This software and documentation contain proprietary information of Informatica Corporation and are provided under a license agreement containing restrictions on use and disclosure and are also protected by copyright law. Reverse engineering of the software is prohibited. 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 Corporation. This Software may be protected by U.S. and/or international Patents and other Patents Pending.

Use, duplication, or disclosure of the Software by the U.S. Government is subject to the restrictions set forth in the applicable software license agreement and as provided in DFARS 227.7202-1(a) and 227.7202-3(a) (1995), DFARS 252.227-7013 (OCT 1988), FAR 12.212(a) (1995), FAR 52.227-19, or FAR 52.227-14 (ALT III), as applicable.

The information in this product or documentation is subject to change without notice. If you find any problems in this product or documentation, please report them to us in writing.

Informatica, Informatica Platform, Informatica Data Services, PowerCenter, PowerCenterRT, PowerCenter Connect, PowerCenter Data Analyzer, PowerExchange, PowerMart, Metadata Manager, Informatica Data Quality, Informatica Data Explorer, Informatica B2B Data Transformation, Informatica B2B Data Exchange and Informatica On Demand are trademarks or registered trademarks of Informatica Corporation in the United States and in jurisdictions throughout the world. All other company and product names may be trade names or trademarks of their respective owners.

Portions of this software and/or documentation are subject to copyright held by third parties, including without limitation: Copyright DataDirect Technologies. All rights reserved. Copyright © Sun Microsystems. All rights reserved. Copyright © RSA Security Inc. All Rights Reserved. Copyright © Ordinal Technology Corp. All rights reserved. Copyright © Aendacht c.v. All rights reserved. Copyright © Genivia, Inc. All rights reserved. Copyright © 2007 Isomorphic Software. All rights reserved. Copyright © Meta Integration Technology, Inc. All rights reserved. Copyright © Intalio. All rights reserved. Copyright © Oracle. All rights reserved. Copyright © Adobe Systems Incorporated. All rights reserved. Copyright © DataArt, Inc. All rights reserved. Copyright © ComponentSource. All rights reserved. Copyright © Microsoft Corporation. All rights reserved. Copyright © Rouge Wave Software, Inc. All rights reserved. Copyright © Teradata Corporation. All rights reserved. Copyright © Yahoo! Inc. All rights reserved. Copyright © Glyph & Cog, LLC. All rights reserved.

This product includes software developed by the Apache Software Foundation (http://www.apache.org/), and other software which is licensed under the Apache License, Version 2.0 (the "License"). You may obtain a copy of the License at http://www.apache.org/licenses/LICENSE-2.0. Unless required by applicable law or agreed to in writing, software distributed under the License is distributed on an "AS IS" BASIS, WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied. See the License for the specific language governing permissions and limitations under the License.

This product includes software which was developed by Mozilla (http://www.mozilla.org/), software copyright The JBoss Group, LLC, all rights reserved; software copyright © 1999-2006 by Bruno Lowagie and Paulo Soares and other software which is licensed under the GNU Lesser General Public License Agreement, which may be found at http://www.gnu.org/licenses/lgpl.html. The materials are provided free of charge by Informatica, "as-is", without warranty of any kind, either express or implied, including but not limited to the implied warranties of merchantability and fitness for a particular purpose.

The product includes ACE(TM) and TAO(TM) software copyrighted by Douglas C. Schmidt and his research group at Washington University, University of California, Irvine, and Vanderbilt University, Copyright (©) 1993-2006, all rights reserved.

This product includes software developed by the OpenSSL Project for use in the OpenSSL Toolkit (copyright The OpenSSL Project. All Rights Reserved) and redistribution of this software is subject to terms available at http://www.openssl.org.

This product includes Curl software which is Copyright 1996-2007, Daniel Stenberg, <daniel@haxx.se>. All Rights Reserved. Permissions and limitations regarding this software are subject to terms available at http://curl.haxx.se/docs/copyright.html. Permission to use, copy, modify, and distribute this software for any purpose with or without fee is hereby granted, provided that the above copyright notice and this permission notice appear in all copies.

The product includes software copyright 2001-2005 (©) MetaStuff, Ltd. All Rights Reserved. Permissions and limitations regarding this software are subject to terms available at http://www.dom4j.org/ license.html.

The product includes software copyright © 2004-2007, The Dojo Foundation. All Rights Reserved. Permissions and limitations regarding this software are subject to terms available at http://svn.dojo.org/dojo/trunk/LICENSE.

This product includes ICU software which is copyright International Business Machines Corporation and others. All rights reserved. Permissions and limitations regarding this software are subject to terms available at http://source.icu-project.org/repos/icu/icu/trunk/license.html.

This product includes software copyright © 1996-2006 Per Bothner. All rights reserved. Your right to use such materials is set forth in the license which may be found at http://www.gnu.org/software/ kawa/Software-License.html.

This product includes OSSP UUID software which is Copyright © 2002 Ralf S. Engelschall, Copyright © 2002 The OSSP Project Copyright © 2002 Cable & Wireless Deutschland. Permissions and limitations regarding this software are subject to terms available at http://www.opensource.org/licenses/mit-license.php.
This product includes software developed by Boost (http://www.boost.org/) or under the Boost software license. Permissions and limitations regarding this software are subject to terms available at http://www.boost.org/LICENSE_1_0.txt.

This product includes software copyright © 1997-2007 University of Cambridge. Permissions and limitations regarding this software are subject to terms available at http://www.pcre.org/license.txt.

This product includes software copyright © 2007 The Eclipse Foundation. All Rights Reserved. Permissions and limitations regarding this software are subject to terms available at http://www.eclipse.org/org/documents/epl-v10.php.

This product includes software licensed under the Academic Free License (http://www.opensource.org/licenses/afl-3.0.php), the Common Development and Distribution License (http://www.opensource.org/licenses/cddl1.php) the Common Public License (http://www.opensource.org/licenses/cpl1.0.php) and the BSD License (http://www.opensource.org/licenses/bsd-license.php).

This product includes software copyright © 2003-2006 Joe WaInes, 2006-2007 XStream Committers. All rights reserved. Permissions and limitations regarding this software are subject to terms available at http://xstream.codehaus.org/license.html. This product includes software developed by the Indiana University Extreme! Lab. For further information please visit http://www.extreme.indiana.edu/.

This Software is protected by U.S. Patent Numbers 5,794,246; 6,014,670; 6,016,501; 6,029,178; 6,032,158; 6,035,307; 6,044,374; 6,092,086; 6,208,990; 6,339,775; 6,640,226; 6,789,096; 6,820,077; 6,823,373; 6,850,947; 6,895,471; 7,117,215; 7,162,643; 7,254,590; 7,281,001; 7,421,458; and 7,584,422, international Patents and other Patents Pending.

DISCLAIMER: Informatica Corporation provides this documentation "as is" without warranty of any kind, either express or implied, including, but not limited to, the implied warranties of non-infringement, merchantability, or use for a particular purpose. Informatica Corporation does not warrant that this software or documentation is error free. The information provided in this software or documentation may include technical inaccuracies or typographical errors. The information in this software and documentation is subject to change at any time without notice.

NOTICES

This Informatica product (the “Software”) includes certain drivers (the “DataDirect Drivers”) from DataDirect Technologies, an operating company of Progress Software Corporation (“DataDirect”) which are subject to the following terms and conditions:

1. THE DATADIRECT DRIVERS ARE PROVIDED “AS IS” WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESSED OR IMPLIED, INCLUDING BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NON-INFRINGEMENT.

2. IN NO EVENT WILL DATADIRECT OR ITS THIRD PARTY SUPPLIERS BE LIABLE TO THE END-USER CUSTOMER FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, CONSEQUENTIAL OR OTHER DAMAGES ARISING OUT OF THE USE OF THE ODBC DRIVERS, WHETHER OR NOT INFORMED OF THE POSSIBILITIES OF DAMAGES IN ADVANCE. THESE LIMITATIONS APPLY TO ALL CAUSES OF ACTION, INCLUDING, WITHOUT LIMITATION, BREACH OF CONTRACT, BREACH OF WARRANTY, NEGLIGENCE, STRICT LIABILITY, MISREPRESENTATION AND OTHER TORTS.

Part Number: IN-AUG-90000-0001
# Table of Contents

**Preface** ................................................................. iv
Informatica Resources. ................................................................. iv
Informatica Customer Portal. ................................................................. iv
Informatica Documentation. ................................................................. iv
Informatica Web Site. ................................................................. iv
Informatica How-To Library. ................................................................. iv
Informatica Knowledge Base. ................................................................. v
Informatica Multimedia Knowledge Base. ................................................................. v
Informatica Global Customer Support. ................................................................. v

**Chapter 1: Introduction to Informatica Analyst** ....................................... 1
Informatica Analyst Overview. ................................................................. 1
Informatica Analyst Interface. ................................................................. 2
Informatica Analyst Navigator. ................................................................. 2
Informatica Analyst Views. ................................................................. 3
Contents View. ................................................................. 3
Properties View. ................................................................. 3
Security View. ................................................................. 3
Logging In to Informatica Analyst. ................................................................. 4

**Chapter 2: Projects** ............................................................... 5
Projects Overview. ................................................................. 5
Creating a Project. ................................................................. 6
Duplicating a Project. ................................................................. 6
Renaming a Project. ................................................................. 7
Deleting a Project. ................................................................. 7
Folders Overview. ................................................................. 7
Creating a Folder. ................................................................. 7
Renaming a Folder. ................................................................. 8
Duplicating a Folder. ................................................................. 8
Moving a Folder. ................................................................. 8
Deleting a Folder. ................................................................. 9
Viewing a Project or Folder. ................................................................. 9
Objects Overview. ................................................................. 9
Viewing an Object. ................................................................. 10
Duplicating an Object. ................................................................. 10
Renaming an Object. ................................................................. 10
Moving an Object. ................................................................. 11
Preface

The Informatica Analyst User Guide is written for data quality analysts. It describes how to use Informatica Analyst. This guide assumes that you have an understanding of data quality concepts, flat file and relational database concepts, and the database engines in your environment.

Informatica Resources

Informatica Customer Portal

As an Informatica customer, you can access the Informatica Customer Portal site at http://my.informatica.com. The site contains product information, user group information, newsletters, access to the Informatica customer support case management system (ATLAS), the Informatica How-To Library, the Informatica Knowledge Base, the Informatica Multimedia Knowledge Base, Informatica Documentation Center, and access to the Informatica user community.

Informatica Documentation

The Informatica Documentation team takes every effort to create accurate, usable documentation. If you have questions, comments, or ideas about this documentation, contact the Informatica Documentation team through email at infa_documentation@informatica.com. We will use your feedback to improve our documentation. Let us know if we can contact you regarding your comments.

The Documentation team updates documentation as needed. To get the latest documentation for your product, navigate to the Informatica Documentation Center from http://my.informatica.com.

Informatica Web Site

You can access the Informatica corporate web site at http://www.informatica.com. The site contains information about Informatica, its background, upcoming events, and sales offices. You will also find product and partner information. The services area of the site includes important information about technical support, training and education, and implementation services.

Informatica How-To Library

As an Informatica customer, you can access the Informatica How-To Library at http://my.informatica.com. The How-To Library is a collection of resources to help you learn more about Informatica products and features. It includes articles and interactive demonstrations that provide solutions to common problems, compare features and behaviors, and guide you through performing specific real-world tasks.
Informatica Knowledge Base

As an Informatica customer, you can access the Informatica Knowledge Base at http://my.informatica.com. Use the Knowledge Base to search for documented solutions to known technical issues about Informatica products. You can also find answers to frequently asked questions, technical white papers, and technical tips. If you have questions, comments, or ideas about the Knowledge Base, contact the Informatica Knowledge Base team through email at KB_Feedback@informatica.com.

Informatica Multimedia Knowledge Base

As an Informatica customer, you can access the Informatica Multimedia Knowledge Base at http://my.informatica.com. The Multimedia Knowledge Base is a collection of instructional multimedia files that help you learn about common concepts and guide you through performing specific tasks. If you have questions, comments, or ideas about the Multimedia Knowledge Base, contact the Informatica Knowledge Base team through email at KB_Feedback@informatica.com.

Informatica Global Customer Support

You can contact a Customer Support Center by telephone or through the WebSupport Service. WebSupport requires a user name and password. You can request a user name and password at http://my.informatica.com.

Use the following telephone numbers to contact Informatica Global Customer Support:

<table>
<thead>
<tr>
<th>North America / South America</th>
<th>Europe / Middle East / Africa</th>
<th>Asia / Australia</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Toll Free</strong></td>
<td><strong>Toll Free</strong></td>
<td><strong>Toll Free</strong></td>
</tr>
<tr>
<td>+1 877 463 2435</td>
<td>00 800 4632 4357</td>
<td>Australia: 1 800 151 830</td>
</tr>
<tr>
<td><strong>Standard Rate</strong></td>
<td><strong>Standard Rate</strong></td>
<td>Singapore: 001 800 4632 4357</td>
</tr>
<tr>
<td>Brazil: +55 11 3523 7761</td>
<td>Belgium: +32 15 281 702</td>
<td><strong>Standard Rate</strong></td>
</tr>
<tr>
<td>Mexico: +52 55 1168 9763</td>
<td>France: +33 1 41 38 92 26</td>
<td>India: +91 80 4112 5738</td>
</tr>
<tr>
<td>United States: +1 650 385 5800</td>
<td>Germany: +49 1805 702 702</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Netherlands: +31 306 022 797</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Spain and Portugal: +34 93 480 3760</td>
<td></td>
</tr>
<tr>
<td></td>
<td>United Kingdom: +44 1628 511 445</td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER 1

Introduction to Informatica Analyst

This chapter includes the following topics:

- Informatica Analyst Overview, 1
- Informatica Analyst Interface, 2
- Logging In to Informatica Analyst, 4

Informatica Analyst Overview

Informatica Analyst is a web-based application client that analysts can use to analyze, cleanse, and standardize data in an enterprise. Use the Analyst tool to collaborate with data quality and data integration developers on data quality integration solutions. You can perform column and rule profiling, manage scorecards, and manage bad records and duplicate records in the Analyst tool. You can also manage and provide reference data to developers in a data quality solution.

The Analyst Service manages the Analyst tool. The Analyst tool uses projects to store folders and objects. The Analyst tool stores projects, folders, and objects in the Model repository. The Analyst tool connects to the Model repository database to create, update, and delete projects and objects in the Analyst tool.

You can import data objects such as tables and flat files into projects and folders. The Analyst Service manages the connection to the directory that stores uploaded flat files that you use as flat file sources in the Analyst tool. The Analyst Service also manages the connection to a database that stores reference tables that you create or import in the Analyst tool.

You can use the data objects to profile data, apply rules to cleanse the data, and create scorecards to measure data quality progress. You can create reference tables from profile results to standardize the data. The Data Quality Assistant is a web-based tool in the Analyst tool that you can use to fix bad records and consolidate duplicate records in tables.

The tasks that you can perform in the Analyst tool depend on the Informatica products that you use. Use the Analyst tool to accomplish the following tasks:

- Run profiles. Create and run profiles to analyze the structure and content of enterprise data and identify strengths and weaknesses. After you run a profile, you can view the rows from the profile results. You can also add profiled columns to scorecards and add column values to reference tables.

- Create rules in profiles. Create and apply rules within profiles. A rule is reusable business logic that defines conditions applied to data when you run a profile. Use rules to further validate the data in a profile and to measure data quality progress.

- Score data. Create scorecards to score the valid values for any column or the output of rules. Scorecards display the value frequency for columns in a profile as scores. Use scorecards to measure and visually represent data quality progress. You can also view trend charts to view the history of scores over time.

Manage bad records and duplicate records. Filter records, edit records, and set record status. Merge groups of duplicate records from a consolidated record table into a single record.

Informatica Analyst Interface

Use the Analyst tool interface to perform Analyst tool tasks. The Analyst tool interface has tabs, headers, views, and a Navigator. Use the Navigator to browse projects and perform tasks on projects and folders.

When you log in to the Analyst tool, the Browse: Projects tab appears. The tab displays views and the Navigator. The tab also displays the icons and the Actions menus that you can use to perform tasks in the Navigator and in views.

When you click the link for an object in a view, the object opens in a tab. After you perform a search, the Analyst tool displays the search results in a tab. You can keep multiple tabs open in the Analyst tool interface. You can navigate between tabs.

The Analyst tool has the following header items:

- **Sign Off.** Log out of the Analyst tool.
- **User Preferences.** Set user preferences to open metadata bookmarks in the Analyst tool or Developer tool.
- **Help.** Access help for the current tab.
- **Search.** Search for objects and folders in projects.

Informatica Analyst Navigator

Use the Navigator to browse projects and folders and their contents. After you log in to the Analyst tool, the Navigator appears in the left pane.

When you select a project in the Navigator, you can select a view to view the project contents, descriptive information about the project, and permissions on the project.

The Navigator displays the following types of objects:

- **Projects.** Highest object in the Navigator hierarchy. It is the top-level container for all projects that you create.
- **Folders.** Child object of a project. Organize domain objects within a project in folders.

Refresh the Navigator to get the latest version of all objects in the Navigator. Multiple users can add projects and folders that appear in the Navigator.

The Navigator has an Actions menu that you can use to perform tasks on projects and folders. You can also right-click projects and folders to perform the same tasks.

Use the Navigator to perform the following tasks:

- Create projects and folders.
- Manage projects and folders.
- Refresh the projects and folders that appear in the Navigator.
Informatica Analyst Views

The Analyst tool has views for the projects and folders that you select in the Navigator. Objects that open in tabs also have views. Use the Actions menu or right-click objects to perform tasks related to the view. You can also click icons in the view panels to perform the common tasks related to the view.

The Contents View and the Properties View are the views for the top-level Projects container. The Projects container contains the projects that you create in the Navigator.

After you select a project or folder, the Analyst tool interface displays the following views:

- **Contents view.** Displays project or folder contents and properties for selected objects.
- **Properties view.** Displays project or folder properties.
- **Security view.** Displays user permissions on the project.

Contents View

Use the Contents view to view project and folder information. In the Contents view, you can create and add data objects and profiles to projects and folders. You can perform project and folder management tasks.

After you select a project or folder in the Navigator, click the Contents view to view project or folder contents.

The Contents view displays project or folder contents in the Contents panel. When you select an object in the Contents panel, the Analyst tool displays the object properties in the Properties panel.

You can perform the following tasks in the Contents view:

- Open an object.
- Duplicate projects, folders, and objects.
- Rename projects, folders, and objects.
- Move folders and objects.
- Delete projects, folders, and objects.
- Add a flat file to a project or folder.
- Add a relational table to a project or folder.
- Create a custom profile.
- Create a reference table.
- Create bad record or duplicate record tables.
- Close all tabs.

Properties View

Use the Properties view to view descriptive information about the project or folder.

After you select a project or folder in the Navigator, click the Properties view to view the project or folder properties.

In the Properties view, you can view the project or folder name and description.

Security View

Use the Security view to view and assign project-level permissions to users.
After you select a project in the Navigator, click the Security view to view user permissions on the project. In the Security view, you can assign the read, write, and grant permissions to users. You can also add users and assign permissions to them.

The Security view displays the following information in the Project-level permissions panel:

♦ User. User name for the user who is assigned permissions on the project.
♦ Security domain. Name of the security domain that the user belongs to. Security domain can be LDAP or Native.
♦ Permission. Permissions assigned to the user. Permissions can include read, write, or grant permission.

Logging In to Informatica Analyst

Use the Analyst tool URL to log in to the Analyst tool interface. When you log in to the Analyst tool, you must specify the user name, password, and the native domain or the LDAP security domain.

1. Start a Microsoft Internet Explorer or Mozilla Firefox browser.
2. In the Address field, enter the URL for the Analyst tool:
   http[s]://<host name>:<port number>/AnalystTool
3. On the login page, enter your user name and password.
4. Select Native or the name of a specific security domain.
   The Security Domain field appears when the Informatica domain contains an LDAP security domain. If you do not know the security domain that your user account belongs to, contact the Informatica domain administrator.
5. Click Login.
   The welcome screen appears.
6. Click Close to exit the welcome screen and access the Analyst tool.
Projects Overview

A project is the top-level container that you use to store folders and objects in the Analyst tool. Use projects to organize and manage the objects that you want to analyze for data quality.

Create a project based on the structure of the data for which you want to analyze data quality. For example, an analyst needs to assess data quality on multiple systems structured by region in a country. The analyst creates projects named East and West to correspond with data for East and West regions. The analyst can import data objects such as relational tables and flat files in the East and West projects and create object types such as profiles, rules, and scorecards to analyze data quality in each project.

You must create or open a project before you can work in the Analyst tool. Use the Navigator to create a project in the Analyst tool. When you create a project, the Analyst tool stores the project in the Model repository.

You can share a project to share the project contents and collaborate with other users on the project. When you share a project in the Analyst tool, the project also appears in the Developer tool.

A project can contain folders and objects. You can organize objects in folders.
The following table describes the tasks you can perform on a project:

<table>
<thead>
<tr>
<th>Task</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manage projects</td>
<td>Manage and share project contents. You can create, duplicate, rename, and delete a project. You can view project contents.</td>
</tr>
<tr>
<td>Manage folders</td>
<td>Organize project content in folders. You can create, duplicate, rename, move, and rename folders within projects.</td>
</tr>
<tr>
<td>Manage objects</td>
<td>You can view object contents, duplicate, rename, move, and delete objects in a project or in a folder within a project.</td>
</tr>
<tr>
<td>Search projects</td>
<td>You can search for folders or objects in projects. You can view search results and select an object from the results to view its contents.</td>
</tr>
<tr>
<td>Assign permissions</td>
<td>You can add users to a project. You can assign the read, write, and grant permissions to users on a project to restrict or provide access to objects within the project.</td>
</tr>
</tbody>
</table>

**Creating a Project**

Create a project to store data objects and object types in the Analyst tool. You can create folders in projects. Use projects to manage the folders and objects in the project.

1. In the Navigator, select **Projects**.
2. Click **Actions > New Project**.
   
   The **New Project** window appears.
3. Enter a name for the project and an optional description.
4. Click **Unshared** if you do not want to share the project or **Shared** if you want to share the project with other users. Default is **Unshared**.
5. Click **OK**.
   
   The project appears in the Navigator.

**Duplicating a Project**

You can duplicate project contents in another project that you create. Duplicate a project to use the same project contents to perform different project tasks. Duplicating a project does not duplicate the user permissions on the project. The owner of the project gets all permissions by default on the duplicate project.

1. In the Navigator, select the project that you want to duplicate.
2. Click **Actions > Duplicate**.
   
   The **Duplicate** window appears.
3. Enter the project name and an optional description.
4. Click **OK**.
   
   The Analyst tool duplicates the project contents in the project.
Renaming a Project

You can rename a project after you create it. Rename a project according to business usage and naming convention. You may need to rename a project because the name is incorrect or the project has a different use.

1. In the Navigator, select the project that you want to rename.
2. Click **Actions > Rename**.
   
The Rename window appears.
3. Enter a name.
4. Click **OK**.
   
The Analyst tool renames the project in the Navigator.

Deleting a Project

Delete a project when the project and its contents become redundant.

1. In the Navigator, select the project that you want to delete.
2. Click **Actions > Delete**.
3. In the **Delete Project** window, click **Yes**.
   
The Analyst tool deletes the project from the Navigator.

Folders Overview

Use folders to organize project contents. You can create a folder to group objects for a particular task in a project. You can create a folder in a project or in another folder.

Create folders to group objects based on business needs. For example, a project requires data analysis for data stored in multiple relational databases across an organization. Each region has a relational database. You can create folders named East and West to store the project metadata for each region.

Folders appear under projects in the Navigator. A folder can contain other folders and objects.

You can perform the following tasks on a folder:

- Create a folder.
- View a folder.
- Rename a folder.
- Duplicate a folder.
- Move a folder.
- Delete a folder.

Creating a Folder

Create a folder to store objects created in the Analyst tool.

1. In the Navigator, select the project or folder where you want to create a folder.
2. Click **Actions > New Folder**.
3. Enter the folder name and optional description.
4. Click OK.
   The Analyst tool creates the folder in the Navigator.

Renaming a Folder

You can rename a folder after you create it. Rename a folder to change its name according to business usage or naming convention.

1. In the Navigator, select the project and the folder in the project that you want to rename.
2. Click Actions > Rename
   The Rename window appears.
3. Enter the folder name.
4. Click OK.
   The Analyst tool renames the folder in the Navigator.

Duplicating a Folder

You can duplicate a folder within a project. Duplicate a folder to organize or enhance the contents of a folder or to use the contents of a folder to perform different tasks.

1. In the Navigator, select the project and the folder in the project that you want to duplicate.
2. Click Actions > Duplicate.
   The Duplicate window appears.
3. Navigate to the location where you want to duplicate the folder.
   Optionally, enter the location.
4. Enter the folder name.
5. Click OK.
   The Analyst tool duplicates the folder in the project in the Navigator.

Moving a Folder

You can move a folder within a project. Move folders to organize project content into a hierarchy of folders.

1. In the Navigator, select the folder in the project you want to move.
2. Click Actions > Move.
   The Move window appears.
3. Navigate to the location in the project where you want to move the folder.
   Optionally, enter the location.
4. Click OK.
   The Analyst tool moves the folder in the project in the Navigator.
Deleting a Folder

Delete a folder if the folder contents become redundant.

1. In the Navigator, select the folder in the project you want to delete.
2. Click **Actions > Delete**.
   The Delete Folder dialog box appears.
3. Click **Yes** to delete the folder.
   The Analyst tool deletes the folder from the Navigator.

Viewing a Project or Folder

You can view objects in a project or folder. For each object, you can view the object name and object type on the **Contents** view. You can view descriptive information such as project name and description about the project or folder on the **Properties** view.

1. To view project or folder contents on the **Contents** view, select a project or folder in the Navigator and view the contents in the **Contents** panel.
   The Analyst tool displays a list of all objects in the project or folder and displays the object name and object type for each object.
2. To view descriptive information about the project or folder on the **Properties** view, select a project or folder from the Navigator and view descriptive information in the **Properties** panel.
   The Analyst tool displays the project name or folder name and description for the project or the folder.

Objects Overview

The types of objects that you use in the Analyst tool depend on the structure of data for which you want to analyze data quality. You can use data objects to structure the data and create object types to analyze data quality in a project.

Data objects can include the relational tables and flat files that you import into the Analyst tool. Logical data objects created in Data Object Models in the Developer tool appear as logical data objects in projects shared by the developer in the Analyst tool. These logical data object can appear as tables or flat files.

Object types include objects such as profiles, rules, scorecards, reference tables, bad record tables, and duplicate record tables that you create in the Analyst tool.

You can store objects in projects or folders.

You can perform the following common tasks on object types and data objects:

- View an object.
- Duplicate an object.
- Rename an object.
- Move an object.
- Delete an object.
Note: You cannot duplicate, rename, move, or delete a logical data object.

Viewing an Object

You can view object properties for each object in a project or folder. You can open the object to preview data in a tab. You can preview the contents of data objects and object types to view the structure of data and analyze data quality results.

1. In the Navigator, select the project or folder that contains the object you want to view.
2. In the Contents panel, select the object you want to view.
   
   The Analyst tool displays the name, type, and location of the object in the project or folder in the Properties panel. You can view connection name, Data Object Model name, table name, and schema name for table objects. Additionally, you can view the file path for flat file objects.
3. Click Actions > Open.
   
   The Analyst tool opens the object contents for preview in a tab. You can preview column metadata for tables and flat files and data quality results for other object types.

Duplicating an Object

You can duplicate objects within a project or within folders in a project to use objects for different tasks.

1. In the Navigator, select the project or folder that contains the object you want to duplicate.
2. In the Contents panel, select the object you want to duplicate.
3. Click Actions > Duplicate.
   
   The Duplicate window appears.
4. Navigate to the location in the project where you want to duplicate the object.
   
   Optionally, enter the location.
5. Enter the name of the object.
6. Click OK.
   
   The Analyst tool duplicates the object to the location in the project or folder.

Renaming an Object

Rename an object to change its name according to business usage and naming convention.

1. In the Navigator, select the project or folder that contains the object you want to copy.
2. In the Contents panel, select the object you want to rename.
3. Click Actions > Rename.
   
   The Rename window appears.
4. Enter the object name.
5. Click OK.
   
   The Analyst tool renames the object with specified name.
Moving an Object

Move an object within a project to another location in the project to organize project contents. You cannot move an object to a target folder that is a child folder of the source folder.

1. In the Navigator, select the project that contains the object you want to move.
2. In the Contents panel, select the object you want to move.
3. Click Actions > Move.
   The Move window appears.
4. Navigate to the location where you want to move the object to a folder.
   Optionally, enter the location where you want to move the object to a folder.
5. Click OK.
   The Analyst tool moves the object to specified location in the project or folder.

Deleting an Object

Delete an object from a project or folder if the object becomes redundant.

1. In the Navigator, select the project or folder that contains the object you want to delete.
2. In the Contents panel, select the object you want to delete.
3. Click Actions > Delete.
4. In the Delete Object dialog box, click Yes.
   The Analyst tool deletes the object from the project or folder.

Object Properties

The following table describes the objects that you can store in a project and the viewable object properties:

<table>
<thead>
<tr>
<th>Object</th>
<th>Data Object / Object Type</th>
<th>Object Properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relational Tables</td>
<td>Data Object</td>
<td>Name. Name of the table in the model repository.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Location. Location of the table in the project or folder.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Connection. Name of the database relational connection.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Schema. Name of the database schema.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Table Name. Name of the relational table source.</td>
</tr>
<tr>
<td>Flat Files</td>
<td>Data Object</td>
<td>Location. Location of the flat file in the project or folder.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>File Path. File path of the flat file on a network drive.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Uploaded. File path of the uploaded flat file.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FileName. Name of the flat file.</td>
</tr>
<tr>
<td>Object</td>
<td>Data Object / Object Type</td>
<td>Object Properties</td>
</tr>
<tr>
<td>------------------------</td>
<td>---------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Logical data object</td>
<td>Data Object</td>
<td><strong>Name.</strong> Name of the table in the model repository. <strong>Location.</strong> Location of the table in the project or folder. <strong>Data Object Model.</strong> Name of the Data Object Model from which the logical data object was created. <strong>Logical Data Object Name.</strong> Logical data object table name.</td>
</tr>
<tr>
<td>Profiles</td>
<td>Object Type</td>
<td><strong>Location.</strong> Location of the profile in the project or folder. <strong>Name.</strong> Name of the profile.</td>
</tr>
<tr>
<td>Rules</td>
<td>Object Type</td>
<td><strong>Location.</strong> Location of the rule in the project or folder. <strong>Name.</strong> Name of the rule.</td>
</tr>
<tr>
<td>Scorecards</td>
<td>Object Type</td>
<td><strong>Location.</strong> Location of the scorecard in the project or folder. <strong>Name.</strong> Name of the scorecard.</td>
</tr>
<tr>
<td>Reference Tables</td>
<td>Object Type</td>
<td><strong>Name.</strong> Name of the table in the model repository. <strong>Location.</strong> Location of the table in the project or folder. <strong>Connection.</strong> Name of the database relational connection. <strong>Schema.</strong> Name of the database schema. <strong>Table Name.</strong> Name of the reference table.</td>
</tr>
<tr>
<td>Bad Record Tables</td>
<td>Object Type</td>
<td><strong>Name.</strong> Name of the table in the model repository. <strong>Location.</strong> Location of the table in the project or folder. <strong>Connection.</strong> Name of the database relational connection. <strong>Schema.</strong> Name of the database schema. <strong>Table Name.</strong> Name of the bad record table.</td>
</tr>
<tr>
<td>Duplicate Record Tables</td>
<td>Object Type</td>
<td><strong>Name.</strong> Name of the table in the model repository. <strong>Location.</strong> Location of the table in the project or folder. <strong>Connection.</strong> Name of the database relational connection. <strong>Schema.</strong> Name of the database schema. <strong>Table Name.</strong> Name of the duplicate record table.</td>
</tr>
</tbody>
</table>
Rules and Guidelines for Projects

Rules and guidelines for working with projects.

Use the following rules and guidelines when you work with projects:

- You cannot move a project in the Navigator.
- You can move folders within a project but you cannot move a folder into one of its own child folders in a project.
- You cannot duplicate a project in another project with the same name.
- You cannot duplicate a folder within a project to another folder in a different project.

Search

You can search for objects and folders in the Analyst tool. Search folders to find objects for a particular task such as profiling data or creating business rules.

You can search for data objects, object types, and folders by name in the Analyst tool. You cannot search for projects by name.

The Model Repository Service uses a search analyzer to index the metadata in the Model repository. The Analyst tool uses the search analyzer to perform searches on objects in the Model repository.

You can create a search query and filter the search results. You can view search results and select an object from the results to view its contents. You can select an object from the search results and view its contents in another tab.

You can search in different languages. To search in a different language, an administrator must change the search analyzer and configure the Model repository to use the search analyzer. You can change the search analyzer in the Model Repository Service. After you change the search analyzer, you must restart the Model Repository Service and re-index the search index. For more information about changing the search analyzer, see the Informatica Administrator Guide.

Search Syntax

Use search syntax to create a search query and filter search results.

The following table describes the search syntax you can use in a search:

<table>
<thead>
<tr>
<th>Search Syntax</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Keywords</td>
<td>Use an exact keyword match in the search.</td>
</tr>
<tr>
<td>Cases</td>
<td>Use upper case and lower case text in the search.</td>
</tr>
<tr>
<td>Wildcards</td>
<td>Use wildcard characters in the search.</td>
</tr>
<tr>
<td>Logical Operators</td>
<td>Use logical operators in the search.</td>
</tr>
</tbody>
</table>
Keyword Matches
Use a keyword match to search for folders and objects that match the keyword.

Enclose a search query in quotation marks (" ") to search for an exact keyword match. The Analyst tool returns objects with the name that matches the keyword exactly.

**Note:** You cannot use wildcards or special characters in a search.

Wildcards
Use wildcards to define one or more characters in a search. Use wildcards as a prefix, suffix, or infix in a search.

The following table describes the wildcards you can use in a search:

<table>
<thead>
<tr>
<th>Wildcard</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>*</td>
<td>Represents characters. For example when you search for customer*, the Analyst tool can return customer, customer_name, and CustomerID.</td>
</tr>
<tr>
<td>?</td>
<td>Represents a single character. For example when you search for Customer?, the Analyst tool can return Customer1, Customer2, and CustomerA.</td>
</tr>
</tbody>
</table>

Operators
Use boolean search operators to logically combine search terms. All boolean operators must be upper case.

The following table describes the search operators you can use in a search:

<table>
<thead>
<tr>
<th>Operator</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AND</td>
<td>Includes both search terms. For example, sales data AND data sales.</td>
</tr>
<tr>
<td>OR</td>
<td>Includes either one of the search terms. For example, sales data OR sales.</td>
</tr>
<tr>
<td>NOT</td>
<td>Excludes a search term. For example, sales data NOT data sales. The NOT operator requires two operands.</td>
</tr>
</tbody>
</table>

Performing a Search
Perform a search to search for folders and objects in the Analyst tool.

> In the **Search** header box, enter a keyword or a search query.

The Analyst tool returns the results of the search in another tab.

Searching Objects Example
For example, an analyst wants to search for objects that contain customer metadata for the East or West regions. These objects could be in many projects that analyze data quality for different regions. The analyst can create a
search query that filters search results to find objects that contain Customer in the object name for either of these two regions.

1. In the Search box, enter the search query CustomersWest Or East.
   The Analyst tool displays all the objects that have CustomerWest or East in the object name in a Search tab.
2. Preview column metadata for search results. The Analyst tool displays the Name, Type, and Location columns for each object that the Analyst tool returns.
3. Optionally, select the Name and Type columns and click Columns and select the columns you want to include in the preview. Default includes all columns.
4. Click the CustomerWest table object in the search results.
   The Analyst tool opens the CustomerWest table in a tab. You can preview the column metadata for the CustomerWest table on the Data Preview view and the table properties on the Properties view.

Manage Security

Manage permissions on projects in the Analyst tool to control access to projects. You can add users to a project and assign permissions for users on a project.

Even if a user has the privilege to perform certain actions, the user may also require permission to perform the action on a particular object.

When you create a project, you are the owner of the project by default. The owner has all permissions, which you cannot change. The owner can assign permissions to users.

The following table describes the permissions you can assign for users on a project.

<table>
<thead>
<tr>
<th>Permission</th>
<th>Grants users the ability to</th>
</tr>
</thead>
<tbody>
<tr>
<td>Read</td>
<td>Read on project, view projects and objects in projects.</td>
</tr>
<tr>
<td>Write</td>
<td>Modify projects, create, edit, and delete objects in projects.</td>
</tr>
<tr>
<td>Grant</td>
<td>Grants users the ability to manage the read, write, and grant permissions on a project.</td>
</tr>
</tbody>
</table>

Assigning Permissions on a Project

You can add users to a project and assign permissions on a project to restrict, provide access, or manage the objects within the project.

1. In the Navigator, select a project, and click the Security tab.
2. In the Project-Level Permissions View on the Contents panel, click the Edit icon to edit user permissions.
   The Edit Project-Level Permissions dialog box appears.
3. Select a user from the Users panel.
4. Optionally, click Add to add another user.
   The Add Users dialog box appears.
5. Select a user or users that need to have access to the project.
6. Click OK.
7. Select or clear the Read, Write, or Grant permissions in the **Permissions** panel.
8. Click **OK**.

**Rules and Guidelines for Security**

Rules and guidelines for managing security.

Use the following rules and guidelines to manage security:

- A user can have the privilege to create a project but can view only those projects on which the user is assigned permissions.
- A user with the privilege to create a project but not the permission to view a project can create a project with a name that already exists. The Analyst tool displays a message stating that the project already exists.
- A user might be able to bypass Analyst tool security and view projects that the user does not have permission to view. The Analyst tool cannot prevent users from guessing and therefore determining project names they do not have the permission to view.
- When you assign the permission to view a project to a user, the user needs to refresh the Navigator to view the project. An administrator can view all projects.
Data Objects

This chapter includes the following topics:

- Data Objects Overview, 17
- Flat Files, 18
- Tables, 22
- Viewing Data Objects, 23

Data Objects Overview

Data objects represent the metadata sources from which you want to extract metadata to analyze in an Analyst tool project. You can import data objects such as tables and flat files to analyze the structure of the data in a project.

Data objects appear when you select the project or folder that contains the object in the Analyst tool. Any table or flat file that you add to a project in the Analyst tool also appears in the Developer tool. A table appears under the name of the connection for the table. A flat file appears with the file object name.

Logical data objects created in Data Object Models in the Developer tool appear as logical data objects in projects shared by the developer in the Analyst tool. These appear as Logical Data Objects in the Analyst tool. You cannot rename, move, or delete logical data objects. You can view logical data objects and create profiles and scorecards for logical data objects.

Use table and flat file objects to profile source data and perform data analysis tasks. You can add data objects by importing them into the Analyst tool. You can store data objects in projects and folders in the Navigator.

Before you can import a data object you must access the metadata source to extract the metadata that you analyze in the data object. The Analyst tool requires a connection to the source relational table to extract metadata for the table data object. The Analyst tool requires the network path or browse location to locate the source flat file to extract metadata for the flat file data object.

After you add tables and files, you can create a profile for the source data that the tables and files represent. When you run the profile, the Analyst tool connects to the database table or flat file.

You can perform the following tasks on data objects:

- **Add data objects.** Add tables and flat files by importing tables and flat files into projects and folders in the Navigator.
- **View data objects.** View object properties and column metadata for tables, flat files, and logical data objects.
- **Duplicate data objects.** Duplicate tables and flat files to use them for different tasks.
- **Rename data objects.** Rename tables and flat files according to their business usage and naming convention.
- **Move data objects.** Move tables and flat files in a project or folders in a project.

- **Delete data objects.** Delete tables and flat files when they become redundant.

  **Note:** If you delete a data object that other object types reference, the Analyst tool displays a message that lists the object types being referenced. Determine the impact of deleting the data object before you choose to delete it.

---

**Flat Files**

A flat file data object contains the metadata for a flat file in the Analyst tool. Use flat files to profile source data. When you add a flat file, the Analyst tool connects to the network path location or the location where you upload the source flat file to extract metadata.

You can add flat files in the Analyst tool by importing the flat files into projects or folders. Before you import a flat file, you can choose to browse a file from your local machine. The Analyst tool uploads a copy of the flat file to a directory in the Informatica Services installation directory that the Analyst tool can access. Or, you can point the Analyst tool to a network location. The Analyst tool uses the location you specify to access the source flat file.

You can specify the directory where you upload flat files in the flat file cache when you configure the Analyst Service in the Administrator tool. For more information about specifying the flat file cache, see the *Informatica Administrator Guide*.

Use the **Add Flat File Wizard** to import flat files into the Analyst tool. To add a flat file in the Analyst tool, select the flat file, configure the file options, and configure the column data types. After you add the flat file, you can preview the flat file properties and column metadata in the flat file.

**Flat File Options**

When you import a flat file, you can configure the flat file options for each column in **Add Flat File wizard**. The options that you configure determine how the wizard reads the data from the source flat file.

The following table describes the flat file options that you configure the in the **Add Flat File** wizard:

<table>
<thead>
<tr>
<th>Options</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Delimiters</strong></td>
<td>Character used to separate columns of data. Use the Other field to enter a different delimiter. Delimiters must be printable characters and must be different from the escape character and the quote character if selected. You cannot select non-printing multibyte characters as delimiters.</td>
</tr>
<tr>
<td><strong>Text Qualifier</strong></td>
<td>Quote character that defines the boundaries of text strings. Choose No Quote, Single Quote, or Double Quotes. If you select a quote character, the wizard ignores delimiters within pairs of quotes.</td>
</tr>
</tbody>
</table>

---

Chapter 3: Data Objects
<table>
<thead>
<tr>
<th>Options</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Column Names</td>
<td>Option to import column names from the first line. The wizard uses data in the first row in the preview for column names. Select this option if column names appear in the first row. Default is not enabled.</td>
</tr>
<tr>
<td>Values</td>
<td>Option to start value import from a line. Indicates the row number in the preview at which the wizard starts reading when it imports the file.</td>
</tr>
</tbody>
</table>

**Flat File Datatypes**

When you import a flat file, you can configure the datatypes for the data in each column in the Add Flat File wizard. The datatypes you configure determine how the wizard imports the data from the source flat file.

You can configure the following data types for the data in each column in the Add Flat File wizard:

- **bigint.** You can specify the format in the Numeric Format window. You can use the default or specify another numeric format and choose to make this the default numeric format.
- **datetime.** You can specify the format in the Datetime Format window. You can use the default or specify another datetime format and choose to make this the default datetime format.
- **double.** You can specify the format in the Numeric Format window. You can use the default or specify another numeric format and choose to make this the default numeric format.
- **int.** You can specify the format in the Numeric Format window. You can use the default or specify another numeric format and choose to make this the default numeric format.
- **nstring.** You cannot specify a format.
- **number.** You can specify the format in the Numeric Format window. You can use the default or specify another numeric format and choose to make this the default numeric format.
- **string.** You cannot specify a format.

**Datetime Datatypes**

When you import a flat file, you can configure the datatypes for file columns in the Add Flat File wizard. When you configure the datetime dataype, you can specify the format in the Datetime Format window. You can use the default or specify another datetime format and choose to make this the default datetime format.

The following table describes the datetime format strings to specify as part of the date:

<table>
<thead>
<tr>
<th>Format String</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AM, a.m., PM, p.m.</td>
<td>Meridian indicator. Use any of these format strings to specify AM and PM hours. AM and PM return the same values as do a.m. and p.m.</td>
</tr>
<tr>
<td>DAY</td>
<td>Name of day, including up to nine characters (for example, Wednesday). The DAY format string is not case sensitive.</td>
</tr>
<tr>
<td>DD</td>
<td>Day of month (1-31).</td>
</tr>
<tr>
<td>DDD</td>
<td>Day of year (001-366, including leap years).</td>
</tr>
<tr>
<td>Format String</td>
<td>Description</td>
</tr>
<tr>
<td>---------------</td>
<td>-------------</td>
</tr>
<tr>
<td>DY</td>
<td>Abbreviated three-character name for a day (for example, Wed). The DY format string is not case sensitive.</td>
</tr>
<tr>
<td>HH, HH12</td>
<td>Hour of day (1-12).</td>
</tr>
<tr>
<td>HH24</td>
<td>Hour of day (0-23), where 0 is 12AM (midnight).</td>
</tr>
<tr>
<td>J</td>
<td>Modified Julian Day.</td>
</tr>
<tr>
<td>MI</td>
<td>Minutes (0-59).</td>
</tr>
<tr>
<td>MM</td>
<td>Month (01-12).</td>
</tr>
<tr>
<td>MONTH</td>
<td>Name of month, including up to nine characters (for example, August). Case does not matter.</td>
</tr>
<tr>
<td>MON</td>
<td>Abbreviated three-character name for a month (for example, Aug). Case does not matter.</td>
</tr>
<tr>
<td>MS</td>
<td>Milliseconds (0-999).</td>
</tr>
<tr>
<td>NS</td>
<td>Nanoseconds (0-999999999).</td>
</tr>
<tr>
<td>RR</td>
<td>Four-digit year (for example, 1998, 2034). Use when source strings include two-digit years.</td>
</tr>
<tr>
<td>SS</td>
<td>Seconds (0-59).</td>
</tr>
<tr>
<td>SSSSSS</td>
<td>Seconds since midnight.</td>
</tr>
<tr>
<td>US</td>
<td>Microseconds (0-999999).</td>
</tr>
<tr>
<td>Y</td>
<td>The current year with the last digit of the year replaced with the string value.</td>
</tr>
<tr>
<td>YY</td>
<td>The current year with the last two digits of the year replaced with the string value.</td>
</tr>
<tr>
<td>YYY</td>
<td>The current year with the last three digits of the year replaced with the string value.</td>
</tr>
<tr>
<td>YYYY</td>
<td>Four digits of a year. Do not use this format string if you are passing two-digit years. Use the RR or YY format string instead.</td>
</tr>
</tbody>
</table>

## Adding a Flat File

Use the **Add Flat File Wizard** to add a flat file to a project or folder. You can add flat files to projects and folders to analyze the structure of the data before you perform data quality tasks on the data.

1. In the Navigator, select the project or folder that you want to add the flat file to.
2. Click **Actions > New > Flat File**. The **Add Flat File wizard** appears.
3. Select **Browse and Upload** and click **Browse** to select the flat file and click **Upload** to upload it to the machine on which Informatica Analyst runs. Or, select **Enter a Network Path** and configure the path and file name of the file.

4. Click **Next**.

   If you chose to upload the file, Informatica Analyst uploads the flat file to an Informatica Services installation directory that Informatica Analyst can access.

5. Configure the name, description, and code page, and preview the flat file data.

   **Note:** Select a code page that matches the code page of the data in the file.

6. Click **Next**.

7. Configure the flat file options.

8. Optionally, click **Show** to preview changes to the flat file data.

9. Click **Next**.

10. Configure the datatype for each column.

11. Optionally, configure the precision, scale, and format for each column. The properties you can edit depend on the datatype.

12. Click **Finish**.

   The Analyst tool adds the flat file to the project or folder in the Navigator.

---

**Rules and Guidelines for Flat Files**

Rules and Guidelines for working with flat files.

Use the following rules and guidelines while working with flat files:

- **Upload small files.** Use the option to upload small files to an Informatica Services installation directory on the machine where the Analyst tool runs. The Analyst tool accesses this location to extract flat file metadata that does not change frequently. When you use small files of sizes up to 10MB, the Analyst tool accesses a copy of the file in the Informatica Services installation directory. If you modify the original file, you need to upload the file again.

- **Upload large files.** Use the option to enable the Analyst tool to connect to a network path location for large files. The Analyst tool accesses this location to extract flat file metadata that changes frequently. The network path location should be a shared directory or file system that the Analyst tool can access. When you use large file sizes greater than 10MB, the Analyst tool can connect to the flat file in the network path. If you modify the original flat file, you must refresh the flat file in the Analyst tool. Refreshing metadata for a large flat file can take time.

- **Blank data rows.** The Analyst tool does not import the blank rows above the first data row, blank middle rows, and blank rows after the last data row when importing a flat file.

- **Previewing data.** After a preview, you can change the row number at which the **Add Flat File** wizard starts reading when it imports the file. This row number corresponds with the preview. If you choose to import column names from the first line, refresh the preview to update the row numbers for the preview data.
Tables

A table object contains the metadata for a relational database source in the Analyst tool. Use tables to profile source data. When you add a table, the Analyst tool uses a database connection to connect to the source database to extract metadata.

You can add tables in the Analyst tool by importing the tables into projects or folders. Before you import a table, you select or create a database connection, and select the database table that you want to add. You can add multiple tables from a connection as data objects.

Use the **Add Table Wizard** to add a table to the project or folder.

You can use the database connection that the staging database uses to select a table. An administrator creates the connection to the staging database that stores reference tables for the Analyst tool before configuring the Analyst Service. For more information on configuring the staging database, see the *Informatica Administrator Guide*.

You can create another database connection to connect to source relational tables. The following table describes the database connection options that you can configure for a database connection:

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name</td>
<td>Name of the connection. Connection names cannot have spaces and cannot be longer than 128 characters.</td>
</tr>
<tr>
<td>Description</td>
<td>Description of the connection.</td>
</tr>
<tr>
<td>Database Type</td>
<td>Type of relational database. You can select an Oracle, Microsoft SQL Server, or IBM DB2 database.</td>
</tr>
<tr>
<td>Username</td>
<td>User name used for authentication when you connect to the relational database.</td>
</tr>
<tr>
<td>Password</td>
<td>Password for the database user name.</td>
</tr>
<tr>
<td>Connection String</td>
<td>Connection string used to access data from the database.</td>
</tr>
<tr>
<td></td>
<td>IBM DB2: <code>&lt;database name&gt;</code></td>
</tr>
<tr>
<td></td>
<td>Microsoft SQL Server: <code>&lt;server name&gt;@&lt;database name&gt;</code></td>
</tr>
<tr>
<td></td>
<td>ODBC: <code>&lt;data source name&gt;</code></td>
</tr>
<tr>
<td></td>
<td>Oracle: <code>&lt;database name listed in TNSNAMES entry&gt;</code></td>
</tr>
<tr>
<td>JDBC URL</td>
<td>JDBC connection URL used to access metadata from the database.</td>
</tr>
<tr>
<td></td>
<td>IBM DB2: <code>jdbc:informatica:db2://&lt;host name&gt;:&lt;port&gt;;DatabaseName=&lt;database name&gt;</code></td>
</tr>
<tr>
<td></td>
<td>ODBC: n/a</td>
</tr>
<tr>
<td></td>
<td>Oracle: <code>jdbc:informatica:oracle://&lt;host_name&gt;:&lt;port&gt;;SID=&lt;database name&gt;</code></td>
</tr>
<tr>
<td></td>
<td>Microsoft SQL Server: <code>jdbc:informatica:sqlserver://&lt;host name&gt;:&lt;port&gt;;DatabaseName=&lt;database name&gt;</code></td>
</tr>
<tr>
<td>Code page</td>
<td>Code page use to read from a source database or write to a target database or file.</td>
</tr>
</tbody>
</table>
Adding a Table

Use the Add Tables Wizard to add a table to a project. Add the tables that you want to profile data for. To add a table, select or create a connection, select the schema and tables, and add the table.

1. In the Navigator, select the project or folder that you want to add the table to.
2. Click Actions > New > Table.
   The Add Tables wizard appears.
3. Select a connection. Optionally, click New Connection to create a connection and configure the connection properties.
4. Click Next.
5. Optionally, unselect Show Default Schema Only to show all schemas associated with the selected connection.
6. Select the table that you want to add or enter a table name in the search box and click Go to search by table name. Click Clear to remove the search results and display all tables.
7. Optionally, click the Properties View to view the properties and column metadata for the table. Or, click the Data Preview View to view the columns and data for the table.
8. Click Next.
   The wizard displays the table to add to your folder or project.
9. Click Finish.

Rules and Guidelines for Tables

Rules and guidelines for working with table data objects.

Use the following rules and guidelines while working with tables:

- **Previewing data.** The Analyst tool displays the first 100 rows by default when you preview the data for a table. The Analyst tool may not display all the data columns in a wide table.
- **Wide tables.** The Analyst tool can import wide tables with more than 30 columns for profiling data. When you import a wide table, the Analyst tool does not display all the columns in the data preview. The Analyst tool displays the first 30 columns in the data preview. However, you can include all the columns in the wide tables and flat files for profiling.

Viewing Data Objects

View data objects to preview data object properties and column metadata.

1. In the Navigator, select the project or folder that contains the table or file data object.
2. Click Actions > Open to open the data object.
   The data object appears in a new tab. The Analyst tool retrieves the first 100 rows for the data object and displays it on the Data Preview view.
3. Optionally, select Columns and select a column to include in the preview. Default includes all columns for preview.
   If you choose not to include a column, the Analyst tool refreshes the preview and does not include the column in the preview.
4. Optionally, click the **Properties** view to view the table and file properties in the **Properties** panel.

   The Analyst tool displays the table name, description, location, connection name, and database schema name for the table data object. The Analyst tool displays the file name, location, upload file path, or network path for the flat file data object.

5. Optionally, view the column metadata for each column in the **Columns** panel.

   You can view column name and datatype for each column in the table or flat file. You can view if the column is nullable and the key for each column in the table. Nullable and key properties are relational database properties.

6. Optionally, click the **Refresh** button to refresh the metadata for the data object.
Profiles and Rules

This chapter includes the following topics:

- Profiles and Rules Overview, 25
- Configure Profiles and Rules, 26
- Rule Types, 29
- Profile Results, 33
- Rules and Guidelines for Profiles, 37

Profiles and Rules Overview

A profile is the analysis of data quality based on the content and structure of data. A rule is business logic that defines conditions applied to data when you run a profile. Use profiles to analyze the structure, content, and anomalies present in the data, and apply or create data rules to cleanse and standardize the data. Use profiles to create scorecards before and after data standardization.

You can create a profile for tables and flat files. You can create an expression rule or apply a predefined rule to a profile to modify the data in the profile columns. You can make a rule reusable to use it within other profiles. Run a profile to analyze the profile results and drill down to find the underlying rows for the values and patterns in the profile results.

You can view profile results as a chart in the Analyst tool. You can export the profile results to a CSV file. You can add comments to a profile to share information about the profile with other Analyst tool users.

For example, an analyst can create a profile for a Customer table by including all columns of the Customer table in the profile. The analyst can analyze and view the value and pattern distribution for the columns. The analyst can decide to drill down on the value and pattern frequencies for the CustomerID column to inspect records that have invalid customer numbers. The analyst can create an expression rule named IsValidCustomerID within the profile to look at the percentage of records with valid customer numbers. The IsValidCustomerID rule that appears in the profile considers all non-numeric customer numbers to be invalid. The analyst selects the output field of the IsValidCustomerID rule and profiles the Customer table again to view valid customer numbers in the CustomerID column.

When you create a profile in the Analyst tool, the profile also appears in the Developer tool. The Analyst tool uses column profiling to profile columns in a data object. Profile objects that are created by a developer in the Developer tool can appear in the Analyst tool if the profiling method used was column profiling. The Developer tool uses Column Profiling, Join Analysis, and Midstream profiling methods to profile data objects. For more information about Join Analysis and Midstream profiling, see the Informatica Developer User Guide.

Use the New Profile Wizard to create a profile. You can create a profile in the following ways:
Create a profile in a project. Create a profile in a project, select a data object from any project, and configure the columns and sampling and drilldown options.

Create a quick profile. Select a data object and create a quick profile to use the default settings for all columns and sampling and drilldown options.

Create a custom profile. Select a data object and create a custom profile to configure the columns and sampling and drilldown options.

Configure Profiles and Rules

When you create a profile, you select the columns in the data object for which you want to profile data. You can set or configure sampling and drilldown options for faster profiling. After you run the profile, you can examine the profiling statistics to understand the data. You can add a rule to the profile to cleanse, modify, or validate the data. Run the profile again to view the data.

The following steps describe how to configure a profile and add rules to to perform data quality analysis on the data:

1. Create a profile. Use the New Profile window to create a profile by selecting the columns in the data object that you want to profile data for. You can create a profile in a project, create a custom profile, or a quick profile.

2. Save and run the profile.

3. View the profile results and examine the profiling statistics to understand the data. You can determine if the data has unique values, data patterns, minimum and maximum values, and look for rows that match selected patterns or values.

4. After you examine the data, you can determine which rule to add to the profile. For example, you can add a rule because you want a cleansed version of one or more of the data object columns. You can add a lookup rule that provides information that the source data does not provide. You can add a rule to validate the correctness of a cleansing rule that you intend to use in a data quality or data integration project.

5. Add the rule to the profile. You can select a predefined rule, or create an expression rule.

6. Save and run the profile.

7. View the profile results.

Wide Tables and Files

You can profile wide tables and flat files that have a large number of columns. You can profile tables with more than 30 columns and flat files with more than 100 columns.

When you create or run a profile, you can choose to select all the columns or select each column you want to include for profiling. The Analyst tool displays the first 30 columns in the data preview. You can select all columns for drilldown and view value frequencies for these columns. You can use rules that have more than 50 output fields and include the rule columns for profiling when you run the profile again.

Profile Options

You can create quick and custom profiles for a data object. You configure the sampling and drilldown options when you create a profile.

Use a quick profile to include all columns for a data object and use the default sampling and drilldown options. Use a custom profile to select the columns for a data object and to configure the sampling and drilldown options.
Sampling Options

The following table describes the sampling options for a profile:

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Rows</td>
<td>The Analyst tool profiles all rows in the data object. This is the default option for a quick profile.</td>
</tr>
<tr>
<td>First &lt;number&gt; Rows</td>
<td>Number of rows that the Analyst tool profiles. The Analyst tool chooses the rows from the first rows in the table or file.</td>
</tr>
<tr>
<td>Random Sample &lt;number&gt; Rows</td>
<td>Number of rows to randomly sample in the table or file. Random sampling forces the drill down to be performed on staged data stored in the profiling warehouse that can slow drill down performance.</td>
</tr>
<tr>
<td>Random sample</td>
<td>The Analyst tool chooses the sample size, based on the number of rows in the data object. Random sampling forces the drill down to be performed on staged data stored in the profiling warehouse that can slow drill down performance.</td>
</tr>
</tbody>
</table>

Drilldown Options

The following table describes the drilldown options for a profile:

<table>
<thead>
<tr>
<th>Options</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enable Row Drilldown</td>
<td>Drill down to specific records from the profile results. You can select one of the following options:</td>
</tr>
<tr>
<td></td>
<td>- <strong>Show Profiled Columns Only.</strong> Only show the columns that you select to profile.</td>
</tr>
<tr>
<td></td>
<td>You can select this option if you configure specific columns to profile.</td>
</tr>
<tr>
<td></td>
<td>- <strong>Show All Columns.</strong> Show all columns in the profile results.</td>
</tr>
<tr>
<td></td>
<td>This is the default option for a quick profile.</td>
</tr>
<tr>
<td>Drilldown</td>
<td>Type of data to drill down to. You can select one of the following options:</td>
</tr>
<tr>
<td></td>
<td>- <strong>on live data.</strong> Access the row data that you drill down to on the source.</td>
</tr>
<tr>
<td></td>
<td>This is the default option for a quick profile.</td>
</tr>
<tr>
<td></td>
<td>- <strong>on staged data.</strong> The Analyst tool stages the row data in the profiling warehouse.</td>
</tr>
</tbody>
</table>

Creating a Profile in a Project

Use the **New Profile Wizard** to create a profile in a project. When you create a profile, you select a project, select the data object, and configure the columns and sampling and drilldown options.

1. In the Navigator, select the project where you want to create the profile.
2. Click **Actions > New Profile.**
   The **New Profile Wizard** appears and offers the option to create a custom profile.
3. Click **Next.**
4. Enter a name and an optional description for the profile.
5. In the Folders panel, select the folder where you want to create the profile.
   The Analyst tool displays the project that you selected and shared projects that contain folders where you can create the profile. The profile objects in the folder appear in the Profiles panel.
6. Click Next.
7. In the Sources panel, select the data object that you want to create a profile for.
   The Columns panel displays the columns and datatype of the columns in the data object.
8. Click Next.
9. In the Select Columns panel, select the columns you want to profile. The columns include any rules you applied to the profile. The Analyst tool lists the name, datatype, precision, and scale for each column.
   Optionally, select Name to include all columns.
10. In the Sampling Options panel, configure the sampling options.
11. In the Drilldown Options panel, configure the drilldown options.
12. Click Save to create the profile or click Save & Run to create the profile and then run the profile.

Creating a Quick Profile

Select a data object and create a quick profile to use the default settings for all columns and sampling and drilldown options. The Analyst tool creates the profile in the same project and folder as the data object.

1. In the Navigator, select the project that contains the data object that you want to create a quick profile for.
2. In the Contents panel, select the data object.
   Do not click the link for the object.
3. Click Actions > New Profile.
   The New Profile wizard appears.
4. Select Quick Profile.
5. Click Next.
6. Click Save to create the profile or click Save & Run to create the profile and then run the profile.

Creating a Custom Profile

Select a data object and create a custom profile to configure the columns and sampling and drilldown options. The Analyst tool creates the profile in the same project and folder as the data object.

1. In the Navigator, select the project that contains the data object that you want to create a custom profile for.
2. In the Contents panel, select the data object.
   Do not click the link for the object.
3. Click Actions > New Profile.
   The New Profile wizard appears.
4. Select Custom Profile.
5. Click Next.
6. Enter a name and an optional description for the profile.
7. In the Folders panel, select the folder where you want to create the profile.
The Analyst tool displays the project that you selected and shared projects that contain folders where you can create the profile. The profile objects in the folder appear in the Profiles panel.

8. Click Next.

9. In the Columns panel, select the columns you want to profile. The columns include any rules you applied to the profile. The Analyst tool lists the name, datatype, precision, and scale for each column.
   Optionally, select Name to select all columns.

10. In the Sampling Options panel, configure the sampling options.

11. In the Drildown Options panel, configure the drilldown options.

12. Click Save to create the profile or click Save & Run to create the profile and then run the profile.

Running a Profile

Run a profile to analyze a data source for content and structure.

1. In the Navigator, select the project or folder that contains the profile you want to run.

2. Click the profile to open it.
   The profile appears in a tab.

3. Click Actions > Run Profile.
   The Column Profile window appears.

4. In the Columns panel, select the columns you want to profile. The columns include rules you applied to the profile.
   Optionally, select Name to select all columns. The Analyst tool lists the name, datatype, precision, and scale for each column.

5. In the Sampling Options panel, configure the sampling options

6. In the Drildown Options panel, configure the drilldown options.

7. Click Run.
   The Analyst tool displays the profile results.

Rule Types

You can add rules after you create a profile. You can add predefined rules to a profile or create expression rules to add to a profile. After you create an expression rule, you can make it reusable.

Add rules to a profile by selecting a predefined rule or create an expression rule.

You can create or apply the following rule types for a profile:

- **Expression rules.** Use expression functions and columns to define rule logic. Create expression rules in the Analyst tool.

- **Predefined rules.** Includes system-defined, user-defined, and reusable rules. System-defined rules are packaged with the content installer for the Developer tool and can appear as reusable rules in the Analyst tool. Rules that are created in the Developer tool as mapplets can appear in the Analyst tool as user-defined rules. An analyst can create an expression rule and promote it to a reusable rule that other analysts can use in multiple profiles.

The output of a rule can be one or more virtual columns. The virtual columns exist only in the profile results. The Analyst tool profiles the virtual columns. For example, you use a predefined rule that splits a column that contains
first and last names into FIRST_NAME and LAST_NAME virtual columns. The Analyst tool profiles the
FIRST_NAME and LAST_NAME columns.

**Note:** If you delete a rule object that other object types reference, the Analyst tool displays a message that lists
the object types being referenced. Determine the impact of deleting the rule before you choose to delete it.

### Expression Rules

Expression rules use expression functions and columns to define rule logic. Create expression rules and add them
to a profile in the Analyst tool.

Use expression rules to modify or validate values for columns in a profile. You can create one or more expression
rules to use in a profile.

Expression functions are SQL-like functions used to transform source data. The Analyst tool provides the following
types of functions that you can use to create expression rule logic:

- Character
- Conversion
- Data Cleansing
- Date
- Encoding
- Financial
- Numerical
- Scientific
- Special
- Test

Use the **New Rule** Wizard to create an expression rule and add it to a profile. You must open a profile before you
create an expression rule. When you create an expression rule, you configure the name and description for the
rule and configure the rule logic using expression functions and columns as parameters. You can then include the
rule in profile results and configure sampling options.

The **New Rule** Wizard includes an expression editor. Use the expression editor to add expression functions,
configure columns as input to the functions, validate the expression, and configure the return type, precision, and
scale.

The output of an expression rule is a virtual column that uses the name of the rule as the column name. The
Analyst tool profiles the virtual column. For example, you use an expression rule to validate a ZIP code. The rule
returns 1 if the ZIP code is valid and 0 if the ZIP code is invalid. Informatica Analyst profiles the 1 and 0 output
values of the rule.

### Creating an Expression Rule

Use the **New Rule** Wizard to create an expression rule and add it to a profile. Create an expression rule to modify
or validate values for columns in a profile.

1. In the Navigator, select the project or folder that contains the profile that you want to add the rule to.
2. In the **Contents** panel, click the profile to open it.
   The profile appears in a tab.
3. Click **Actions > Add Rule**.
   The **New Rule** window appears.
4. Select **Create new rule**.
5. Click **Next**.
6. Enter a name and optional description for the rule.
7. Optionally, choose to promote the rule as a reusable rule and configure the project and folder location.
   
   If you promote a rule to a reusable rule, you or other users can use the rule in another profile as a predefined rule.
8. In the **Functions** tab, select a function and click the right arrow to enter the parameters for the function.
9. In the **Columns** tab, select an input column and click the right arrow to add the expression in the **Expression** editor. You can also add logical operators to the expression.
10. Click **Validate**. You can proceed to the next step if the expression is valid.
11. Optionally, click **Edit** to configure the return type, precision, and scale.
12. Click **Next**.
13. In the **Columns** panel, select the columns you want to profile. The columns include any rules you applied to the profile. Optionally, select **Name** to select all columns.
   
   The Analyst tool lists the name, datatype, precision, and scale for each column.
14. In the **Sampling Options** panel, configure the sampling options.
15. In the **Drilldown Options** panel, configure the drilldown options.
16. Click **Save** to create the rule or click **Save & Run** to create the rule and then run the profile.

### Creating an Expression Rule from Profile Results

You can run a profile on a data object and create an expression rule based on selected value frequencies for profile columns. You cannot create an expression rule from a virtual column in the profile. The rule uses the value frequencies from the profile results to create the rule syntax.

1. In the Navigator, select the project or folder that contains the profile that you want to add the rule to.
2. Click the profile to open it.
   
   The profile appears in a tab.
3. In the **Column Profiling** view, select a column and select the value frequencies for the column data in the **Values** view.
4. Click **Actions > Add Rule**.
   
   The **New Rule** window appears.
5. Select the option to **Create a new Value Frequency Rule**.
6. Click **Next**.
7. Enter a name and optional description for the rule.
8. Optionally, choose to promote the rule as a reusable rule and configure the project and folder location.
   
   If you promote a rule to a reusable rule, you or other users can use the rule in another profile as a predefined rule.
9. The Analyst tool adds the expression rule for the value frequencies you selected in the Expression editor.
   
   Optionally, in the **Functions** tab, select a function and click the right arrow to enter the parameters for the function. In the **Columns** tab, select an input column and click the right arrow to add the expression in the **Expression** editor. You can add logical operators to the expression.
10. Click **Validate**. You can proceed to the next step if the expression is valid.
11. Click **Next**.
12. In the Columns panel, select the columns you want to profile. The columns include any rules you applied to the profile. Optionally, select Name to include all columns.

   The Analyst tool lists the name, datatype, precision, and scale for each column.

13. In the Sampling Options panel, configure the sampling options.

14. In the Drilldown Options panel, configure the drilldown options.

15. Click Save to create the rule or click Save & Run to create the rule and then run the profile.

Predefined Rules

Predefined rules are rules created in the Developer tool or provided with the Developer tool and the Analyst tool. Apply predefined rules to the Analyst tool profiles to modify or validate source data.

Predefined rules use transformations to define rule logic. You can use predefined rules with multiple profiles. In the Model repository, a predefined rule is a mapplet with an input group, an output group, and transformations that define the rule logic.

You can apply the following types of predefined rules:

- **System-defined.** Packaged with the content installer for the Developer tool and can appear as reusable rules in the Analyst tool.
- **User-defined.** Created in the Developer tool as a mapplet and validated as a rule. Can appear as reusable rules in the Analyst tool.

Use the New Rule wizard to apply a predefined rule to a profile. You must open a profile before you apply a predefined rule. When you apply a predefined rule, you configure the name and description for the rule and configure the columns in the profile or reference tables as parameters for the rule. You can then include the rule in profile results and configure sampling and drilldown options.

Applying a Predefined Rule

Use the New Rule Wizard to apply a predefined rule to a profile. When you apply a predefined rule, you select the rule and configure the input and output columns for the rule. Apply a predefined rule to use a rule promoted as a reusable rule or use a rule created by a developer.

1. In the Navigator, select the project or folder that contains the profile that you want to add the rule to.
2. Click the profile to open it.
   The profile appears in a tab.
3. Click Actions > Add Rule.
   The New Rule window appears.
4. Select the option to Apply a Rule.
5. Click Next.
6. In the Rules panel, select the rule that you want to apply.
   The name, datatype, description, and precision columns appear for the Inputs and Outputs columns in the Rules Parameters panel.
7. Click Next.
8. In the Inputs section, select an input column. The input column is a column name in the profile.
9. Optionally, in the Outputs section, configure the label of the output columns.
10. Click Next.
11. In the **Columns** panel, select the columns you want to profile. The columns include any rules you applied to the profile. Optionally, select **Name** to include all columns.

   The Analyst tool lists the name, datatype, precision, and scale for each column.

12. In the **Sampling Options** panel, configure the sampling options.

13. In the **Drilldown Options** panel, configure the drilldown options.

14. Click **Save** to apply the rule or click **Save & Run** to apply the rule and then run the profile.

---

**Profile Results**

View profile results to understand the structure of the data and analyze data quality. You can view the profile results after you run a profile. You can view a summary of the columns and rules in the profile and the values, patterns, and statistics for columns and rules. You can view properties for the profile and properties for the columns and rules in the profile. You can preview profile data.

After you run a profile, you can view the profile results in the **Column Profiling**, **Properties**, and **Data Preview** views.

In the **Column Profiling** view, you can view the summary information for columns for a profile run. You can view values, patterns, and statistics for each column in the **Values**, **Patterns**, and **Statistics** views.

The Analyst tool displays rules as columns in profile results. The profile results for a rule appear as a profiled column. The profile results that appear depend on the profile configuration and sampling options.

The following table describes the profiling results that appear in the **Column Profiling** view:

- **Summary**: Summary information for the profile run, including the number of unique and null values, inferred datatype, and last run date and time.
- **Values**: Values for columns and the frequency in which the value appears for the column. The frequency appears as a number, a percentage, and a chart.
- **Patterns**: Value patterns for the profiled columns and the frequency in which the pattern appears. The frequency appears as a number and a percentage.
- **Statistics**: Statistics about the column values, such as average, length, and top and bottom values.

**Note**: You can select a value or pattern and view profiled rows that match the value or pattern on the Details panel.

In the **Properties** view, you can view profile properties on the **Properties** panel. You can view properties for columns and rules on the **Columns** and **Rules** panel.

In the **Data Preview** view, you can preview the profile data. The Analyst tool includes all columns in the profile displays the first 100 rows of data.

---

**Profile Summary**

The summary for a profile run includes the number of unique and null values expressed as a number and percentage, inferred datatype, and last run date and time.
You can click each profile summary property to sort on values of the property. The following table describes the profile summary properties:

<table>
<thead>
<tr>
<th>Property</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Column</td>
<td>Name of the column in the profile.</td>
</tr>
<tr>
<td>Unique Values</td>
<td>Number of unique values for the column.</td>
</tr>
<tr>
<td>Unique %</td>
<td>Percentage of unique values for the column.</td>
</tr>
<tr>
<td>Null Values</td>
<td>Number of null values for the column.</td>
</tr>
<tr>
<td>Null %</td>
<td>Percentage of null values for the column.</td>
</tr>
<tr>
<td>Inferred Data Type</td>
<td>Data type derived from the values for the column. The Analyst tool can derive the following datatypes from the datatypes of values in columns: - String - Varchar - Decimal - Integer - &quot;-&quot; for Nulls</td>
</tr>
<tr>
<td>Preview</td>
<td>If selected, includes the column in the data preview.</td>
</tr>
<tr>
<td>Sampling Policy</td>
<td>Sampling performed on a number of rows.</td>
</tr>
<tr>
<td>Last Run On</td>
<td>Date and time you last ran the profile.</td>
</tr>
</tbody>
</table>

Column Values

The column values include values for columns and the frequency in which the value appears for the column.

The following table describes the properties for the column values:

<table>
<thead>
<tr>
<th>Property</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value</td>
<td>List of all values for the column in the profile.</td>
</tr>
<tr>
<td>Frequency</td>
<td>Number of times a value appears for a column, expressed as a number, a percentage, and a chart.</td>
</tr>
<tr>
<td>Percent</td>
<td>Percentage that a value appears for a column.</td>
</tr>
<tr>
<td>Chart</td>
<td>Chart for the percentage.</td>
</tr>
</tbody>
</table>

Note: You can sort the Value and Frequency columns by selecting the columns. When you sort the results of the Frequency column, the Analyst tool sorts the results based on the datatype of the column.

Column Patterns

The column patterns include the value patterns for the columns and the frequency in which the pattern appears.
The following table describes the properties for the column patterns:

<table>
<thead>
<tr>
<th>Property</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pattern</td>
<td>Pattern for the column in the profile.</td>
</tr>
<tr>
<td>Frequency</td>
<td>Number of times a pattern appears for a column, expressed as a number.</td>
</tr>
<tr>
<td>Percent</td>
<td>Percentage that a pattern appears for a column.</td>
</tr>
<tr>
<td>Chart</td>
<td>Chart for the percentage.</td>
</tr>
</tbody>
</table>

**Column Statistics**

The column statistics include statistics about the column values, such as average, length, and top and bottom values.

For a string column, the Analyst tool displays the Maximum Length, Minimum Length, Bottom, and Top statistics.

For an integer column, the Analyst tool displays the Average, Standard Deviation, Maximum Length, Minimum Length, Bottom, and Top statistics.

The following table describes the types of column statistics:

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average</td>
<td>Average of the values for the column.</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>The standard deviation, or variability between column values, for all values of the column.</td>
</tr>
<tr>
<td>Maximum Length</td>
<td>Length of the longest value for the column.</td>
</tr>
<tr>
<td>Minimum Length</td>
<td>Length of the shortest value for the column.</td>
</tr>
<tr>
<td>Bottom</td>
<td>Three lowest values for the column.</td>
</tr>
<tr>
<td>Top</td>
<td>Three highest values for the column.</td>
</tr>
</tbody>
</table>

**Viewing Profile Results**

View profile results to see a summary of the columns and rules in the profile and the values, patterns, and statistics for columns and rules. You can view the rows from the data source that represent a specific value or pattern in the profile results.

1. In the Navigator, select the project or folder that contains the profile.
2. Click the profile to open it.
   The profile appears in a tab.
3. In the **Summary** view, select a column name to view the profile results for the column.
The values, patterns, and statistics for the column appear in the Results panel.

4. Select a column value on the Values tab or select a column pattern on the Patterns tab. Use CONTROL and SHIFT keys to select multiple values.

The Drilldown panel displays the rows that contain the values or patterns. The column value or pattern appears at the top of the panel.

Note: If you chose to drill down on live data and the profile does not contain any rules, the rows that appear are for the data in the source. Otherwise, the data that appears is staged in the profiling warehouse.

Previewing a Table or File

You can preview the rows for the table or file for a profile. You can preview the rows before or after you run the profile. You can preview the first 100 rows.

1. In the Navigator, select the project or folder that contains the profile that you preview rows for or create a profile.
2. Click the profile to open it.
   The profile appears in a tab.
3. Click the Data Preview view.
   The rows for the data source appear in the panel.
4. Click Refresh to refresh the data.

Viewing and Adding Profile Comments

You can view and add comments to any profile object. Use comments to share information about profiles with other Analyst tool users.

1. In the Navigator, select the project and folder that contains the profile.
2. Click the profile to open it.
   The profile appears in a tab.
3. Click Actions > Show Comments.
   The Comments panel appears. All comments appear in descending order by date.
4. To add a comment, enter the text for the comment in the comment box and click Add.
   The comment appears in the Comments panel.

Exporting Profile Results

You can export the results of a profile to a CSV file to view the data in a file. For each column in the profile, you can select the type of data to export. You can select value frequencies, pattern frequencies, or drilldown data for selected values and patterns.

1. In the Navigator, select the project or folder that contains the profile.
2. Click the profile to open it.
   The profile opens in a tab.
3. In the Column Profiling view, select the column that you want to export.
4. Click Actions > Export Data.
   The Export Data to a file window appears.
5. Enter the file name. Optionally, use the default file name.
6. Select the type of data to export.
   You can select one of the following data options to export:
   ♦ Value frequencies for the selected column.
   ♦ Pattern frequencies for the selected column.
   ♦ Drilldown data for the selected values or pattern.
7. Enter a file format. Default format is CSV.
8. Select the code page of the file.
9. Click OK.

Rules and Guidelines for Profiles

Rules and Guidelines for working with profiles.

Use the following rules and guidelines when you work with profiles:

♦ **Exporting Profile Results.** The Analyst tool displays only the first 200 values for value frequencies, pattern frequencies, and drilldown data for the selected value or column. When you export the profile results, the Analyst tool exports all values to the CSV file.

♦ **Inferred Datatypes.** The Analyst tool cannot derive the datatype from the values of a numeric column that has a precision greater than 38. The Analyst tool cannot derive the datatype from the values of a string column that has a precision greater than 255.

♦ **Column Patterns.** The Analyst tool cannot derive the pattern for a numeric column that has a precision greater than 38. The Analyst tool cannot derive the pattern for a string column that has a precision greater than 255.

♦ **Datatypes not included in profiling.** The Analyst tool excludes the CLOB, BLOB, Raw, and Binary datatypes in column values in a profile.

♦ **Column names.** You cannot add a column with the same name to an existing profile. You cannot add the same column twice to a profile even if you change the column name.
Scorecards Overview

A scorecard is the graphical representation of valid values for a column or the output of a rule in profile results. Use scorecards to measure data quality progress. You can create a scorecard from a profile and monitor the progress of data quality over time.

Scorecards display the value frequency for columns in a profile as scores. Scores reflect the percentage of valid values for a column.

For example, you can create a scorecard to measure data quality before you apply data quality rules. After you apply data quality rules, you can create another scorecard to compare the effect of the rules on data quality. You can monitor the progress of performing data quality on the data by setting up the scorecard to run daily. You can view the score trend chart to see how a particular score improves over time.

After you run a profile, you can add columns from the profile to a scorecard. You can add columns from multiple profiles to a scorecard. Select the valid values for the column and run the scorecard to see the scores for the columns in a chart.

When you view a scorecard in the Contents view, the Analyst tool opens the scorecard in another tab. After you run the scorecard, you can view the scores on the Scorecard view. You can select the data object and navigate to the data object from a score within a scorecard. The Analyst tool opens the data object in another tab.

You can complete the following tasks with scorecards:

- **Add columns to a scorecard.** Use the Add to Scorecard Wizard to add columns from a profile to a scorecard and configure the valid values for the columns. You can add columns to an existing scorecard or you can add columns to a new scorecard.
- **Run a scorecard.** Run a scorecard to generate scores for columns.
- **View a scorecard.** View a scorecard to see the scores for the columns.
- **Drill down on columns.** Drill down on the columns for a score to select columns that appear when you view the valid or invalid data rows.
- **Edit a scorecard.** Edit valid values for columns in a scorecard. You must run a scorecard before you can edit it.
Define thresholds. Define thresholds that specify the ranges of bad data for that are acceptable for each column in a record.

View trend charts. View trend charts to determine data quality progress over time.

Adding Columns to a Scorecard

Use the Add to Scorecard Wizard to add columns from a profile to a scorecard and configure the valid values for the columns. You can add columns to an existing scorecard or you can add columns to a new scorecard.

1. In the Navigator, select the project or folder that contains the profile.
2. Click the profile to open it.
   The profile appears in a tab.
3. Click Actions > Run Profile to run the profile.
4. Click Actions > Add to Scorecard.
   The Add to Scorecard Wizard appears.
5. Select the columns and rules that you want to add to a scorecard. Optionally, click the check box in the left column header to select all columns. Optionally, select Column Name to sort column names.
6. Optionally, select a score name to change the score name and add a description.
7. Click Next.
8. Select the scorecard that you want to add the columns to or click New, configure the name, description, and location of the scorecard in the New Scorecard window, and click OK.
9. Click Next.
10. Select each column in the Scores panel and configure the valid values from the list of all values in the Score using: Values panel.
11. Select each column in the Scores panel and select Set Custom Thresholds for this Score in the Score Settings panel to configure score thresholds.
    You can set thresholds for Good, Acceptable, and Unacceptable scores.
12. Click Finish.

Running a Scorecard

Run a scorecard to generate scores for columns.

1. In the Navigator, select the project or folder that contains the scorecard.
2. Click the scorecard to open it.
   The scorecard appears in a tab.
3. Click Actions > Run Scorecard.
   The Analyst tool runs the scorecard and generates scores.

Viewing a Scorecard

View a scorecard to see the scores for each column in a record. You can select a score to view the rows of valid or invalid data for each column. You must run the scorecard to view the scores.
1. Run a scorecard to view the scores.
2. Select a column that contains the score you want to view.
3. Click **Actions > Show Rows** to view the rows of valid or invalid data for the column.
   The Analyst tool displays the rows of valid data by default in the **Drilldown** panel. Optionally, click **Invalid** to view the rows of invalid data.

### Drilling Down on Columns

Drilldown on the columns for a score to select columns that appear when you view the valid or invalid data rows. The columns you select to drill down on appear in the **Drilldown** panel.

1. Run a scorecard to view the scores.
2. Select a column that contains the score you want to view.
3. Click **Actions > Show Rows** to view the rows of valid or invalid data for the column.
4. Click **Actions > Drilldown Columns**.
   The **Drilldown Columns** window appears.
5. Select the columns you want to include in the Drilldown panel.
   Optionally, select the **Name** column to include all columns.
6. Click **OK**.
   The columns appear in the **Drilldown** panel for the selected score. The Analyst tool displays the rows of valid data for the columns by default. Optionally, click **Invalid** to view the rows of invalid data.

### Editing a Scorecard

Edit valid values for columns in a scorecard. You must run a scorecard before you can edit it.

1. In the Navigator, select the project or folder that contains the scorecard.
2. Click the scorecard to open it.
   The scorecard appears in a tab.
3. Click **Actions > Edit**.
   The **Edit Scorecard** window appears.
4. Select each column in the **Scores** panel and configure the valid values from the list of all values in the **Score using: Values** panel.
5. Select **Set Custom Thresholds for this Score** in the **Score Settings** panel to configure score thresholds.
6. Click **Save** to save changes to the scorecard.

### Defining Thresholds

You can set thresholds for each score in a scorecard. A threshold specifies the range in percentage of bad data that is acceptable for columns in a record. You can set thresholds for good, acceptable, or unacceptable ranges of data. You can define thresholds for each column when you add columns to a scorecard, or when you edit a scorecard.
Complete the following prerequisite tasks before you define thresholds for columns in a scorecard:

- In the Navigator, select the project or folder that contains the profile and add columns from the profile to the scorecard in the Add to Scorecard window.
- Optionally, in the Navigator, select the project or folder that contains the scorecard and click the scorecard to edit it in the Edit Scorecard window.

1. In the Add to Scorecard window, or the Edit Scorecard window, select each column in the Score panel.
2. Select Set Custom Thresholds for this Score in the Score Settings column.
3. Enter the thresholds that represent the upper bound of the unacceptable range and the lower bound of the good range.
4. Click Finish or Save.

Viewing Trend Charts

You can view trend charts for each score to monitor how the score changes over time.

1. In the Navigator, select the project or folder that contains the scorecard.
2. Click the scorecard to open it.
   The scorecard appears in a tab.
3. In the Scorecard view, select a score.
4. Click Actions > Show Trend Chart.
   The Trend Chart Detail window appears. You can view score values that have changed over time. The Analyst tool uses historical scorecard run data for each date and the latest valid score values to calculate the score. The Analyst tool uses the latest threshold settings in the chart to depict the color of the score points.

Rules and Guidelines for Scorecards

Rules and guidelines for working with scorecards.

Use the following rules and guidelines when you work with scorecards:

- You cannot add a column with the same name to an existing scorecard.
- You cannot add the same column twice to a scorecard even if you change the column name.
Reference Tables Overview

A reference table contains reference data that you can use to standardize source data. Reference data can include valid and standard values.

Create reference tables to establish relationships between source data values and the valid and standard values. You can share reference data with a developer for use in Standardizer and Lookup transformations in the Developer tool.

For example, during a data quality project, you create a reference table that contains the list of valid values for an address column in source data. A developer can use the reference data in the Developer tool to create a Standardizer transformation in a mapplet or mapping and standardize on the valid values for the address.

When you create reference tables in the Analyst tool, a developer can view these tables in the Developer tool. A developer can open a reference table to view the contents of the reference table and use them in Lookup and Standardizer transformations. A developer can also launch the Analyst tool from the Developer tool to edit the reference table.

To create a reference table, you can create the table manually, create the table from a profile column, or import a reference table. You can also create a reference table from the column values and pattern values in a profile column.

After you create a reference table, you can edit the reference table to add column or rows and add or edit standard and valid values. You can also search and replace values in the reference table rows. You create and manage reference tables on the Reference Table view. The Analyst tool tracks editing activities in the audit trail log. You can view the audit trail events to see the changes made to a reference table on the Audit Trail view. You can view properties for the reference table in the Properties view.

Reference Table Properties

When you create reference tables manually or from profile columns, configure column properties for each column you include in the reference table. When you import a reference table from a flat file, configure the flat file properties for the delimited flat file.
You can configure the following column properties for each column in a reference table:

<table>
<thead>
<tr>
<th>Property</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valid</td>
<td>Appears when you create a reference table manually or import it as a flat file. Table record contains a valid value to use in a Lookup or Standardizer transformation in the Developer tool.</td>
</tr>
<tr>
<td>Name</td>
<td>Name of the column.</td>
</tr>
<tr>
<td>Data Type</td>
<td>Datatype for the column. You can choose one of the following datatypes: - bigint - date/time - decimal - integer - string The values you can configure for precision and scale depend on the datatype you choose.</td>
</tr>
<tr>
<td>Precision</td>
<td>Precision for the column. Precision is the maximum number of digits or the maximum number of characters that the column can accomodate.</td>
</tr>
<tr>
<td>Scale</td>
<td>Scale for the column. Scale is the maximum number of digits that a column can accommodate to the right of the decimal point. Applicable for decimal columns.</td>
</tr>
<tr>
<td>Description</td>
<td>Description for the column.</td>
</tr>
</tbody>
</table>

You can configure the following flat file properties when you import a reference table from a delimited flat file:

<table>
<thead>
<tr>
<th>Property</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delimiters</td>
<td>Character used to separate columns of data. Use the Other field to enter a different delimiter. Delimiters must be printable characters and must be different from the escape character and the quote character if selected. Default is comma.</td>
</tr>
<tr>
<td>Text Qualifier</td>
<td>Quote character that defines the boundaries of text strings. Choose No Quote, Single Quote, or Double Quotes. If you select a quote character, the wizard ignores delimiters within pairs of quotes. Default is Double Quotes.</td>
</tr>
<tr>
<td>Column Names</td>
<td>Use data in the first row for column names. Select this option if column names appear in the first row.</td>
</tr>
<tr>
<td>Values</td>
<td>Indicates the row number at which the wizard starts reading when it imports the file.</td>
</tr>
</tbody>
</table>
Create Reference Tables

Use the reference table editor, profile results, or a flat file to create reference tables. Create reference tables to share reference data with developers in the Developer tool.

Use the following methods to create a reference table:

- **Create a reference table manually.** Use the reference table editor to create a reference table, add columns, and configure attributes.
- **Create a reference table from profile columns.** Select a column in a profile and add it to a reference table or create a reference table to add the column. Select a column in a profile and select the column values to add to a reference table or create a reference table to add the column values. Select a column in the profile and select the pattern values to add to a reference table or create a reference table to add the pattern values.
- **Import a reference table.** Import a reference table from a delimited flat file.

Creating a Reference Table Manually

Use the New Reference Table Wizard and the reference table editor to create a reference table manually. You can use the reference table editor to define the structure, columns, and data for the table.

1. In the Navigator, select the project and folder where you want to create the reference table.
2. Click **Actions > New Reference Table.**
   The New Reference Table Wizard appears.
3. Select the option to **Use the reference table editor.**
4. Click **Next.**
5. Enter the table name and optional description and default value.
   The Analyst tool uses the default value for any table record that does not contain a value.
6. For each column you want to include in the reference table, click the **Add New Column** icon and configure the column properties for each column.
   **Note:** You can reorder the columns or delete columns.
7. Optionally, choose to create a description column for rows in the reference table. Configure the name and precision for the column.
8. Optionally, enter an audit note.
   The audit note appears in the audit trail log.
9. Click **Finish.**

Creating a Reference Table from Profile Columns

You can create a reference table from a profile column. You can add a profile column to an existing reference table. The New Reference Table Wizard adds the column to the reference table.

1. In the Navigator, select the project or folder that contains the profile with the column that you want to add to a reference table.
2. Click the profile name to open it in another tab.
3. In the **Column Profiling** view, select the column that you want to add to a reference table.
4. Click **Actions > Add to Reference Table.**
   The New Reference Table Wizard appears.
5. Select the option to **Create a new reference table.**
Optionally, select **Add to existing reference table**, and click **Next**. Navigate to the reference table in the project or folder, preview the reference table data and click **Next**. Select the column to add and click **Finish**.

6. Click **Next**.
7. The column name appears by default as the table name. Optionally enter another table name, a description, and default value.
   The Analyst tool uses the default value for any table record that does not contain a value.
8. Click **Next**.
9. In the **Column Attributes** panel, configure the column properties for the column.
10. Optionally, choose to create a description column for rows in the reference table.
    Enter the name and precision for the column.
11. Preview the column values in the **Preview** panel.
12. Click **Next**.
13. The column name appears as the table name by default. Optionally, enter another table name and a description.
14. In the **Save in** panel, select the location where you want to create the reference table.
    The **Reference Tables** panel lists the reference tables in the location you select.
15. Optionally, enter an audit note.
16. Click **Finish**.

**Creating a Reference Table from Column Values**

You can create a reference table from the column values in a profile column. Select a column in a profile and select the column values to add to a reference table or create a reference table to add the column values.

1. In the Navigator, select the project or folder that contains the profile with the column that you want to add to a reference table.
2. Click the profile name to open it in another tab.
3. In the **Column Profiling** view, select the column that you want to add to a reference table.
4. In the **Values** view, select the column values you want to add. Use the CONTROL or SHIFT keys to select multiple values.
5. Click **Actions > Add to Reference Table**.
   The **New Reference Table** Wizard appears.
6. Select the option to **Create a new reference table**.
   Optionally, select **Add to existing reference table**, and click **Next**. Navigate to the reference table in the project or folder, preview the reference table data and click **Next**. Select the column to add and click **Finish**.
7. Click **Next**.
8. The column name appears by default as the table name. Optionally enter another table name, a description, and default value.
   The Analyst tool uses the default value for any table record that does not contain a value.
9. Click **Next**.
10. In the **Column Attributes** panel, configure the column properties for the column.
11. Optionally, choose to create a description column for rows in the reference table.
    Enter the name and precision for the column.
12. Preview the column values in the **Preview** panel.
13. Click **Next**.
14. The column name appears as the table name by default. Optionally, enter another table name and a description.
15. In the **Save in** panel, select the location where you want to create the reference table.
   
   The Reference Tables: panel lists the reference tables in the location you select.
16. Optionally, enter an audit note.
17. Click **Finish**.

Creating a Reference Table from Column Patterns
You can create a reference table from the column patterns in a profile column. Select a column in the profile and select the pattern values to add to a reference table or create a reference table to add the pattern values.

1. In the Navigator, select the project or folder that contains the profile with the column that you want to add to a reference table.
2. Click the profile name to open it in another tab.
3. In the **Column Profiling** view, select the column that you want to add to a reference table.
4. In the **Patterns** view, select the column patterns you want to add. Use the CONTROL or SHIFT keys to select multiple values
5. Click **Actions > Add to Reference Table**.
   
   The New Reference Table Wizard appears.
6. Select the option to Create a new reference table.
   
   Optionally, select Add to existing reference table, and click **Next**. Navigate to the reference table in the project or folder, preview the reference table data and click **Next**. Select the column to add and click **Finish**.
7. Click **Next**.
8. The column name appears by default as the table name. Optionally enter another table name, a description, and default value.
   
   The Analyst tool uses the default value for any table record that does not contain a value.
9. Click **Next**.
10. In the **Column Attributes** panel, configure the column properties for the column.
11. Optionally, choose to create a description column for rows in the reference table.
   
   Enter the name and precision for the column.
12. Preview the column values in the **Preview** panel.
13. Click **Next**.
14. The column name appears as the table name by default. Optionally, enter another table name and a description.
15. In the **Save in** panel, select the location where you want to create the reference table.
   
   The Reference Tables: panel lists the reference tables in the location you select.
16. Optionally, enter an audit note.
17. Click **Finish**

Importing a Reference Table
Import a reference table from a delimited flat file.
1. In the Navigator, select the project or folder where you want to create the reference table.
2. Click Actions > New Reference Table. The New Reference Table Wizard appears.
3. Select the option to Import a flat file.
4. Click Next.
5. Click Browse to select the flat file.
6. Click Upload to upload the file to a directory in the Informatica Services installation directory that the Analyst tool can access.
7. Enter the table name, and optional description and default value. The Analyst tool uses the default value for any table record that does not contain a value.
8. Select a code page that matches the data in the flat file.
9. Preview the data in the Preview of file panel.
10. Click Next.
11. Configure the flat file properties.
12. In the Preview panel, click Show to update the preview.
13. Click Next.
14. On the Column Attributes panel, configure the column properties for each column.
15. Optionally, choose to create a description column for rows in the reference table. Enter the name and precision for the column.
16. Click Finish.

Reference Table Management

You can perform tasks to manage reference tables. You can find and replace column values, add or remove columns and rows, edit column values, and export a reference table to a file.

You can perform the following tasks to manage reference tables:

- **Manage columns.** Use the Edit column properties window to add, edit, or delete columns in a reference table.
- **Manage rows.** Use the Add Rows window to add rows and the Edit Row window to edit rows in a reference table. Use the Delete icon to delete rows in a reference table.
- **Find and replace values.** You can find and replace values in individual reference table columns. You can find a value in a column and replace it with another value. You can replace all values in columns with another value.
- **Export a reference table.** Export a reference table to a comma-separated values (CSV) file, dictionary file, or Excel file.

Managing Columns

Use the Edit column properties window to add, edit, or delete columns in a reference table.

1. In the Navigator, select the project or folder that contains the reference table that you want to edit.
2. Click the reference table name to open it in a tab. The Reference Table tab appears.
3. Click Actions > Edit Table or click the Edit Table icon. The Edit column properties window appears.
4. To add a column, click the Add New Column icon in the Column Attributes panel and edit the column properties. Or, to edit an existing column, click the property you want to edit.

You cannot edit the datatype, precision, and scale of the column. You can rename the column and change the column description.

5. To delete a column, click the column and click the Delete icon.

6. Optionally, you can enter an audit note on the Audit Note panel. The audit note appears in the audit log for any action you perform in the Edit column properties window.

7. Click OK.

Managing Rows

You can add, edit, or delete rows in a reference table.

1. In the Navigator, select the project or folder containing the reference table that you want to edit.
2. Click the reference table name to open it in a tab. The Reference Table tab appears.
3. To add a row, click Actions > Add Row or click the Add Row icon. In the Add Row window, enter the value for each column and enter an optional audit note. Click OK.
4. To edit rows, select the rows and click Actions > Edit or click the Edit icon. In the Edit Rows window, enter the value for each column, select the columns to apply the changes to, and enter an optional audit note. Optionally, click Previous to edit the previous row and click Next to edit the next row. Click Apply to apply the changes.

The new column values appear in the tab.

5. To delete rows, select the rows you want to delete and click Actions > Delete or click the Delete icon. In the Delete Rows window, enter an optional audit note and click OK.

Finding and Replacing Values

You can find and replace values in individual reference table columns.

1. In the Navigator, select the project or folder containing the reference table that you want to find and replace values in.
2. Click the reference table name to open it in a tab. The Reference Table tab appears.
3. Click Actions > Find and Replace or click the Find and Replace icon. The Find and Replace toolbar appears.
4. Enter the search criteria in the Find box. Select all columns or a column that you want to find in the list. Enter the value you want to replace with, and click one of the following buttons:

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Next/Previous</td>
<td>Scroll through the column values that match the search criteria.</td>
</tr>
<tr>
<td>Highlight All</td>
<td>Highlight all the column values that match the search criteria.</td>
</tr>
<tr>
<td>Replace</td>
<td>Replace the currently highlighted column value.</td>
</tr>
<tr>
<td>Replace All</td>
<td>Replace all occurrences of the search criteria in column values.</td>
</tr>
</tbody>
</table>
Exporting a Reference Table

Export a reference table to a comma-separated values (CSV) file, dictionary file, or Microsoft Excel file.

1. In the Navigator, select the project or folder containing the reference table that you want to view the audit trail for.
2. Click the reference table name to open it in a tab. The Reference Table tab appears.
3. Click Actions > Export Data.
   The Export data to a file window appears.
4. Configure the following options:

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>File Name</td>
<td>File name for the exported data.</td>
</tr>
<tr>
<td>File Format</td>
<td>Format of the exported file. You can select the following formats:</td>
</tr>
<tr>
<td></td>
<td>✦ csv. Comma-separated values file.</td>
</tr>
<tr>
<td></td>
<td>✦ xls. Microsoft Excel file.</td>
</tr>
<tr>
<td></td>
<td>Optionally, select Export field names as first row to export the column names as a header row in the exported file.</td>
</tr>
<tr>
<td>Code Page</td>
<td>Code page of the reference data.</td>
</tr>
</tbody>
</table>

5. Click OK.
   The options to save or open the file depend on your browser.

Audit Trail Events

Use the Audit Trail view for a reference table to view audit trail log events.

The Analyst tool creates audit trail log events when you make a change to a reference table and enter an audit trail note. Audit trail log events provide information about the reference tables that you manage.

You can configure query options on the Audit Trail tab to filter the log events that you view. You can specify filters on the date range, type, user name, and status. The following table describes the options you configure when you view audit trail log events:

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Date</td>
<td>Start and end dates for the log events to search for. Use the calender to choose dates.</td>
</tr>
<tr>
<td>Type</td>
<td>Type of audit trail events. You can filter and view the following events types:</td>
</tr>
<tr>
<td></td>
<td>- Data. Events related to data in the reference table. Events include creating, editing, deleting, and replacing all rows.</td>
</tr>
<tr>
<td></td>
<td>- Metadata. Events related to reference table metadata. Events include creating reference tables, adding, deleting, and editing columns, and updating valid columns.</td>
</tr>
<tr>
<td>Option</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>-------------</td>
</tr>
<tr>
<td>User</td>
<td>User who edited the reference table and entered the audit trail comment. The Analyst tool generates the list of users from the Analyst tool users configured in the Administrator tool.</td>
</tr>
<tr>
<td>Status</td>
<td>Status of the audit trail log events. Status corresponds to the action performed in the reference table editor.</td>
</tr>
</tbody>
</table>

Audit trail log events also include the audit trail comments and the column values that were inserted, updated, or deleted.

**Viewing Audit Trail Events**

View audit trail log events to get more information about changes made to a reference table.

1. In the Navigator, select the project or folder that contains the reference table that you want to view the audit trail for.
2. Click the reference table name to open it in a tab. The Reference Table tab appears.
3. Click the Audit Trail view.
4. Configure the filter options.
5. Click Show.
   The log events for the specified query options appear.
This chapter includes the following topics:

- Data Quality Assistant Overview, 51

Data Quality Assistant Overview

Data Quality Assistant is a web-based tool in the Analyst tool that you can use to manage database records. Use Data Quality Assistant to fix bad records and consolidate duplicate records.

You can analyze bad records and duplicate records by creating tables for the records in the Analyst tool. Use the Data Quality Assistant to create bad record and duplicate record tables. When you create a bad record or duplicate record table, the Data Quality Assistant imports these tables into the Analyst tool. The tables that the Data Quality Assistant imports originate in the Informatica Data Quality product. Informatica Data Quality runs data quality plans on these tables and converts them into a format that the Data Quality Assistant can use.

Use Data Quality Assistant to complete the following tasks:

- Manage bad records. Filter records, edit records, and set record status.
- Consolidate records. Merge groups of duplicate records from a consolidation record table into a single record.
- View the audit trail. View the history of changes made to bad record and consolidation record tables.

A record is a row in a database table. To fix bad records or consolidate records in a database table, use the Data Quality Assistant to import a bad record or duplicate record table in the Analyst tool. When you open the table from a project or folder, the Analyst tool opens it in Data Quality Assistant. You can then fix the bad records or consolidate duplicate records.

Bad Records

Use Data Quality Assistant to filter and edit bad records and to flag bad records to reprocess. Bad records may contain invalid values or formats.

When you open a bad record table, Data Quality Assistant marks the bad fields in bold.

To manage bad records, use Data Quality Assistant to complete the following tasks:

- Filter records. Filter records by column name or by a filter criteria.
- Edit records. Select a record and edit the value for the record. You can select and edit multiple records.
- Flag records. Flag a record and assign a status to the record. You can select and flag multiple records.
The following table describes the status with which you can flag a record:

<table>
<thead>
<tr>
<th>Status</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accept</td>
<td>Flag a record as correct or acceptable for use.</td>
</tr>
<tr>
<td>Reject</td>
<td>Flag a record as not acceptable for use.</td>
</tr>
<tr>
<td>Reprocess</td>
<td>Flag a record for reprocessing. Select this option if a record may not be valid.</td>
</tr>
</tbody>
</table>

To manage bad records, use the following process:

1. **Import a bad record table.** In the Analyst tool, import the database table that contains the records that you want to analyze.
2. **Open the bad record table.** When you open the table, the Analyst tool opens the table in Data Quality Assistant.
3. **Filter the records.** Select the **Quality Issue** and the column to list the rows that have the specified quality issue in the selected column. Optionally, click **Actions > Filter** to specify the query to filter the listed records.
4. **Edit records.** Select bad records and edit the values.
5. **Flag records.** For each bad record identified by Data Quality Assistant, flag the record as Accept, Reject, or Reprocess from the **Actions** menu.
6. **View the audit trail.** View the audit trail events for actions performed on the bad record table.

**Consolidation Records**

Use Data Quality Assistant to merge duplicate records into a single record.

When you open a consolidation record table, Data Quality Assistant groups similar records into a cluster. A cluster contains a master record and potential duplicate records. Use the master record to identify the final values for records and merge the master record with other duplicate records in a cluster.

To manage consolidation records, use the following process:

1. **Import consolidation record table.** In the Analyst tool, import the database table that contains the records that you want to consolidate.
2. **Open the consolidation record table.** When you open the table, the Analyst tool opens the table in the Data Quality Assistant.
3. **Edit the clusters.** Filter the clusters, extract records that are not part of the final record, and select a master record.
4. **Consolidate the records.** Consolidate the records in a cluster into a single, final record.
5. **Edit the final record.** Optionally, you can edit the values for the final record.
6. **View the audit trail.** View the audit trail events for actions performed on the consolidation record table.

**Audit Trail**

Use the audit trail to view the history of changes made to bad record and consolidation record tables.

Data Quality Assistant tracks changes made to records during bad record and consolidation management in the audit trail log. View the audit trail on the Audit Trail tab for the bad record or consolidation record table.
The Audit Trail tab for a bad record table or consolidation record table shows the following types of log events:

- Updated, accepted, rejected, and reprocessed records for a bad record table.
- Consolidated and extracted rows for a consolidation table.

You can configure query options on the Audit Trail tab to filter the log events that you view. You can specify filters on the date range, user name, and status. Audit trail log events also include the audit trail comment and the record values that were modified.

The following table describes the options you configure when you view audit trail log events:

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Date</td>
<td>Start and end dates for the log events to search for. Use the calendar to choose dates.</td>
</tr>
<tr>
<td>User</td>
<td>User who edited the table and entered the audit trail comment. The Analyst tool generates the list of users from the Analyst tool users configured in the Administrator tool.</td>
</tr>
<tr>
<td>Status</td>
<td>Status of the audit trail log events. Status corresponds to the action performed in Data Quality Assistant.</td>
</tr>
</tbody>
</table>
### Metadata Bookmarks Overview

A metadata bookmark is a link to an object in a Model repository. Use a metadata bookmark to share an object with other Analyst tool users.

Because the Analyst tool is a web-based tool, you can access objects through a link to the object in the Analyst tool. Each object you view in the Analyst tool has a unique URL. You can share an object with other Analyst tool users by sharing the URL for the object. You can create a metadata bookmark for any object that you can open in the Analyst tool.

The following example shows a metadata bookmark for a relational table:

```
http://styx:8080/AnalystTool/com.informatica.at.AnalystTool/
index.jsp?p=lewis$i=U:VeF2HpstEd66x8vKMFuKtQ4c=com.informatica.metadata.relational.datasource.Repository
```

To share a metadata bookmark, open the object you want to share in the Analyst tool. Copy the link location in the location bar. You can then send the link in an email or add it to a document. You can also bookmark the link in your browser to access it again.

To access a metadata bookmark, you can click the link or copy and paste the link into the location bar in a browser. When you access a metadata bookmark, the Analyst tool prompts you to log in if you are not already logged in, and then displays the object.

To access a metadata bookmark, you must have the following permissions:

- Permission to use the the Analyst tool.
- Permission to access the project that contains the object.

### Creating a Metadata Bookmark

Create a metadata bookmark to share an object in Informatica Analyst with other users.

1. In the Analyst tool, open the object that you want to create a metadata bookmark for.
2. Copy the URL in the location bar of the browser.

   You can then paste the bookmark into an email or a document and distribute the email or document to other users.
Opening a Metadata Bookmark

Open a metadata bookmark to access an object in the Analyst tool.

1. Click an active link for a metadata bookmark or copy and paste the link into the location bar in a browser.
   If you are not already logged in, the Analyst tool displays the login page.
2. Log in to the Analyst tool.
   The Analyst tool displays the object in a tab.
INDEX

A
adding a flat file  
flat files 20
adding a table  
tables 23
adding columns to a scorecard  
scorecards 39
applying a predefined rule  
rules 32
assign permissions  
projects 5
assigning permissions  
projects 15

C
column patterns  
profile results 35
column properties  
reference tables 43
column statistics  
profile results 35
column values  
profile results 34
configuring  
profiles 26
rules 26
contents view  
Informatica Analyst interface 3
creating a custom profile  
profiles 28
creating a folder  
folders 7
creating a metadata bookmark  
metadata bookmark 54
creating a profile in a project  
profiles 27
creating a project  
projects 6
creating a quick profile  
profiles 28
creating a reference table from column patterns  
reference tables 46
creating a reference table from column values  
reference tables 45
creating a reference table from profile columns  
reference tables 44
creating a reference table manually  
reference tables 44
creating an expression rule  
rules 30
creating an expression rule from profile results  
rules 31

D
data objects  
deleting an object 17
flat files 17
renaming an object 17
tables 17
viewing an object 17
datetime datatypes  
flat files 19
defining thresholds  
scorecards 41
deleting a folder  
folders 9
deleting a Project  
projects 7
deleting an object  
objects 11
drilldown columns  
scorecards 40
duplicating a folder  
folders 8
duplicating a project  
projects 6
duplicating an object  
objects 10

E
editing a scorecard  
scorecards 40
exporting a reference table  
reference tables 49
exporting profile results  
profiles 36

F
finding and replacing values  
reference tables 48
flat file datatypes  
flat files 19
flat file options  
flat files 18
flat file properties  
reference tables 43
flat files  
adding a flat file 20
data objects 17
datetime datatypes 19
flat file datatypes 19
flat file options 18
viewing data objects 23
folders
creating a folder 7
deleting a folder 9
duplicating a folder 8
moving a folder 8
renaming a folder 8
viewing folders contents 9

I
importing a reference table
reference tables 47
Informatica Analyst
Navigator 2
Informatica Analyst interface
contents view 3
Informatica Analyst 2
Informatica Analyst views 3
log in 4
properties view 3
security view 4
Informatica Analyst views
Informatica Analyst interface 3

L
log in
Informatica Analyst interface 4

M
manage folders
projects 5
manage projects
projects 5
managing columns
reference tables 47
managing rows
reference tables 48
metadata bookmark
creating a metadata bookmark 54
opening a metadata bookmark 55
moving a folder
folders 8
moving an object
objects 11

N
Navigator
Informatica Analyst 2

O
objects
deleting an object 11
duplicating an object 10
moving an object 11
previewing data 10
renaming an object 10
viewing object properties 10
opening a metadata bookmark
metadata bookmark 55

P
performing a search
search 14
previewing a table or file
profile results 36
previewing data
objects 10
profile comments
profiles 36
profile results
column patterns 35
column statistics 35
column values 34
previewing a table or file 36
profile summary 34
viewing profile results 35
profile summary
profile results 34
profiles
configuring 26
creating a custom profile 28
creating a profile in a project 27
creating a quick profile 28
exporting profile results 36
profile comments 36
running a profile 29
Profiles
profile results 33
projects
assign permissions 5
assigning permissions 15
creating a Project 6
deleting a project 7
duplicating a project 6
manage folders 5
manage projects 5
renaming a project 7
search projects 5
viewing project contents 9
properties view
Informatica Analyst interface 3

R
reference tables
column properties 43
creating a reference table from column patterns 46
creating a reference table from column values 45
creating a reference table from profile columns 44
creating a reference table manually 44
exporting a reference table 49
finding and replacing values 48
flat file properties 43
importing a reference table 47
managing columns 47
managing rows 48
viewing audit trail tables 50
renaming a Folder
folders 8
renaming a Project
projects 7
renaming an object
objects 10
rules
applying a predefined rule 32
configuring 26
creating an expression rule 30
creating an expression rule from profile results 31
running a profile
profiles 29
running a scorecard
scorecards 39

S
scorecards
  adding columns to a scorecard 39
  defining thresholds 41
  drilldown columns 40
  editing a scorecard 40
  running a scorecard 39
  viewing a scorecard 40
  viewing trend charts 41
search
  performing a search 14
  search syntax 13
  searching objects example 15
search projects
  projects 5
search syntax
  search 13
  searching objects example
  search 15
security view

Informatica Analyst interface 4

T
tables
  adding a table 23
  data objects 17
  viewing data objects 23

V
viewing a scorecard
  scorecards 40
viewing audit table events
  reference tables 50
viewing data objects
  flat files 23
  tables 23
viewing folder contents
  folders 9
viewing object properties
  objects 10
viewing profile results
  profile results 35
viewing project contents
  projects 9
viewing trend charts
  scorecards 41