Informatica

Informatica® Test Data Management
10.5.0

Getting Started Guide

Informatica Test Data Management Getting Started Guide
10.5.0
March 2021

© Copyright Informatica LLC 2003, 2021

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

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

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 © Aandacht c.v. All rights reserved. Copyright Genivia, Inc. All rights reserved. Copyright 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 © Rogue 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. Copyright © Thinkmap, Inc. All rights reserved. Copyright © Clearpace Software Limited. All rights reserved. Copyright © Information Builders, Inc. All rights reserved. Copyright © OSS Nokalva, Inc. All rights reserved. Copyright Edifecs, Inc. All rights reserved. Copyright Cleo Communications, Inc. All rights reserved. Copyright © International Organization for Standardization 1986. All rights reserved. Copyright © ej-technologies GmbH. All rights reserved. Copyright © Jaspersoft Corporation. All rights reserved. Copyright © International Business Machines Corporation. All rights reserved. Copyright © yWorks GmbH. All rights reserved. Copyright © Lucent Technologies. All rights reserved. Copyright © University of Toronto. All rights reserved. Copyright © Daniel Veillard. All rights reserved. Copyright © Unicode, Inc. Copyright IBM Corp. All rights reserved. Copyright © MicroQuill Software Publishing, Inc. All rights reserved. Copyright © PassMark Software Pty Ltd. All rights reserved. Copyright © LogiXML, Inc. All rights reserved. Copyright © 2003-2010 Lorenzi Davide, All rights reserved. Copyright © Red Hat, Inc. All rights reserved. Copyright © The Board of Trustees of the Leland Stanford Junior University. All rights reserved. Copyright © EMC Corporation. All rights reserved. Copyright © Flexera Software. All rights reserved. Copyright © Jinfonet Software. All rights reserved. Copyright © Apple Inc. All rights reserved. Copyright © Telerik Inc. All rights reserved. Copyright © BEA Systems. All rights reserved. Copyright © PDFlib GmbH. All rights reserved. Copyright © Orientation in Objects GmbH. All rights reserved. Copyright © Tanuki Software, Ltd. All rights reserved. Copyright © Ricebridge. All rights reserved. Copyright © Sencha, Inc. All rights reserved. Copyright © Scalable Systems, Inc. All rights reserved. Copyright © jQWidgets. All rights reserved. Copyright © Tableau Software, Inc. All rights reserved. Copyright© MaxMind, Inc. All Rights Reserved. Copyright © TMate Software s.r.o. All rights reserved. Copyright © MapR Technologies Inc. All rights reserved. Copyright © Amazon Corporate LLC. All rights reserved. Copyright © Highsoft. All rights reserved. Copyright © Python Software Foundation. All rights reserved. Copyright © BeOpen.com. All rights reserved. Copyright © CNRI. All rights reserved.

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

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 various versions of 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 and http://www.openssl.org/source/license.html.

This product includes Curl software which is Copyright 1996-2013, 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://dojotoolkit.org/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 and at http://www.eclipse.org/org/documents/edl-v10.php.

This product includes software licensed under the terms at http://www.tcl.tk/software/tcltk/license.html, http://www.bosrup.com/web/overlib/?License, http:// www.stlport.org/doc/ license.html, http://asm.ow2.org/license.html, http://www.cryptix.org/LICENSE.TXT, http://hsqldb.org/web/hsqlLicense.html, http:// httpunit.sourceforge.net/doc/ license.html, http://jung.sourceforge.net/license.txt , http://www.gzip.org/zlib/zlib_license.html, http://www.openldap.org/software/ release/license.html, http://www.libssh2.org, http://slf4j.org/license.html, http://www.sente.ch/software/OpenSourceLicense.html, http://fusesource.com/downloads/ license-agreements/fuse-message-broker-v-5-3- license-agreement; http://antlr.org/license.html; http://aopalliance.sourceforge.net/; http://www.bouncycastle.org/ licence.html; http://www.jgraph.com/jgraphdownload.html; http://www.jcraft.com/jsch/LICENSE.txt; http://jotm.objectweb.org/bsd_license.html; . http://www.w3.org/ Consortium/Legal/2002/copyright-software-20021231; http://www.slf4j.org/license.html; http://nanoxml.sourceforge.net/orig/copyright.html; http://www.json.org/ license.html; http://forge.ow2.org/projects/javaservice/, http://www.postgresql.org/about/licence.html, http://www.sqlite.org/copyright.html, http://www.tcl.tk/ software/tcltk/license.html, http://www.jaxen.org/faq.html, http://www.jdom.org/docs/faq.html, http://www.slf4j.org/license.html; http://www.iodbc.org/dataspace/ iodbc/wiki/iODBC/License; http://www.keplerproject.org/md5/license.html; http://www.toedter.com/en/jcalendar/license.html; http://www.edankert.com/bounce/ index.html; http://www.net-snmp.org/about/license.html; http://www.openmdx.org/#FAQ; http://www.php.net/license/3_01.txt; http://srp.stanford.edu/license.txt;

http://www.schneier.com/blowfish.html; http://www.jmock.org/license.html; http://xsom.java.net; http://benalman.com/about/license/; https://github.com/CreateJS/EaselJS/blob/master/src/easeljs/display/Bitmap.js; http://www.h2database.com/html/license.html#summary; http://jsoncpp.sourceforge.net/LICENSE; http://jdbc.postgresql.org/license.html; http://protobuf.googlecode.com/svn/trunk/src/google/protobuf/descriptor.proto; https://github.com/rantav/hector/blob/master/LICENSE; http://web.mit.edu/Kerberos/krb5-current/doc/mitK5license.html; http://jibx.sourceforge.net/jibx-license.html; https://github.com/lyokato/libgeohash/blob/master/LICENSE; https://github.com/hjiang/jsonxx/blob/master/LICENSE; https://code.google.com/p/lz4/; https://github.com/jedisct1/libsodium/blob/master/LICENSE; http://one-jar.sourceforge.net/index.php?page=documents&file=license; https://github.com/EsotericSoftware/kryo/blob/master/license.txt; http://www.scala-lang.org/license.html; https://github.com/tinkerpop/blueprints/blob/master/LICENSE.txt; http://gee.cs.oswego.edu/dl/classes/EDU/oswego/cs/dl/util/concurrent/intro.html; https://aws.amazon.com/asl/; https://github.com/twbs/bootstrap/blob/master/LICENSE; https://sourceforge.net/p/xmlunit/code/HEAD/tree/trunk/LICENSE.txt; https://github.com/documentcloud/underscore-contrib/blob/master/LICENSE, and https://github.com/apache/hbase/blob/master/LICENSE.txt.

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), the Sun Binary Code License Agreement Supplemental License Terms, the BSD License (http:// www.opensource.org/licenses/bsd-license.php), the new BSD License (http://opensource.org/licenses/BSD-3-Clause), the MIT License (http://www.opensource.org/licenses/mit-license.php), the Artistic License (http://www.opensource.org/licenses/artistic-license-1.0) and the Initial Developer's Public License Version 1.0 (http://www.firebirdsql.org/en/initial-developer-s-public-license-version-1-0/).

This product includes software copyright © 2003-2006 Joe Walnes, 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 product includes software Copyright (c) 2013 Frank Balluffi and Markus Moeller. All rights reserved. Permissions and limitations regarding this software are subject to terms of the MIT license.

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

DISCLAIMER: Informatica LLC provides this documentation "as is" without warranty of any kind, either express or implied, including, but not limited to, the implied warranties of noninfringement, merchantability, or use for a particular purpose. Informatica LLC 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.

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

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

Publication Date: 2021-03-17

Table of Contents

Preface 6

Informatica Resources 6

- Informatica Network 6
- Informatica Knowledge Base 6
- Informatica Documentation 6
- Informatica Product Availability Matrices 7
- Informatica Velocity 7
- Informatica Marketplace 7
- Informatica Global Customer Support 7

Chapter 1: Introduction to Test Data Management 8

Test Data Management Overview 8

- TDM Example 9

TDM Architecture 9

Test Data Manager User Interface 11

Informatica Administrator 12

TDM Process 13

Chapter 2: Downloading and Running Scripts 14

Downloading and Running Scripts Overview 14

Download and Run the Script 14

Chapter 3: Creating Users and Groups in Informatica Administrator 15

Creating Users and Groups in Informatica Administrator Overview 15

Step 1. Log In to Informatica Administrator 16

Step 2. Create a User 16

Step 3. Create a Group 18

Step 4. Assign Users to Groups 19

Step 5. Assign Roles and Privileges to Users and Groups 20

Chapter 4: Setting Up Test Data Manager 22

Setting Up Test Data Manager Overview 22

Step 1. Log In to Test Data Manager 23

Step 2. Create Source and Target Connections 23

Step 3. Create a Project 24

Step 4. Import Data Sources 25

Chapter 5: Creating Logical Relationships in TDM 27

Creating Logical Relationships in TDM Overview 27

Step 1. Create a Logical Relationship Between Tables 28

Chapter 6: Defining Data Subset Components. 31

Defining Data Subset Components Overview. 31

Step 1. Create an Entity. 32

Chapter 7: Creating a Data Masking Rule. 36

Creating a Data Masking Rule Overview. 36

Step 1. Create a Standard Data Masking Rule. 37

Step 2. Add the Standard Masking Rule to the Project. 38

Step 3. Assign the Masking Rule. 39

Chapter 8: Creating a Plan. 40

Creating a Plan Overview. 40

Step 1. Create a Plan. 40

Chapter 9: Managing the Workflow. 43

Managing the Workflow Overview. 43

Step 1. Generate the Workflow. 44

Step 2. Run the Workflow. 44

Step 3. Monitor the Workflow. 45

Appendix A: Glossary. 47

Preface

Follow the instructions in the Informatica *Test Data Management Getting Started Guide* to learn how to perform data subset and data masking operations. The tutorial shows you how to create TDM users and groups in the Administrator tool. It also provides information on how to create TDM connections, create a project and import data sources, create logical relationships, create an entity, create a data masking rule, create a plan, and run the workflow.

Informatica Resources

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

Informatica Network

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

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

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

Informatica Knowledge Base

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

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

Informatica Documentation

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

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

Informatica Product Availability Matrices

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

Informatica Velocity

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

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

Informatica Marketplace

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

Informatica Global Customer Support

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

To find your local Informatica Global Customer Support telephone number, visit the Informatica website at the following link:
https://www.informatica.com/services-and-training/customer-success-services/contact-us.html.

To find online support resources on the Informatica Network, visit https://network.informatica.com and select the eSupport option.

CHAPTER 1

Introduction to Test Data Management

This chapter includes the following topics:

- Test Data Management Overview, 8
- TDM Architecture, 9
- Test Data Manager User Interface, 11
- Informatica Administrator, 12
- TDM Process, 13

Test Data Management Overview

Test Data Management (TDM) integrates with PowerCenter®, PowerExchange®, and Informatica applications to manage nonproduction data in an organization.

Organizations create multiple copies of application data to use for testing and development. Organizations often maintain strict controls on production systems, but data security in nonproduction systems is not as secure. An organization must maintain knowledge of the sensitive data in production systems and ensure that sensitive data does not appear in the test environment. Development must not have to rewrite code to create test data.

With TDM, an organization can create a smaller copy of the production data and mask the sensitive data. An organization can discover the sensitive columns in the test data and ensure that the sensitive columns are masked in the test data. An organization can also create test data that does not contain sensitive data from the production database.

Manage data subset and data masking in Test Data Manager. Use data subset to create a small environment for testing and development. You can define the type of data that you want to include in the subset database. You might create a subset database with data based on time, function, or geographic location. Create data masking rules to apply to source columns and data domains. You can assign multiple rules to the same column.

To perform data subset and masking operations, you can generate and run workflows from data subset and data masking plans in Test Data Manager.

TDM users have roles and privileges that determine the tasks they can perform through Test Data Manager. The administrator manages roles and privileges for users from the Informatica Administrator.

TDM Example

With TDM, an organization can create a smaller copy of the production data and mask the sensitive data.

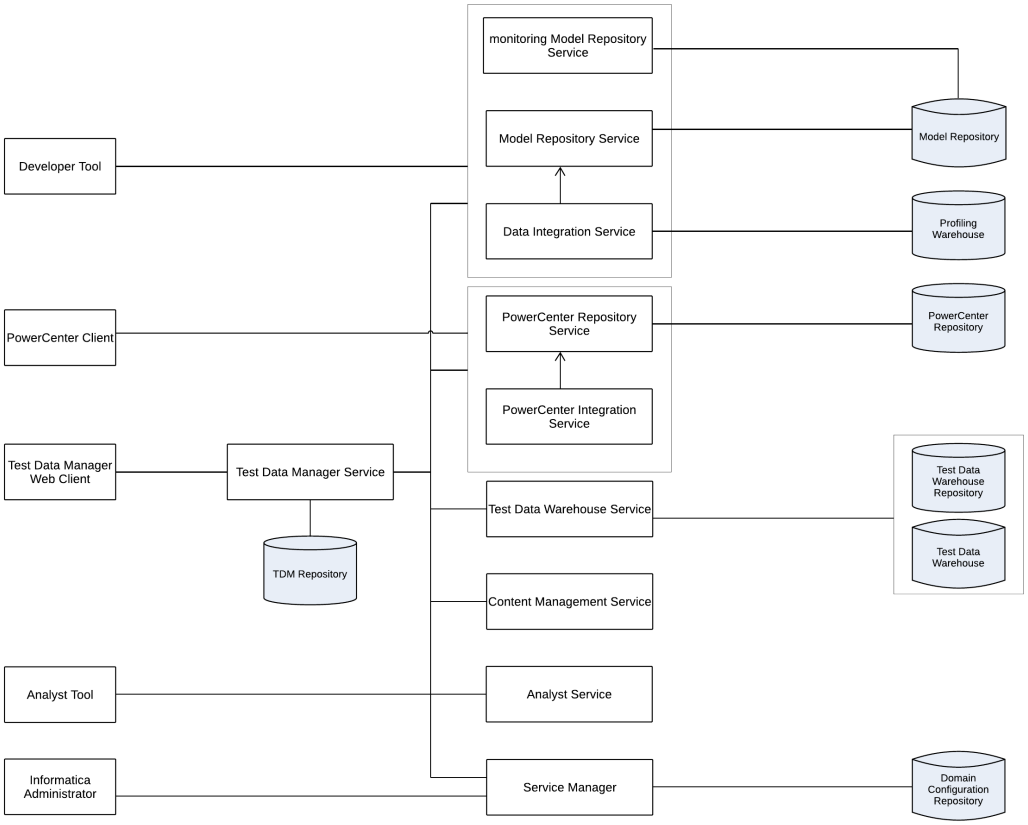
An organization wants to create a subset of data that contains details of all the customers based on the identification numbers. The organization wants to mask the credit card numbers of the customers to ensure data security.

To perform the tasks, the organization uses Test Data Manager to create TDM connections, create a project and import data sources, create an entity, create a masking rule, configure a plan, and run the workflow.

TDM Architecture

The TDM architecture consists of tools, the TDM server, the Test Data Manager Service and other application services, and databases.

The following image shows the components of TDM:



The TDM architecture consists of the tools that you use to modify the data.

The TDM architecture includes the following components.

Test Data Manager

A web-based client application that you can use to configure data masking, data subset, data generation, and profiles for data discovery. You can also configure connections, and manage project permissions for users and user groups.

Informatica Developer

A client application that you use to create and export profiles for data discovery.

Informatica Administrator

A web-based client that a domain administrator uses to manage application services and create users and user groups.

PowerCenter Client

A client application that you use to configure permissions on folders and on connection objects for the PowerCenter repository.

TDM Server

The TDM server is the interface between Test Data Manager and the application services.

Data Integration Service

An application service that performs the data discovery operations. The Data Integration Service connects to the Model Repository Service to store metadata from data discovery profiles in the Model repository. When you create a Data Integration Service in the Administrator tool, you select the Data Profiling warehouse to store data from data discovery profiles.

Model Repository Service

An application service that manages the Model repository for data discovery operations.

PowerCenter Integration Service

An application service that runs data subset, data generation, and data masking workflows. When you create the Test Data Manager Service in the Informatica Administrator, you select the PowerCenter Integration Service that runs the workflows.

PowerCenter Repository Service

An application service that manages the PowerCenter repository. The PowerCenter Repository Service accepts requests from the PowerCenter Integration Service when a workflow runs.

Test Data Manager Service

An application service that creates and manages the TDM repository. The Test Data Manager accesses the Test Data Manager Service to use database content from the TDM repository.

TDM repository

A relational database that contains tables that TDM requires to run and the tables that store metadata.

Model repository

A relational database that stores table metadata for data discovery profiles and the connections that you create in Test Data Manager.

PowerCenter repository

A relational database that stores metadata for PowerCenter sources and targets. The PowerCenter repository also stores metadata for the workflows that you generate from plans in Test Data Manager.

Profiling warehouse

A relational database that stores profile results for data discovery.

Domain configuration repository

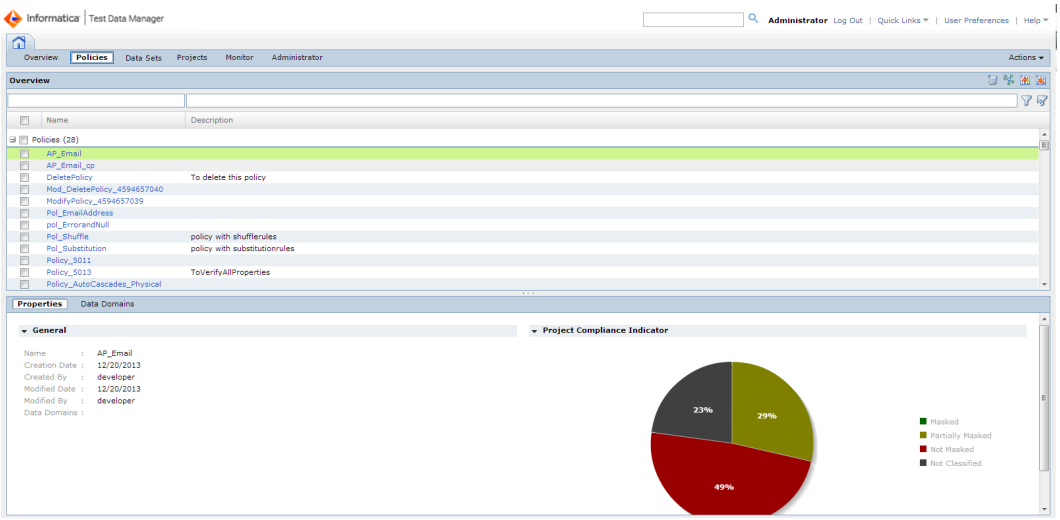
A relational database that stores connections and metadata for the Informatica domain.

Test Data Manager User Interface

Test Data Manager is a web-based user interface that you use to perform data discovery, data subset, data masking, and data generation operations.

Test Data Manager contains options to view and edit TDM components. Open a view in Test Data Manager based on the task that you need to perform.

The following image shows a view in Test Data Manager:



The Contents panel shows an overview of the items in a view. The Details panel shows additional details for a single item in the contents panel.

Test Data Manager contains the following views:

Overview

View dashboard reports about projects in the TDM repository.

Policies

Define policies, masking rules, and generation rules that you can add to projects.

Data Sets

Manage data sets that you store in the test data warehouse.

Projects

Define a project that contains source data and the data subset, data masking, data profiling, or data generation operations for the data.

Monitor

View the status of jobs that import sources or perform data subset, data masking, data profiling, or data generation operations. Stop or abort jobs.

Administrator

Manage connections, dictionaries, and workflow options.

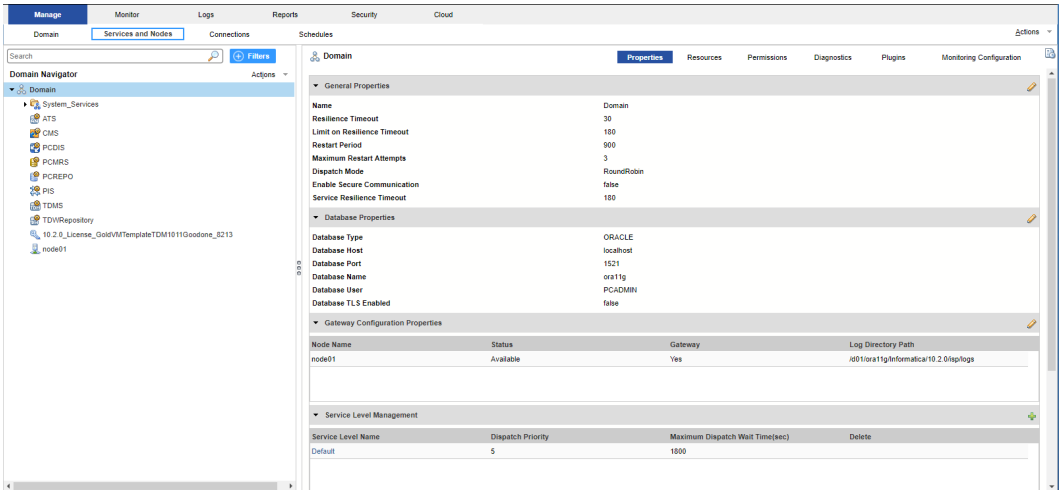
Note: By default, an administrator can access the Administrator view of Test Data Manager. A user must have privileges to access the other views in Test Data Manager.

Informatica Administrator

Informatica Administrator is an application client that consolidates the administrative tasks for domain objects such as services, connections, and licenses.

Use the Administrator tool to manage the domain and the security of the domain.

The following image shows the Administrator tool:



The Administrator tool has the following tabs:

Domain

View and edit the properties of the domain and objects within the domain.

Logs

View log events for the domain and services within the domain.

Monitoring

View the status of profile jobs, scorecard jobs, preview jobs, mapping jobs, SQL data services, web services, and workflows for each Data Integration Service.

Reports

Run a Web Services Report or License Management Report.

Security

Manage users, groups, roles, and privileges. If you have PowerCenter Express Personal Edition, you do not have access to the Security tab.

Cloud

View the details of the organizations, Secure Agents, and connections in the Cloud tab. You must have sufficient privileges to view the Cloud tab.

TDM Process

Create source and target connections, import data sources from the source database into a project, create a subset of the data, and mask the subset data.

Use Test Data Manager to create connections, import metadata, establish foreign key relationships, create entities, create masking rules, create plans, run workflows, and monitor the workflow progress.

Step 1. Create TDM connections

TDM connects to databases, repositories, and services to perform data subset, masking, generation, and profiles for discovery operations. To perform data subset and masking operations, you need to create TDM connections in Test Data Manager.

Step 2. Create a project and import data sources

Create a project to store the TDM components. Import data sources from the database connections that you create in Test Data Manager.

Step 3. Create a logical relationship between tables

Relationships that you create in TDM are logical relationships. Identify relationships that you want to add to the TDM repository and then create the logical relationships as constraints in Test Data Manager.

Step 4. Create an entity

An entity consists of a driving table and related tables. When you create an entity, you select the driving table. Test Data Manager retrieves the related tables based on the constraints.

Step 5. Create and assign a data masking rule

Data masking rules define how to mask sensitive and confidential data in a target database. When you create a data masking rule, you define the logic to replace sensitive data. Assign the data masking rule to a column to mask the sensitive data.

Step 6. Create a plan

A plan includes the components that you need to generate a workflow. You create a plan, add data subset and data masking components, and configure the plan properties.

Step 7. Manage the workflow

Generate and run the workflow to populate the subset of data in the target database. TDM masks the sensitive column based on the masking rule. View the progress of the workflow job from the Monitor view.

CHAPTER 2

Downloading and Running Scripts

This chapter includes the following topics:

- Downloading and Running Scripts Overview, 14
- Download and Run the Script, 14

Downloading and Running Scripts Overview

You must download and run the SQL queries to populate the sample tables in the database.

You need an Oracle database to run the SQL script and to perform the tasks in this tutorial. You must database administrator privileges to run the script.

Download the sample data script from the following location:

https://kb.informatica.com/proddocs/Product%20Documentation/3/TDM_GettingStarted_SampleSourceScript.zip

After you run the sample script, you can view a list of source tables. From the list of tables, you can use the following tables to perform the tasks in this tutorial:

- CUSTOMER
- CUSTS_WITH_INCOMPLETE_COUPONS
- STATEMENT_LINE

Download and Run the Script

Download and run the sample script to populate the tables in the Oracle database.

1. Start a Microsoft Internet Explorer or Google Chrome browser.
2. In the Address field, enter the following URL:

   https://kb.informatica.com/proddocs/Product%20Documentation/3/TDM_GettingStarted_SampleSourceScript.zip
3. Download the `TDM_GettingStarted_SampleSourceScript.zip` file.
4. Run the SQL queries in the script to populate the source tables in the database.

   Note: This sample script is for Oracle database. You might need to modify the script for other databases.

CHAPTER 3

Creating Users and Groups in Informatica Administrator

This chapter includes the following topics:

- Creating Users and Groups in Informatica Administrator Overview, 15
- Step 1. Log In to Informatica Administrator, 16
- Step 2. Create a User, 16
- Step 3. Create a Group, 18
- Step 4. Assign Users to Groups, 19
- Step 5. Assign Roles and Privileges to Users and Groups, 20

Creating Users and Groups in Informatica Administrator Overview

In this lesson, you log in to Informatica Administrator to create a user and user group and add the user to the group. You assign the required roles and privileges to the user group.

Lesson Concepts

To access the TDM service, other application services, and objects in the Informatica domain, and to use the application clients, you must have a user account. The tasks that you can perform depend on the type of user account you have and the type of license.

You can set up different types of user accounts in the Informatica domain. Users can perform tasks based on the roles, privileges, and permissions assigned to them.

You can create, edit, and delete groups, and add users to the groups. You can assign roles and privileges to a group. The roles and privileges assigned to the group determine the tasks that users in the group can perform within the Informatica domain.

You create users and groups and assign roles and privileges from Informatica Administrator.

Lesson Objectives

In this lesson, you perform the following tasks:

- Log in to Informatica Administrator.
- Create a user.

- Create a group.
- Add the user to the group.
- Assign roles and privileges to the group.

Lesson Prerequisites

Before you start this lesson, verify the following prerequisites:

- You have installed Informatica services.
- You have installed Test Data Management.
- You have created a Test Data Manager service.
- The Informatica services are running in the domain.

Lesson Timing

Set aside about 15 minutes to complete the tasks in this session.

Step 1. Log In to Informatica Administrator

Create users and groups, and assign privileges and roles to users and groups from Informatica Administrator. You must have a valid user ID to log in to the Administrator tool.

1. Start Microsoft Internet Explorer or Google Chrome.
2. In the Address field, enter the following URL for the Administrator tool login page:

   ```
   http://<host>:<port>/administrator
   ```

   The host is the gateway node host name. The port is the Informatica Administrator port number.
3. On the Informatica Administrator login page, enter the 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.
5. Click **Log In**.

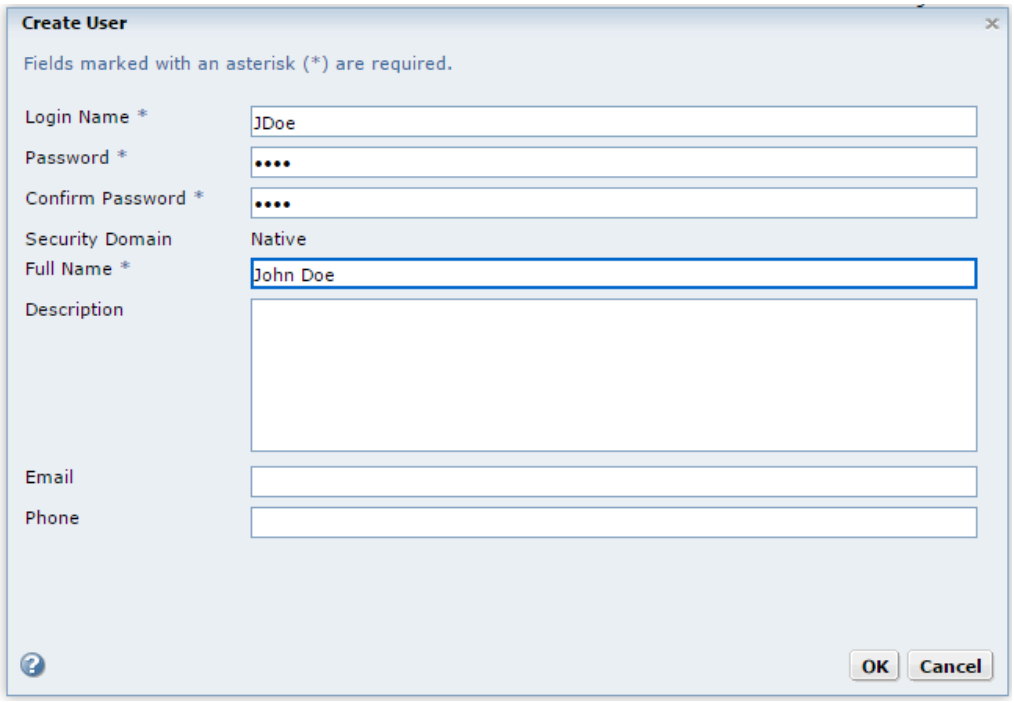
You logged in to Informatica Administrator. You can now create users and user groups, and assign roles and privileges to the users and user groups.

Step 2. Create a User

Create a user account that can log in to Test Data Manager. Create users from the **Security** tab of the Informatica Administrator.

1. In the Administrator tool, click the **Security** tab.
2. On the Security tab **Actions** menu, click **Create User**.

The following image shows the properties that you can set for a user:



3. Enter the following required fields:

Login Name

Login name for the user account. The login name for a user account must be unique within the security domain to which it belongs. The name is not case sensitive and cannot exceed 128 characters. It cannot include a tab, newline character, or the following special characters:
, + " \ < > ; / * % ? &

The name can include an ASCII space character except for the first and last character. No other space characters are allowed.

Password

Password for the user account. The password can be from 1 through 80 characters long.

Confirm Password

Enter the password again to confirm. You must retype the password. Do not copy and paste the password.

Full Name

Full name for the user account. The full name cannot include the following special characters:
< > “

4. Click **OK**.

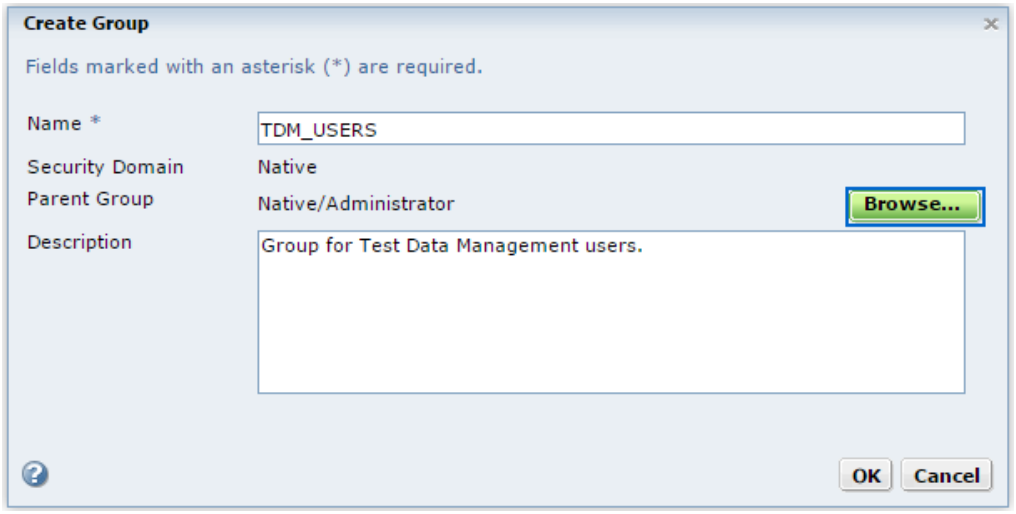
You created a user account that, with the correct privileges, can log in to application clients, such as Test Data Manager, the Administrator tool, or the Analyst tool.

Step 3. Create a Group

Create a user group called TDM_USERS to add users to and assign roles and privileges. Create groups from the **Security** tab of Informatica Administrator.

1. In the Administrator tool, click the **Security** tab.
2. On the Security **Actions** menu, click **Create Group**.

   The following image shows the properties that you can set for a group:



3. Enter the following details for the group:

   Name

   Name of the group. The name is not case sensitive and cannot exceed 128 characters. It cannot include a tab, newline character, or the following special characters:
   , + " \ < > ; / * % ?

   The name can include an ASCII space character except for the first and last character. No other space characters are allowed.

   Parent Group

   Group to which the new group belongs. If you select a native group before you click **Create Group**, the selected group is the parent group. Otherwise, the Parent Group field displays Native indicating that the new group does not belong to a group.

   Description

   Description of the group. The group description cannot exceed 765 characters or include the following special characters:
   < > “

4. To select a different parent group, click **Browse**.

   You can create more than one level of groups and subgroups.

5. To save the group, click **OK**.

You created the user group TDM_USERS to which you can assign users, privileges, and roles.

Step 4. Assign Users to Groups

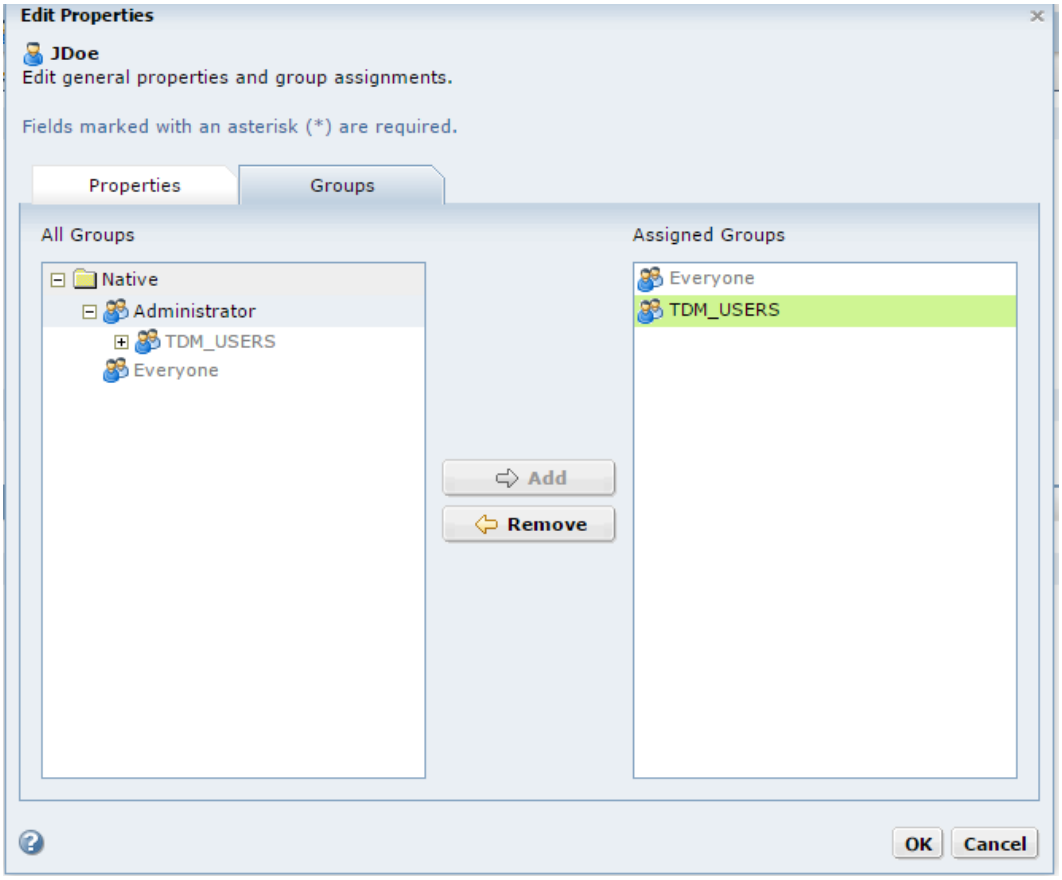
Assign users to groups to assign roles and privileges to multiple users at a time and to manage user roles and privileges. Assign users to groups from the Security tab.

1. In the Administrator tool, click the **Security** tab.
2. In the **Users** section of the Navigator, select the native user account that you created, and then click **Edit**.
3. Click the **Groups** tab.
4. To assign the user to the group TDM_USERS you created, select TDM_USERS from the **All Groups** column and click **Add**.

   If nested groups do not appear in the **All Groups** column, expand each group to show all nested groups.

   You can assign a native user to more than group. Use the Ctrl or Shift keys to select multiple groups at the same time.
5. To save the group assignments, click **OK**.

The following image shows the **Groups** tab of the user properties page from where you can assign a user to user groups:



You assigned the user to the user group TDM_USERS. You can add multiple users to a group and assign privileges and roles to the group. All users in the group inherit the privileges and roles of the group.

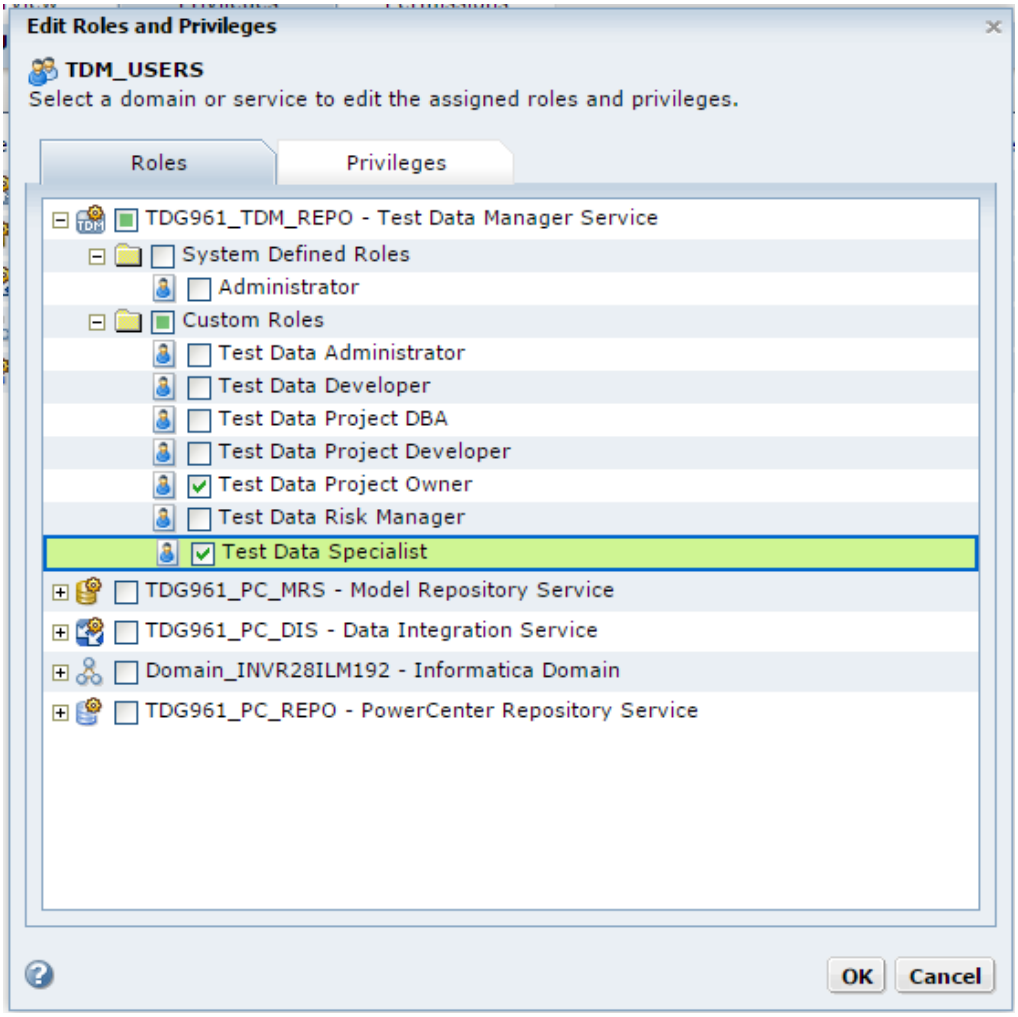
Step 5. Assign Roles and Privileges to Users and Groups

Assign roles and privileges to individual users or to a user group to manage multiple user roles and privileges. Assign roles and privileges to users and user groups from the **Security** tab of the Informatica Administrator. In this exercise we assign all the Test Data Manager service privileges to the user group TDM_USERS.

1. In the Administrator tool, click the **Security** tab.
2. In the Navigator, select the user group TDM_USERS.
3. Click the **Privileges** tab.
4. Click **Edit**.

   The **Edit Roles and Privileges** dialog box appears.
5. To assign roles, expand the Test Data Manager service on the **Roles** tab.
6. To grant roles, select the Test Data Manager service roles to assign to the group.

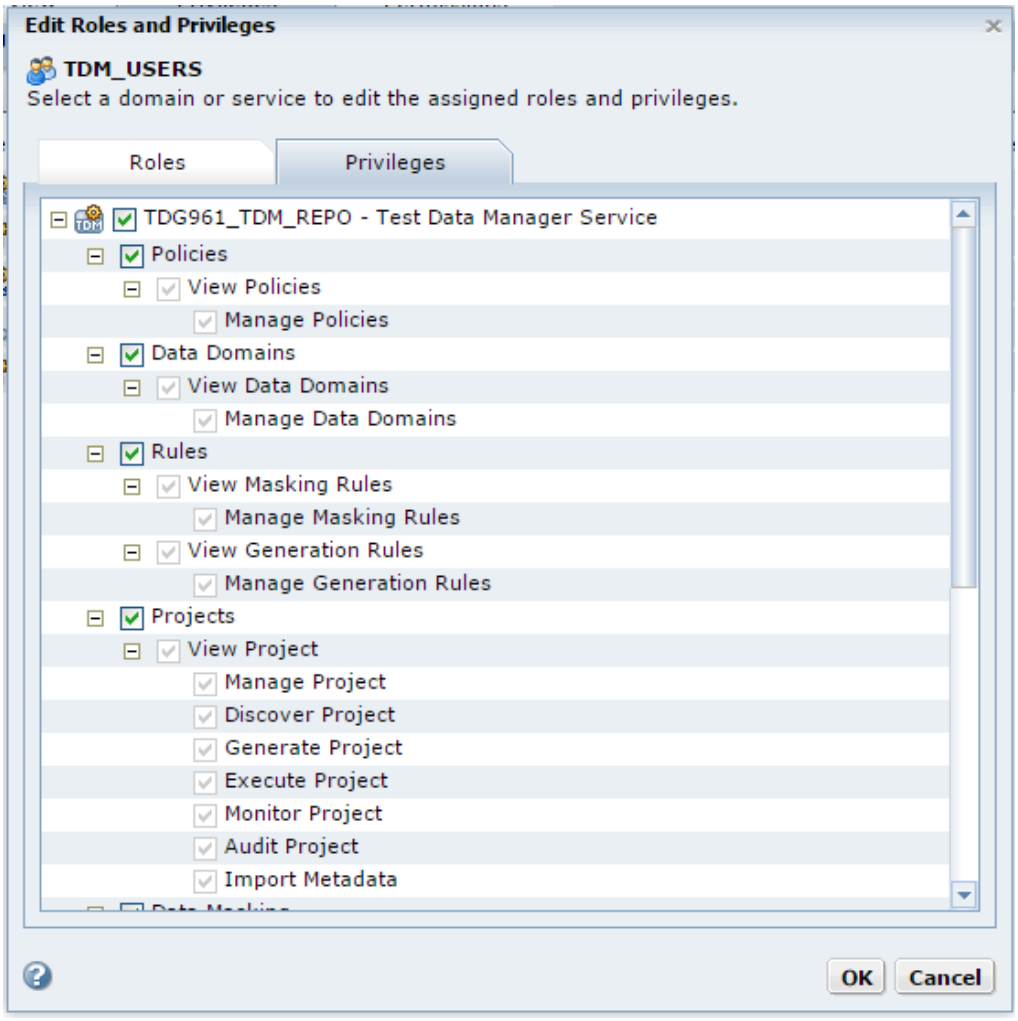
   The following image shows the custom roles available to choose from:



You can select any role that includes privileges for the selected domain or application service type.

7. To assign privileges, click the **Privileges** tab.
8. Expand the Test Data Manager service.
9. To grant privileges, select all the privileges to assign to the user group.

   The following image shows the privilege groups available:



   You cannot revoke privileges inherited from a role or group.

10. Click **OK**.

You assigned roles and privileges to the user group TDM_USERS. Users in the user group TDM_USERS can now perform TDM service tasks.

CHAPTER 4

Setting Up Test Data Manager

This chapter includes the following topics:

- Setting Up Test Data Manager Overview, 22
- Step 1. Log In to Test Data Manager, 23
- Step 2. Create Source and Target Connections, 23
- Step 3. Create a Project, 24
- Step 4. Import Data Sources, 25

Setting Up Test Data Manager Overview

In this lesson, you set up Test Data Manager to perform the data masking and data subset operations. You must log in to Test Data Manager and create source and target connections. To store the TDM components, you must create a project and import data sources from the connections that you created.

Lesson Concepts

You need an Informatica user account to log in to Test Data Manager. Create source and target connections to perform data subset and data masking operations. You must have administrator privileges to create source and target connections. Create a project to store the data masking and data subset components. Import data sources to define the source and target connections. You can import data sources from the PowerCenter repository or an external database to a project in Test Data Manager.

Lesson Objectives

In this lesson, you perform the following tasks:

- Log in to Test Data Manager with the Informatica user account.
- Create source and target connections.
- Create a project to store the TDM components.
- Import data sources into the project.

Lesson Prerequisites

Before you start this lesson, complete the previous lesson in the tutorial.

Lesson Timing

Set aside 15 to 20 minutes to complete the tasks in this lesson.

Step 1. Log In to Test Data Manager

Log in to Test Data Manager with the Informatica user account that you created in the Informatica Administrator tool.

1. Start a Microsoft Internet Explorer or Google Chrome browser.
2. In the **Address** field, enter the URL for Test Data Manager:

   ```
   http://<HostName>:<PortNumber>/tdm
   ```

   HostName represents the host name of the machine where you install TDM. *PortNumber* is the startup port number for TDM. The default port number is 6605.

   If you configure TDM to use HTTPS, the URL opens the HTTPS site:

   ```
   https://<HostName>:<HTTPSPortNumber>/tdm
   ```
3. On the login page, enter the Informatica user name and the password.
4. Click **Login**.

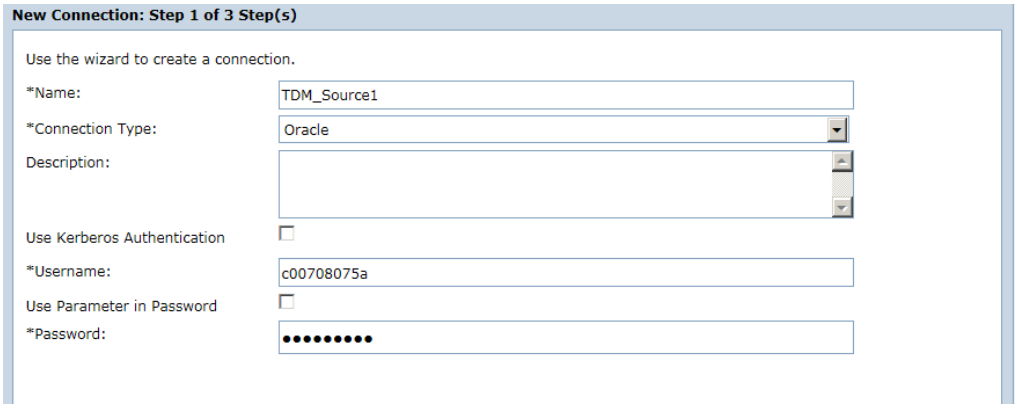
Step 2. Create Source and Target Connections

Create a source connection and a target connection in Test Data Manager.

1. In Test Data Manager, click **Administrator**.
2. Click **Connections**.
3. Click **Actions** > **New Connection**.

   The New Connection wizard opens.
4. Enter the connection name TDM_Source1 for the source.
5. Select Oracle connection type.
6. Optionally, enter a description for the source connection.
7. Enter the user name and password for the source database.

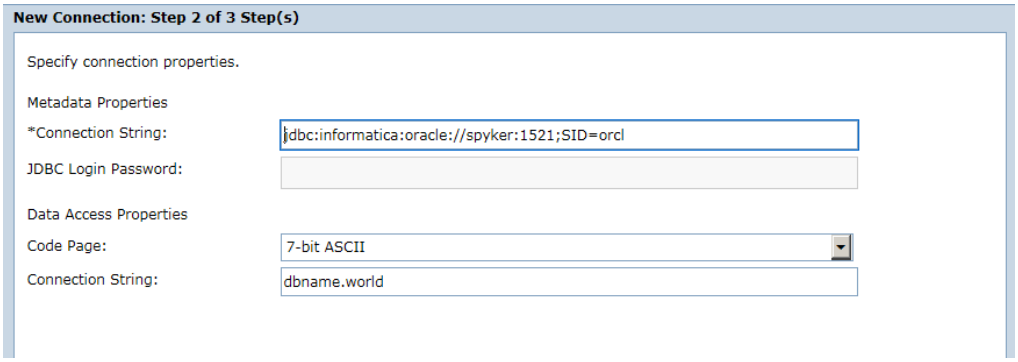
   The following image shows connection properties for a sample source database:



8. Click **Next**.
9. Enter the following connection string for the Oracle connection:

   ```
   jdbc:informatica:oracle://<hostname>:1521;SID=<sid>
   ```

The following image shows the metadata properties for a sample source database:



10. To test the connection, click **Test Connection**.
11. To save the connection, click **Finish**.

    The connection is visible in the **Administrator** | **Connections** view.
12. Repeat the above steps to create a target connection named TDM_Target1 with a separate database source connection.

Step 3. Create a Project

Create a project to store the data discovery, data subset, and data masking components that you can apply to the data source.

1. Click **Projects**.
2. Click **Actions** > **New**.
3. In the **New Project** dialog box, enter the following project properties:

   Name

   Enter the name CustDetails for the folder.

   Description

   Optional. Enter a description.

   PowerCenter Repository

   The default is the name of the PowerCenter repository.

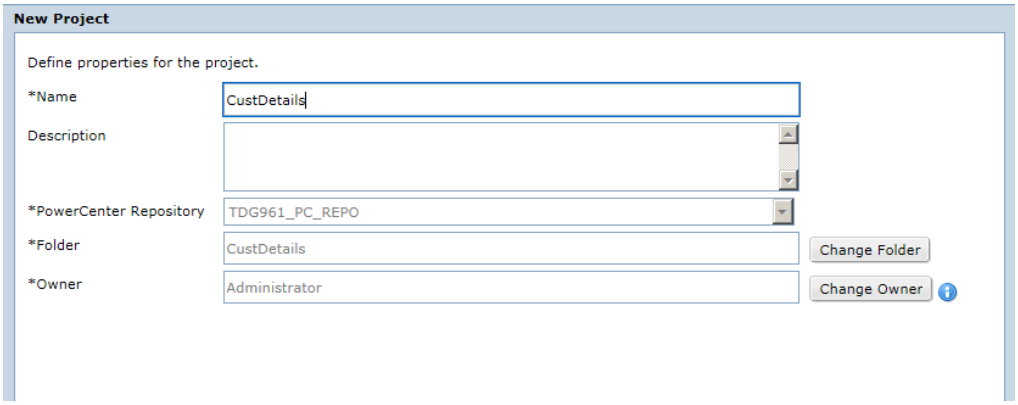
   Folder

   The default is the project name.

   Owner

   The default is the name of the user that created the folder.

The following image shows the properties of the project:



4. Click **OK**.

   View the properties of the CustDetails project that you created.

Step 4. Import Data Sources

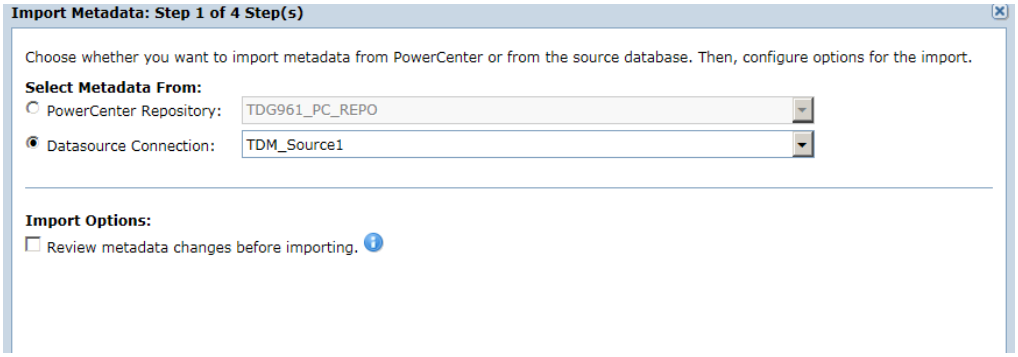
Import data sources from the TDM_Source1 database connection.

1. In the project ,click **Actions** > **Import Metadata**.

   The **Import Metadata** dialog box appears.

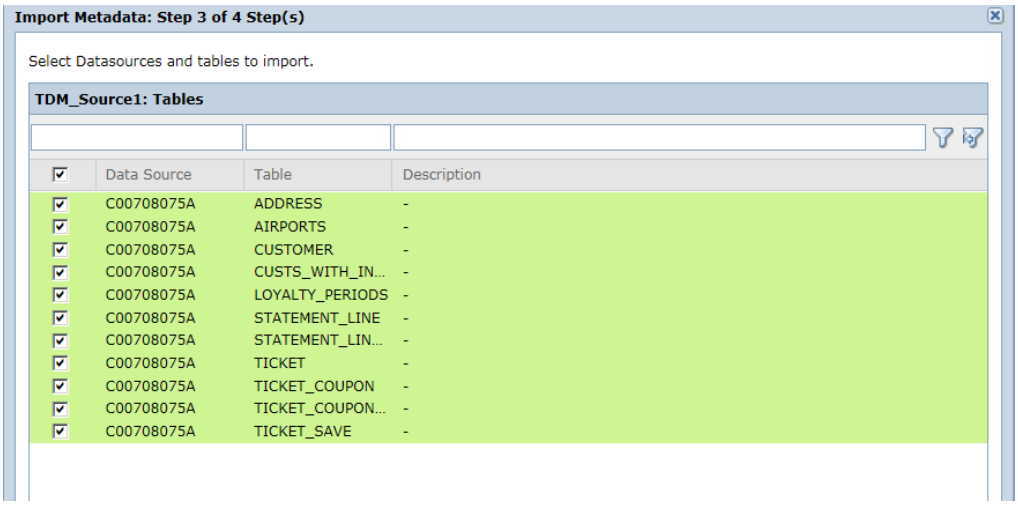
2. To import metadata from a database, choose **Datasource Connection** and select TDM_Source1 database connection.

   The following image shows the options to import metadata:



3. Choose whether you want to review the metadata changes before you import the data sources. You can choose to skip the import options.
4. Click **Next**.
5. Select the C00708075A metadata schema to import from the database. You can filter folders by folder name or description.
6. Click **Next**.
7. Select all the tables within the data source C00708075A to import.

The following image shows the sample tables in a database connection:



8. Click **Next**.
9. To import the data source immediately, select **Import Now**.
10. Click **Finish**.

    While the import job runs, view the progress of the import job in the **Monitor** view. After the import job finishes, you can access the imported metadata through the **Data Sources** details view.

CHAPTER 5

Creating Logical Relationships in TDM

This chapter includes the following topics:

- Creating Logical Relationships in TDM Overview, 27
- Step 1. Create a Logical Relationship Between Tables, 28

Creating Logical Relationships in TDM Overview

In this lesson, you create logical relationships between tables in the source data.

Lesson Concepts

Relationships might exist between data in tables even though a primary key relationship does not exist in the data source. In the absence of physical keys in the data, there must be a way to indicate related tables so that the tables are included in a data subset you create in TDM.

You can create relationships between tables that you want to include in a subset operation in TDM. Relationships that you create in TDM are logical relationships. Logical relationships do not impact the data in the data source. When you create an entity, TDM uses the relationships between tables to determine the tables to include in the entity. By using the entity in a subset operation, you ensure that all related tables are included in the subset operation.

Identify relationships that you want to add to the TDM repository and then create the logical relationships as constraints in Test Data Manager. You can create relationships with either major or minor severity. The severity determines the scope of the data that a data subset receives based on the constraints. When you define a constraint with a major severity level, all of the children of the parent are included in the data subset. When you do not want the parent table to select additional child table records that are related to the parent, you assign a minor constraint between two tables.

Lesson Objectives

The source data contains customer information spread across tables with no physical keys in the source data. You must create a subset that contains all information on a particular set of customers. The tables CUSTOMER, CUSTS_WITH_INCOMPLETE_COUPONS, and STATEMENT_LINE have customer information that you must include in the subset data. By creating a relationship between these tables, you ensure that all the tables are included in an entity when you create an entity with one of the tables. All three tables have a customer ID column. Create foreign key relationships on the CUSTS_WITH_INCOMPLETE_COUPONS and STATEMENT_LINE tables that link them to the customer ID column in the CUSTOMER table.

In this lesson, you complete the following tasks:

- Create a constraint on the CUSTS_WITH_INCOMPLETE_COUPONS table to create a logical relationship between the CUSTOMER and CUSTS_WITH_INCOMPLETE_COUPONS tables.
- Create a constraint on the STATEMENT_LINE table to create a logical relationship between the CUSTOMER and STATEMENT_LINE tables.

Lesson Prerequisites

Before you start this lesson, complete the earlier lessons in this tutorial.

Lesson Timing

Set aside 20 minutes to complete the tasks in this lesson.

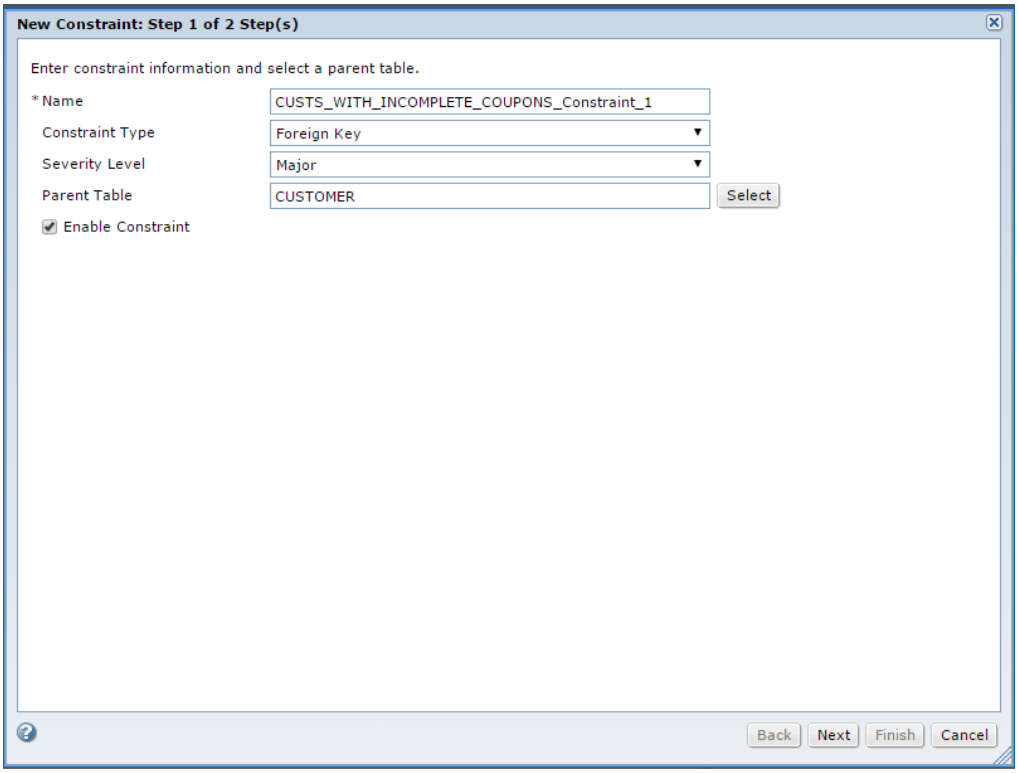
Step 1. Create a Logical Relationship Between Tables

Select a column from a table that you identify as the parent table to establish a foreign key relationship in another table. In this lesson, we select the CUST_ID column in the CUSTOMER table.

Perform the following steps to create a foreign key constraint in the CUSTS_WITH_INCOMPLETE_COUPONS table. The foreign key relates the CUST_ID column in this table with the CUST_ID column in the parent table CUSTOMER. Repeat the steps to create a foreign key constraint in the STATEMENT_LINE table.

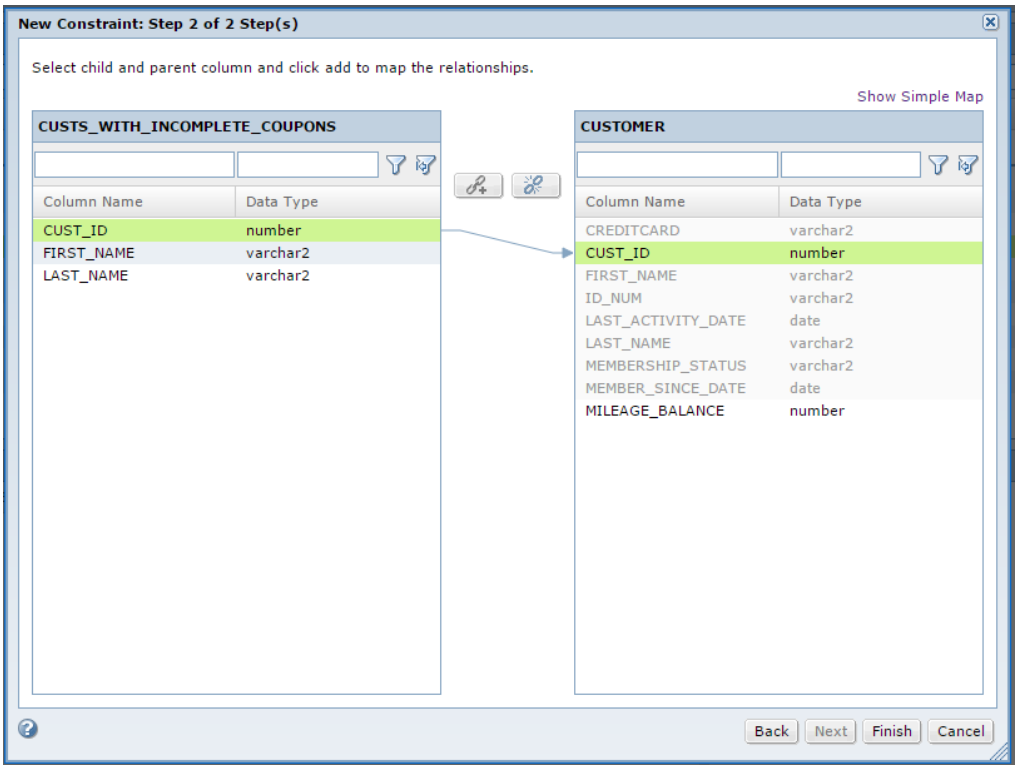
1. In the project, click **Discover** > **Tables**.
2. Select the CUSTS_WITH_INCOMPLETE_COUPONS to create a foreign key.
3. Click **Constraints**.
4. Click **Create New Constraint**.

5. Enter the name. TDM uses the table name and appends a constraint number to it by default. You can edit this if required. In this exercise, we retain the default name.



6. Select **Foreign Key** from the constraint type list.
7. Select the **Major** severity level from the list.
8. Click **Select** and browse to select the CUSTOMER table as the parent table to which the foreign key must relate.
9. To enable the constraint, select the **Enable Constraint** check box and click **Next**.
10. From the left pane, click the CUST_ID column. From the right panel, click the CUST_ID column in the CUSTOMER table. Click the **Link** icon to map the child-parent relationship.

The following image shows mapping between the tables:



11. Click **Finish**.

You created a foreign key constraint in the CUSTS_WITH_INCOMPLETE_COUPONS table that relates to the CUSTOMER table. You repeated the steps to create a foreign key constraint in the STATEMENT_LINE table that relates to the CUSTOMER table. You can now create an entity with the CUSTOMER table as the driving table.

CHAPTER 6

Defining Data Subset Components

This chapter includes the following topics:

- Defining Data Subset Components Overview, 31
- Step 1. Create an Entity, 32

Defining Data Subset Components Overview

In this lesson, you create an entity CUSTOMER_DATA_ENTITY with the CUSTOMER table as the driving table.

The CUSTS_WITH_INCOMPLETE_COUPONS table, and the STATEMENT_LINE table contain foreign keys that relate to the CUSTOMER table.

Lesson Concepts

You can create a subset of production data if you need a small, targeted, and referentially intact copy of the production data to use in a nonproduction environment.

In the source data, customer information exists in multiple tables. You need data from more than one table when you create a subset of specific customers. The source data does not contain physical keys. You created logical constraints to create relationships between these tables. To ensure that the subset operation includes all the related tables, create an entity to use in the subset operation. An entity defines a set of related tables. The entity defines the structure for copying related data to the subset database. When you include the entity in the subset operation, the operation considers all tables in the entity.

An entity consists of a driving table and related tables. A driving table is the starting point for defining relationships between tables in the entity. TDM defines the tables that relate to the driving table based on physical and logical constraints. You can add tables that have cyclic or circular relationships within the table or between tables. You must add a staging connection before you can add these tables to an entity.

If you perform data discovery in TDM, you can run an entity profile to discover possible entities in the data. You can also manually create entities. In this lesson you manually create an entity. When you create an entity, select parameters to filter data in the columns that you want to include in the subset database. In this lesson, we create a data subset that contains information on customers with customer IDs greater than 100. The subset data would therefore include all data related to the specific customers.

Lesson Objectives

In this lesson, you perform the following task:

- Create an entity named CUSTOMER_DATA_ENTITY with the CUSTOMER table as the driving table.

Lesson Prerequisites

Before you start this lesson, complete the earlier lessons in this tutorial.

Lesson Timing

Set aside 15 minutes to complete the tasks in this lesson.

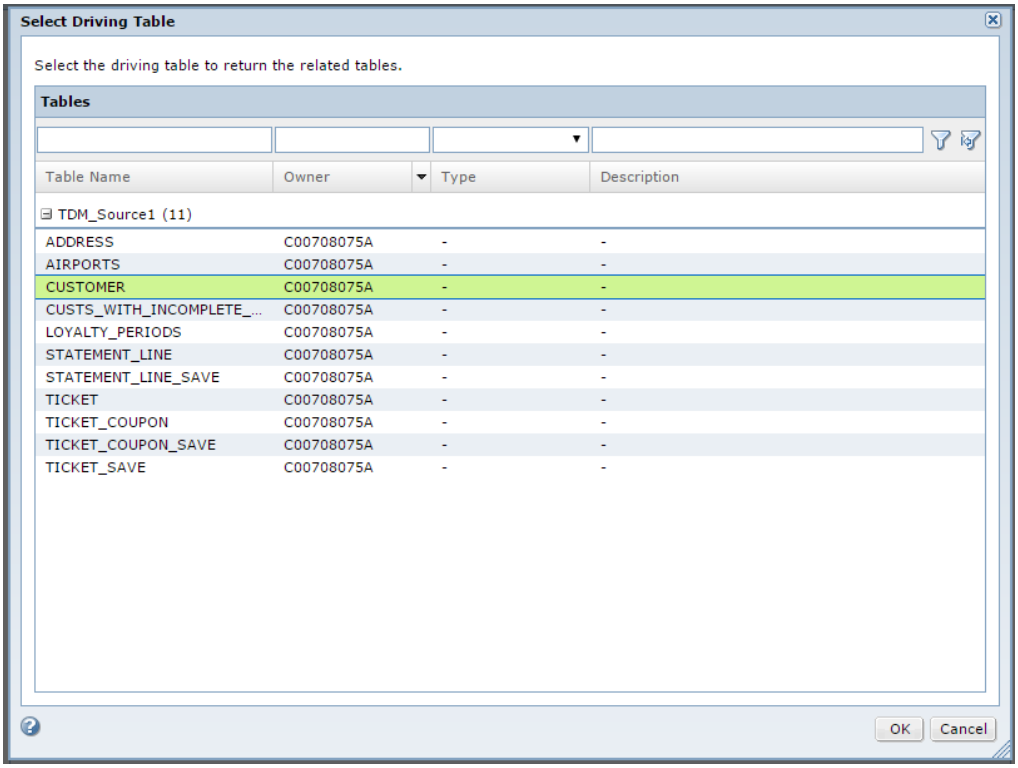
Step 1. Create an Entity

When you create an entity, you select the driving table. Test Data Manager retrieves the related tables based on the constraints. You create an entity CUSTOMER_DATA_ENTITY that uses the CUSTOMER table as the driving table. Create a filter on the CUST_ID column in the CUSTOMER table.

Before you create an entity, identify relationships in the data and create constraints to define the child tables.

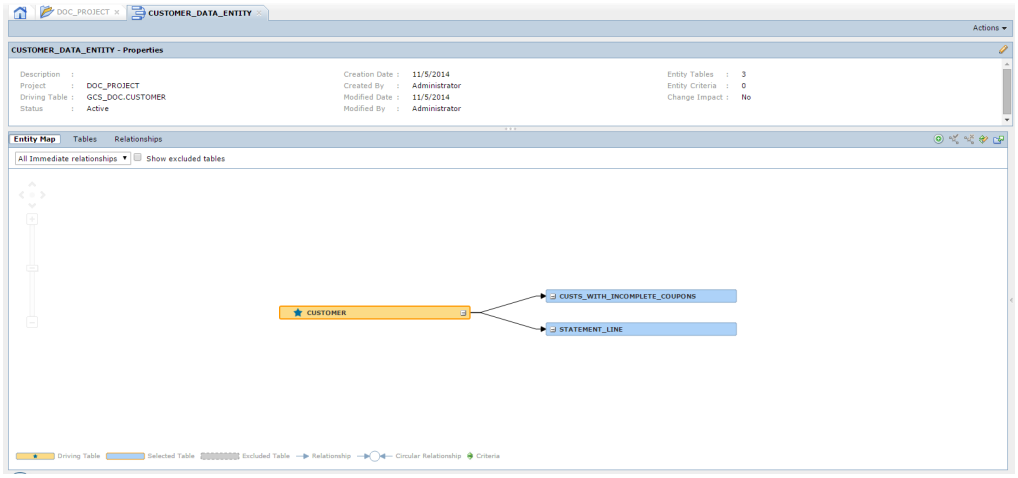
1. In the project, click **Define** > **Data Subset**.
2. Click **Actions** > **New** > **Entities**.
3. In the **New Entity** dialog box, enter the name CUSTOMER_DATA_ENTITY and optional description for the entity.
4. To select a driving table from the list, click **Select Table**.
5. Select the CUSTOMER table and click **OK**.

The following image shows the list of tables from which you select the driving table:



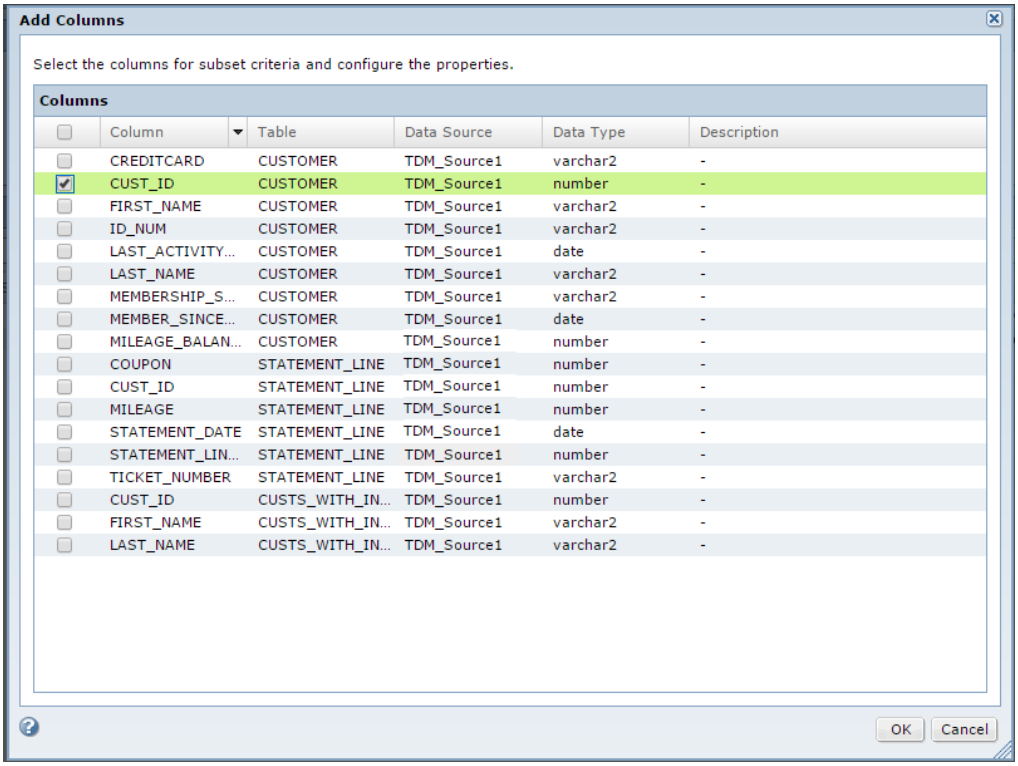
Test Data Manager creates the entity by including all tables related to the driving table. The entity opens with a map of the relationship between the tables in the **Entity Map** tab. You can view a list of the tables on the **Tables** tab. You can view a list that shows the relationships between the tables on the **Relationships** tab.

The following image shows the entity map view of the entity CUSTOMER_DATA_ENTITY that you created:



6. To create a subset of the data based on filter criteria, click **Edit** in the **Properties** pane.
7. Click the **Entity Criteria** tab to add criteria.
8. Click **Add Criteria**.
9. To filter data, select the column and table on which you want to filter data and click **OK**.

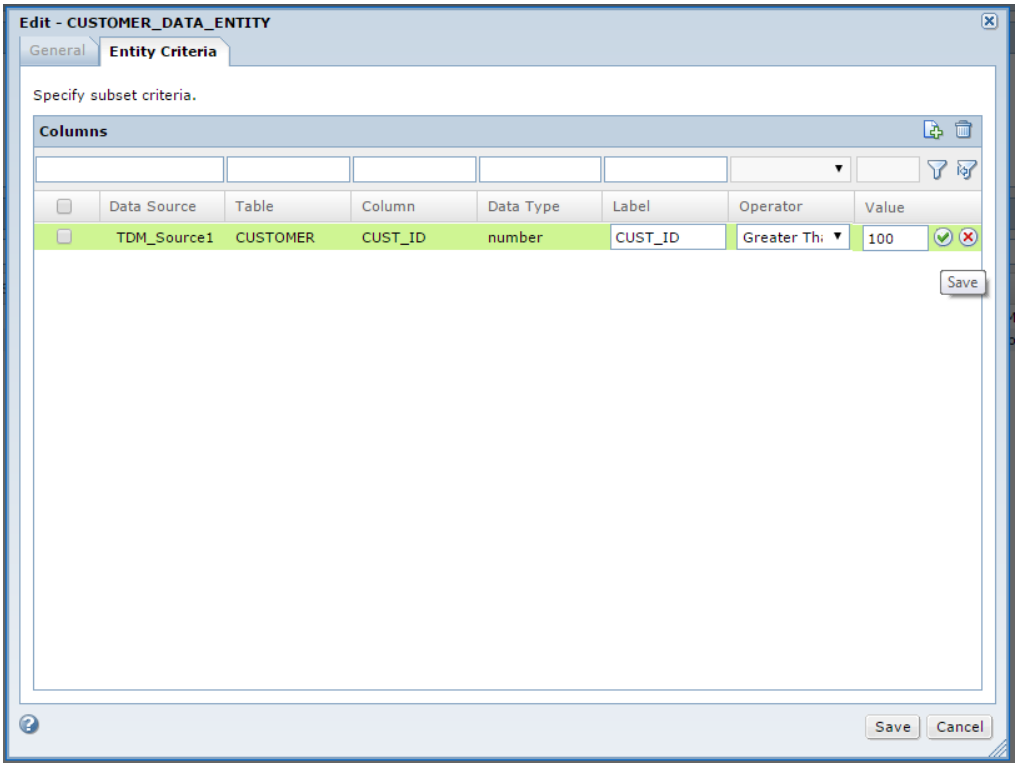
The following image shows the list of columns from which you select a column to apply filter criteria:



You can indicate a filter condition in the entity, but define the expression in the plan.

10. To define the filter expression in the entity, select the operator **Greater Than** from the list to filter the data.

The following image shows the **Entity Criteria** tab where you enter the filter criteria:



11. Enter `100` in the value to complete the filter expression. This creates a filter for customers with customer IDs greater than 100.
12. For each filter criteria that you create, click **Save**. If you define multiple filters in an entity, the filter conditions are considered as "AND."
13. Select the row and click **Save**.

You can use the entity in a subset operation. Add the entity to a plan from the **Add Subset Components** page in the plan workflow.

CHAPTER 7

Creating a Data Masking Rule

This chapter includes the following topics:

- Creating a Data Masking Rule Overview, 36
- Step 1. Create a Standard Data Masking Rule, 37
- Step 2. Add the Standard Masking Rule to the Project, 38
- Step 3. Assign the Masking Rule, 39

Creating a Data Masking Rule Overview

In this lesson, you create a standard credit card masking rule, add the rule to the project, and assign the rule to a column to mask the credit card numbers.

Lesson Concepts

Create a data masking rule to replace source data in sensitive columns with realistic test data for nonproduction environments. Data masking rules define how to mask sensitive and confidential data in a target database. When you create data masking rules, you define the logic to replace sensitive data.

Use repeatable masking when you generate a data masking workflow more than once and you need to return the same masked values each time it runs. Apply a seed value to create repeatable output for data masking output. The seed value is a starting point for generating masked values.

To configure the sensitive columns that you want to mask, assign data masking rules to source columns.

Lesson Objectives

In this lesson, you perform the following tasks:

- Create a credit card masking rule.
- Add the masking rule to the CustDetails project that you created.
- Assign the rule to the CREDITCARD column to mask the credit card numbers.

Lesson Prerequisites

Before you start this lesson, complete the earlier lessons in this tutorial.

Lesson Timing

Set aside 15 to 20 minutes to complete the tasks in this lesson.

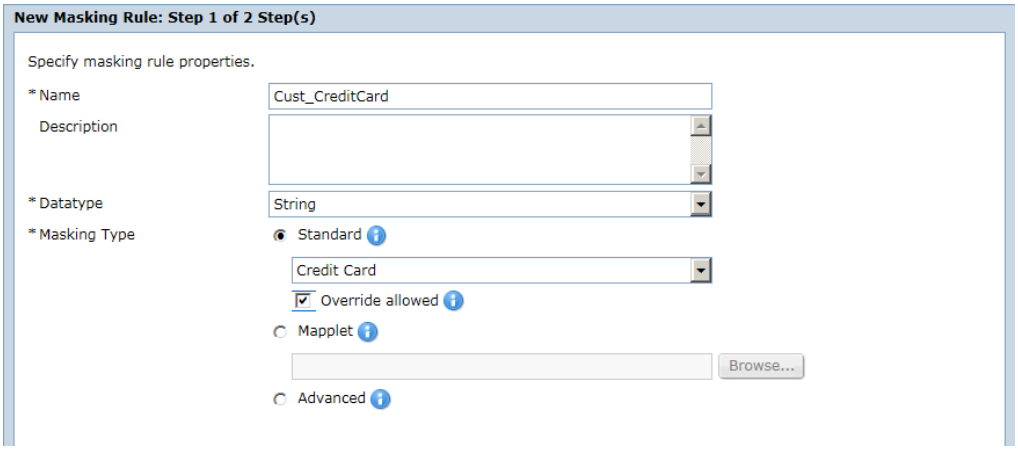
Step 1. Create a Standard Data Masking Rule

Create a rule to define a masking technique, the data type to mask, and masking parameters that define how to apply the technique.

1. In Test Data Manager, click **Policies**.
2. Click **Actions** > **New** > **Masking Rule**.

   The **New Masking Rule** wizard appears.
3. Enter the name Cust_CreditCard and optional description for the rule.
4. Select the **String** data type.
5. Choose **Standard** and select the **Credit Card** masking rule.
6. To enable users to override masking parameters for the rule, select the **Override Allowed** option.

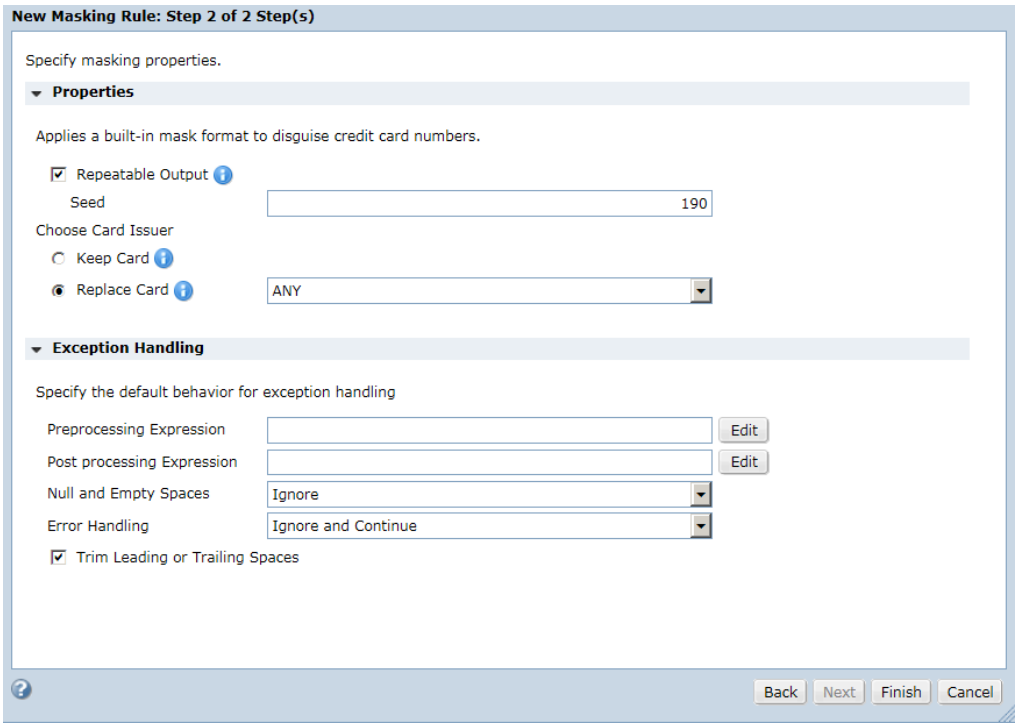
   The following image shows sample rule properties for credit card masking:



7. Click **Next**.

   Note: The **Masking Parameters** dialog box changes based on the **Masking Type** that you select.
8. Select **Repeatable Output** and enter 190 for the seed value.
9. Choose **Replace Card** and select **Any**.
10. To avoid null and empty spaces, select **Ignore**.
11. To configure error handling, select **Ignore and Continue**.
12. To trim spaces before comparing values, select **Trim Leading or Trailing Spaces**.

The following image shows sample credit card masking rule parameters:



13. Click **Finish**.

Step 2. Add the Standard Masking Rule to the Project

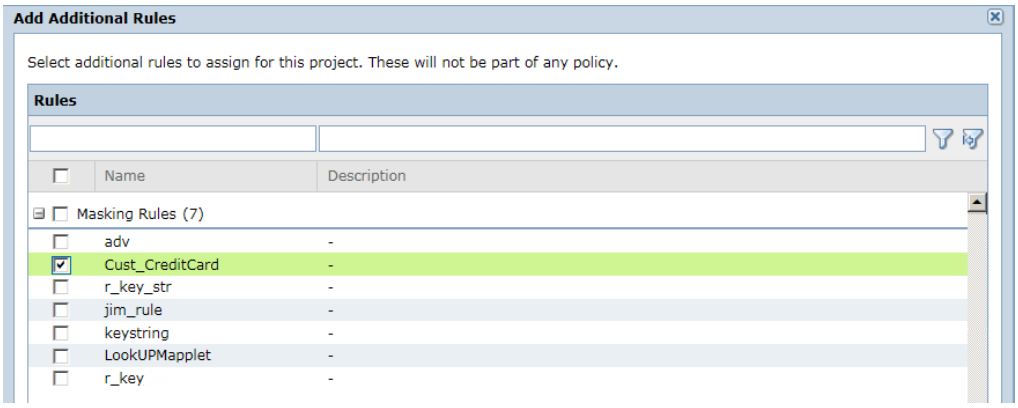
Add the standard credit card masking rule Cust_CreditCard to the Cust_Details project that you created.

1. To view the list of projects, click **Projects**.
2. Open the Cust_Details project that you created.

   The project opens in another tab.
3. Click **Overview** > **Policies**.
4. Click **Actions** > **Add Additional Rules**.

   The **Add Additional Rules** dialog box appears.
5. From the Masking Rules list, select the Cust_CreditCard masking rule.

The following image shows a sample list of rules:



6. Click **OK**.

   The credit card masking rule appears in the **Additional Rules** list.

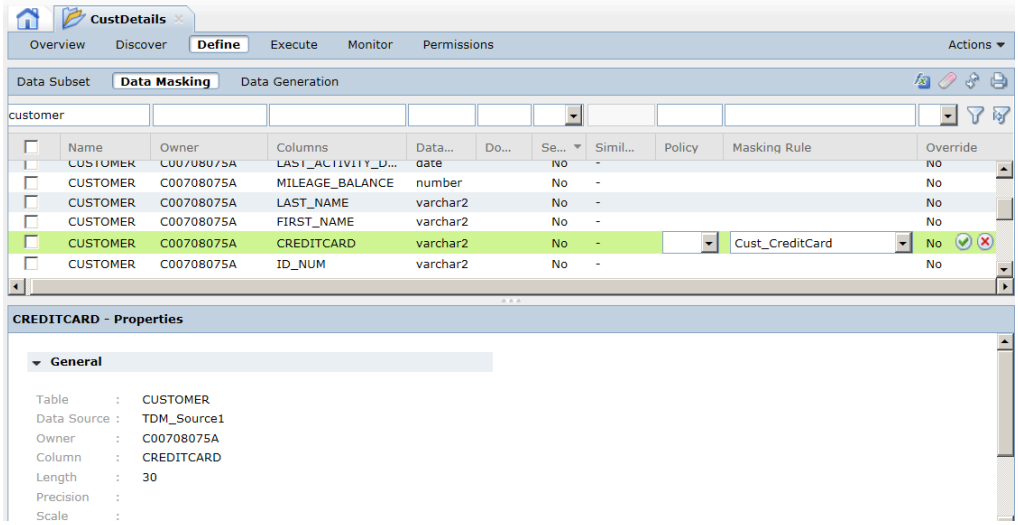
Step 3. Assign the Masking Rule

Assign the credit card masking rule that you added to the source table column that you want to mask.

1. In the project, click **Define** > **Data Masking**.
2. Select the CREDITCARD column to assign the masking rule to.
3. Click inside the **Masking Rule** column.

   A list of masking rules applicable for the string data type appears.
4. Select the Cust_CreditCard masking rule that you created.

   The following image shows a sample masking rule assignment:



5. Click **Save**.

CHAPTER 8

Creating a Plan

This chapter includes the following topics:

- Creating a Plan Overview, 40
- Step 1. Create a Plan, 40

Creating a Plan Overview

In this lesson, you create a plan and define the source and target connections. Define workflow properties in the plan, such as commit properties, update strategies, and recovery strategies.

Lesson Concepts

A plan includes the components that you need to generate a workflow. You can add data masking and data subset components to the same plan. You can create multiple workflows from one plan. When you add both subset and masking components to a plan, TDM performs the subset operation first. TDM then applies masking rules to columns in the subset data that have masking assignments.

Lesson Objectives

In this lesson, perform the following task:

- Create a plan and add the credit card masking rule and the entity. Configure the source and target databases with the connections and other plan properties.

Lesson Prerequisites

Before you start this lesson, complete the earlier lessons in this tutorial.

Lesson Timing

Set aside 10 to 15 minutes to complete the tasks in this lesson.

Step 1. Create a Plan

When you create a plan, add data subset and data masking components. To perform the data subset operation, add the CUSTOMER_DATA_ENTITY entity that you created. To perform the data masking operation, add the Cust_CreditCard masking rule that you created.

1. In the project, click **Execute**.
2. Click **Actions** > **New**.

3. In the **New Plan** dialog box, enter the name Cust_Details_Plan and optional description for the plan.
4. Click **Next**.
5. To add a data masking operation to the plan, click **Add Masking Components**.
6. Select the Cust_CreditCard masking rule to add to the plan. Click **Next**.
7. To add a data subset operation to the plan, click **Add Subset Components**.
8. Select the CUSTOMER_DATA_ENTITY entity to add to the plan. Click **Next**.
9. To skip adding a data generation component, click **Next**.
10. Review the masking and subset components.
11. Click **Next**.
12. In the **Source Connection** field, select **Relational**. Click **Select** and choose the TDM_Source1 database connection that you created.
13. In the Target Connection field, select **Relational**. Click **Select** and choose the TDM_Target1 database connection that you created.
14. Configure the following target settings:

    Truncate Tables

    To truncate the table before the data loads, select **Truncate Tables**.

    Disable Indexes

    To disable indexes for faster loading of data, select **Disable Indexes**.

    Disable Constraints

    To disable physical constraints before the data is loaded to the target database, select **Disable Constraints**. After the data loads, the constraints are enabled automatically.
15. Configure the following update strategy options:

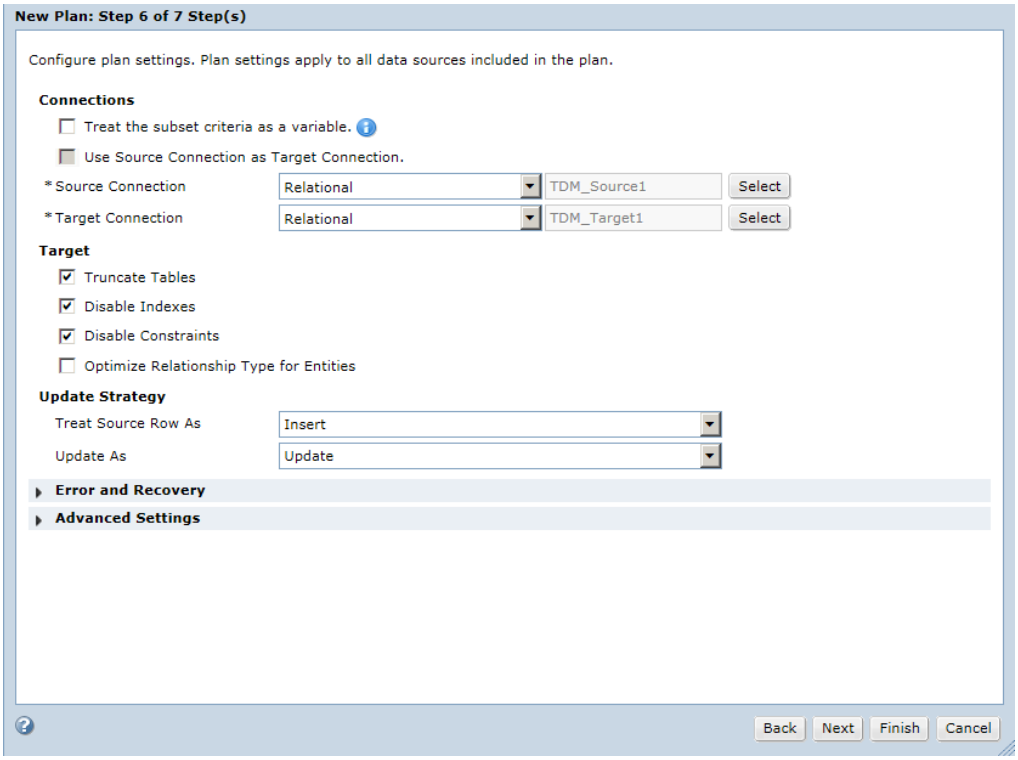
    Treat Source Rows As

    To insert the source rows, select **Insert**.

    Update As

    To update all the flagged rows that are marked for update, select **Update**.

The following image shows sample plan settings that you can configure:



16. Click **Next**.
17. Review the plan, tables, and data source settings.
18. Click **Finish**.

    The plan appears in the project.

CHAPTER 9

Managing the Workflow

This chapter includes the following topics:

- Managing the Workflow Overview, 43
- Step 1. Generate the Workflow, 44
- Step 2. Run the Workflow, 44
- Step 3. Monitor the Workflow, 45

Managing the Workflow Overview

In this lesson, you must generate and run the workflow to populate the data in the target database. You can monitor the progress of the workflow tasks.

Lesson Concepts

When you start a workflow, the PowerCenter Integration Service performs the plan operations. Generate a workflow from the plan to create PowerCenter mappings to perform data subset and data masking operations. You can monitor the workflow and session logs for a job. After you run the workflow, you can view the subset of the data with masked columns in the target database.

Lesson Objectives

In this lesson, you perform the following tasks:

- Generate a workflow for the plan that you created.
- Run the workflow.
- Monitor the workflow.

Lesson Prerequisites

Before you start this lesson, complete the earlier lessons in this tutorial.

Lesson Timing

Set aside 20 to 25 minutes to complete the tasks in this lesson.

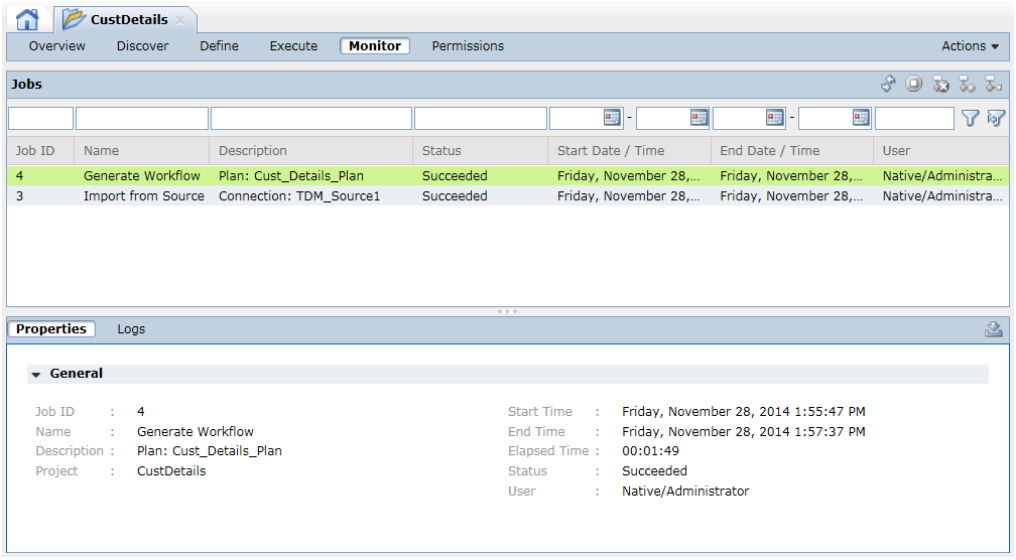
Step 1. Generate the Workflow

After you create the plan, generate the workflow.

1. In the project, click **Execute** to access the plans in the project.
2. Select Cust_Details_Plan.
3. Click **Actions** > **Generate Workflow**.

   The **Generate Workflow** dialog box appears.
4. Select **Schedule Now**.
5. Click **Generate Workflow**.

   View the status of the workflow generation in the **Monitor** tab.
   The following image shows a sample job status of the workflow generation:

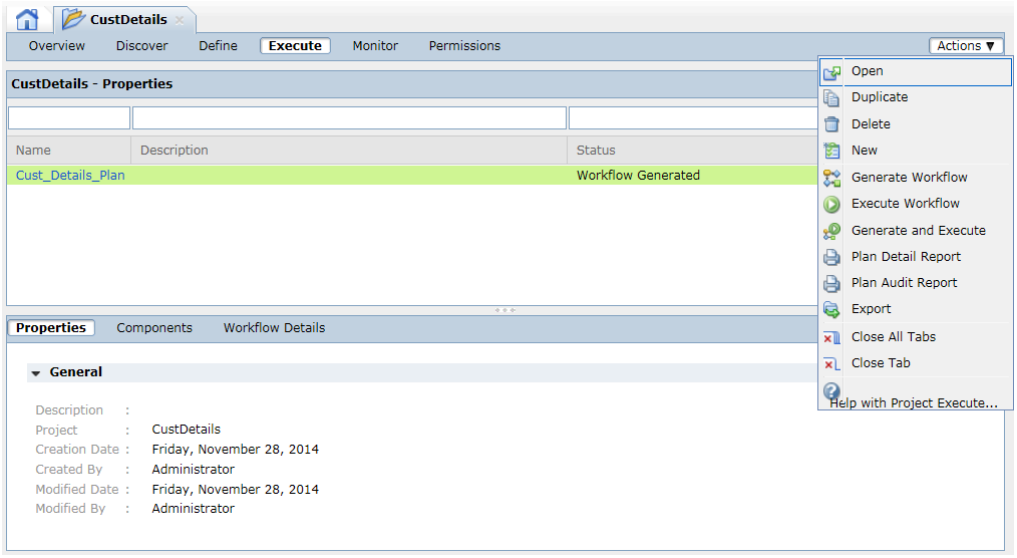


Step 2. Run the Workflow

After you generate the workflow, you must run the workflow to perform the data subset and masking operations.

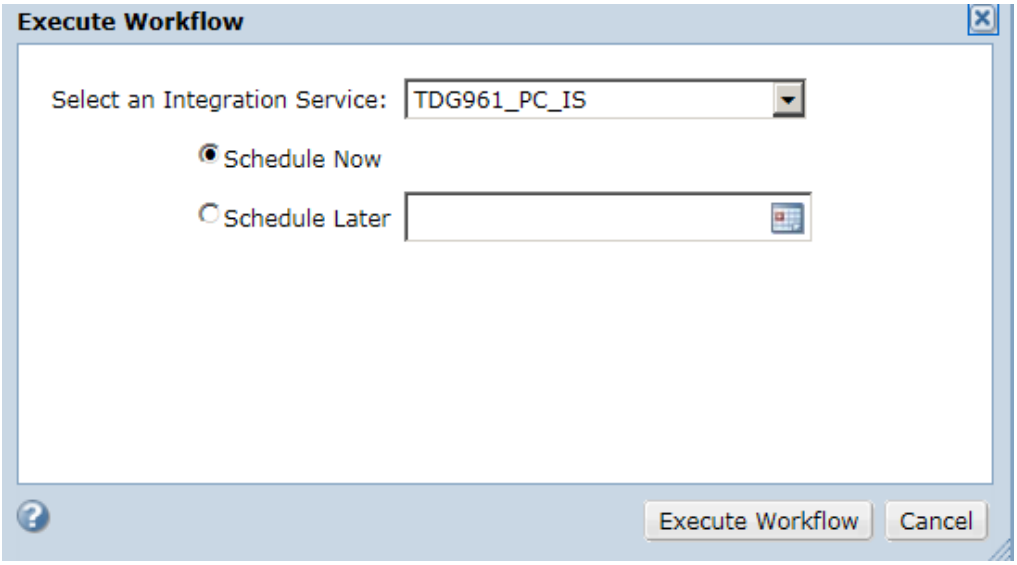
1. In the project, click **Execute** to access the plans in the project.
2. Select Cust_Details_Plan.
3. Click **Actions** > **Execute Workflow**.

The following image shows the **Execute** view from where you can run the workflow:



4. Select the PowerCenter Integration Service.
5. Select **Schedule Now**.
6. Click **Execute Workflow**.

   The following image shows the **Execute Workflow** dialog box:



Step 3. Monitor the Workflow

You can view the status of the workflow that runs in Test Data Manager.

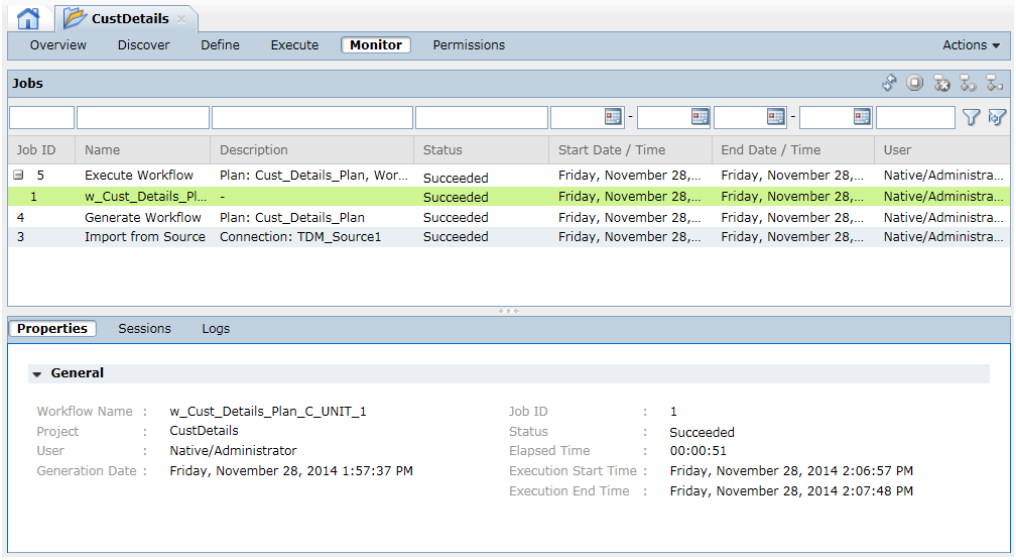
1. In the project, click **Monitor** to view the status of the workflow.
2. To refresh the view, click **Auto Refresh On**.

   View the progress in the Status column.

3. To view the workflow logs, select the workflow Job ID.

   View the summary of the workflow in the **Properties** tab.

   The following image shows a sample workflow job status:



APPENDIX A

Glossary

data discovery

The process of discovering the metadata of source systems that includes content, such as data values and frequencies, and structure such as primary keys, foreign keys, and functional dependencies.

data generation

The process to generate realistic test data for the testing environment without using the production data.

data masking

The process of replacing sensitive columns of source data with realistic test data.

data subset

A small, targeted, and referentially intact copy of production data.

entity

Defines a set of tables that are related based on physical or logical constraints. An entity can contain parent tables and child tables. An entity maintains relational hierarchy in the tables. Define filter parameters for ports in the entity to extract the columns from the tables in an entity. When you run a workflow generated from a data subset plan, the PowerCenter Integration Service extracts source data based on the filter parameters defined in the entity and loads the data to the subset tables.

foreign key profile

A type of data analysis that finds column values in a data source that match the primary key column values in another data source.

plan

Defines data subset, data masking, or data generation operations. You can add entities, groups, templates, policies, rules, and tables to a plan. When you generate and run workflows from a plan, the PowerCenter Integration Service generates and runs the workflows.

project

A container component for entities, groups, templates, and one or more sources that you want to use in data discovery, subset, and masking operations. When you create a project, you add one or more sources to the project. Any source that you add to a project is not available to other projects.

rule

Defines the data masking technique, an optional rule qualifier, and masking parameters.

TDM repository

A relational database that stores the components that you define in Test Data Manager, such as policies, projects, entities, and data masking rules. The TDM repository stores metadata that you import into Test Data Manager from a source database or from the PowerCenter repository. The TDM repository stores constraints that define relationships between the source tables in a project.

TDM Server

Server that runs Test Data Manager and integrates with Informatica application services to perform data subset, data masking, and data discovery operations.

Test Data Management (TDM)

The Informatica solution that bundles Data Subset, Data Generation, and Data Masking to protect sensitive data and create lean nonproduction systems for test and development purposes.

Test Data Manager

The web-based user interface that you use to configure and run data subset, data masking, and discovery operations.