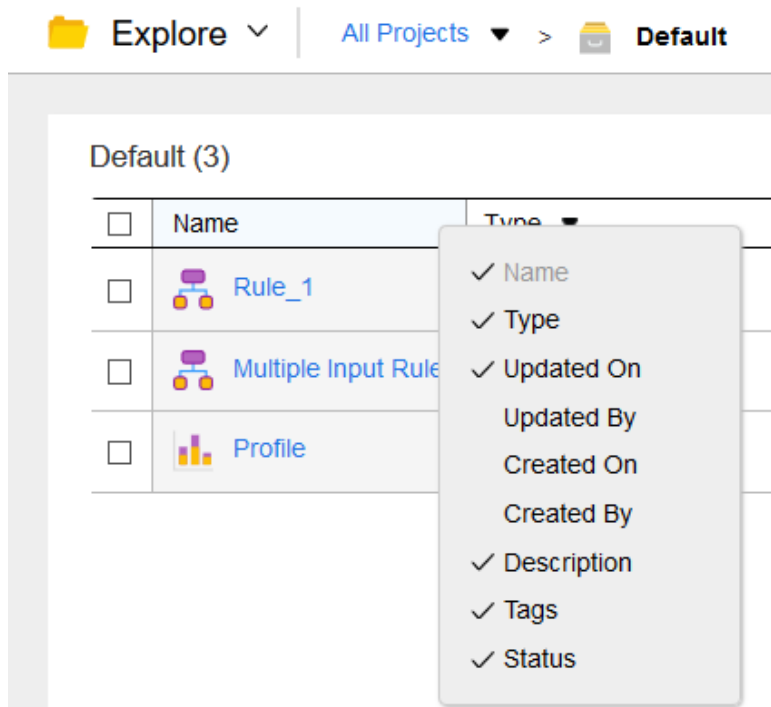
After you select the objects, click the **Actions** icon in the row of any of the selected objects. The Actions menu lists the actions available for the selected assets based on your user role and privileges.

Alternatively, use the **Selection** menu to choose the action, as shown in the following image:



Customizing the Explore page

You can display, hide, or rearrange columns on the **Explore** page. To display or hide columns, right-click the column heading area and check or uncheck the columns as shown in the following image:



To rearrange columns, click a column heading and drag it to a different location.

CHAPTER 3

Informatica resources

In addition to the online help, you can find information about Informatica Intelligent Cloud Services using the following resources.

Informatica Intelligent Cloud Services web site

You can access the Informatica Intelligent Cloud Services web site at <http://www.informatica.com/cloud>. This site contains information about Informatica Cloud integration services.

Informatica Intelligent Cloud Services Communities

Use the Informatica Intelligent Cloud Services Community to discuss and resolve technical issues. You can also find technical tips, documentation updates, and answers to frequently asked questions.

Access the Informatica Intelligent Cloud Services Community at:

<https://network.informatica.com/community/informatica-network/products/cloud-integration>

Developers can learn more and share tips at the Cloud Developer community:

<https://network.informatica.com/community/informatica-network/products/cloud-integration/cloud-developers>

Informatica Intelligent Cloud Services Marketplace

Visit the Informatica Marketplace to try and buy Data Integration Connectors, templates, and mapplets:

<https://marketplace.informatica.com/>

Informatica Intelligent Cloud Services Trust Center

The Informatica Intelligent Cloud Services Trust Center provides information about Informatica security policies and real-time system availability.

You can access the trust center at <https://www.informatica.com/trust-center.html>.

Subscribe to the Informatica Intelligent Cloud Services Trust Center to receive upgrade, maintenance, and incident notifications. The [Informatica Intelligent Cloud Services Status](#) page displays the production status of all the Informatica cloud products. All maintenance updates are posted to this page, and during an outage, it will have the most current information. To ensure you are notified of updates and outages, you can subscribe to receive updates for a single component or all Informatica Intelligent Cloud Services components. Subscribing to all components is the best way to be certain you never miss an update.

To subscribe, on the [Informatica Intelligent Cloud Services Status](#) page, click **SUBSCRIBE TO UPDATES**. You can choose to receive notifications sent as emails, SMS text messages, webhooks, RSS feeds, or any combination of the four.

Informatica Product Availability Matrices

Product Availability Matrices (PAMs) indicate the versions of the operating systems, databases, and types of data sources and targets that a product release supports. You can browse the Informatica PAMs at <https://network.informatica.com/community/informatica-network/product-availability-matrices>.

Informatica Documentation

Use the Informatica Documentation Portal to explore an extensive library of documentation for current and recent product releases. To explore the Documentation Portal, visit <https://docs.informatica.com>.

If you have questions, comments, or ideas about the product documentation, contact the Informatica Documentation team at infa_documentation@informatica.com.

Informatica Knowledge Base

Use the Informatica Knowledge Base to find product resources such as how-to articles, best practices, video tutorials, and answers to frequently asked questions.

To search the Knowledge Base, visit <https://search.informatica.com>. If you have questions, comments, or ideas about the Knowledge Base, contact the Informatica Knowledge Base team at KB_Feedback@informatica.com.

Informatica Global Customer Support

You can contact a Global Support Center through the Informatica Network or by telephone.

To find online support resources on the Informatica Network, click **Contact Support** in the Informatica Intelligent Cloud Services Help menu to go to the **Cloud Support** page. The **Cloud Support** page includes system status information and community discussions. Log in to Informatica Network and click **Need Help** to find additional resources and to contact Informatica Global Customer Support through email.

The telephone numbers for Informatica Global Customer Support are available from the Informatica web site at <https://www.informatica.com/services-and-training/support-services/contact-us.html>.

INDEX

A

assets [12](#)

C

Cloud Application Integration community
URL [15](#)
Cloud Developer community
URL [15](#)

D

Data Integration community
URL [15](#)

E

Explore page [12](#)

I

Informatica Global Customer Support
contact information [5](#), [17](#)
Informatica Intelligent Cloud Services
web site [15](#)

M

maintenance outages [16](#)

P

project folders [11](#), [12](#)
projects [11](#), [12](#)

S

security [10](#)
status
Informatica Intelligent Cloud Services [16](#)
system status [16](#)

T

TLS
security [10](#)
trust site
description [16](#)

U

upgrade notifications [16](#)

W

web site [15](#)