Informatica® Test Data Management
10.1.0

Release Guide

Informatica Test Data Management Release Guide
10.1.0
December 2016

Publication Date: 2018-07-03

Table of Contents

Preface

The *Test Data Management Release Guide* lists new features and enhancements, behavior changes between versions, and tasks that you might need to complete after you upgrade from a previous version.

Informatica Resources

Informatica Network

Informatica Network hosts Informatica Global Customer Support, the Informatica Knowledge Base, and other product resources. To access Informatica Network, visit https://network.informatica.com.

As a member, you can:

- Access all of your Informatica resources in one place.
- Search the Knowledge Base for product resources, including documentation, FAQs, and best practices.
- View product availability information.
- 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 search Informatica Network for product resources such as documentation, how-to articles, best practices, and PAMs.

To access the Knowledge Base, visit https://kb.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

To get the latest documentation for your product, browse the Informatica Knowledge Base at https://kb.informatica.com/_layouts/ProductDocumentation/Page/ProductDocumentSearch.aspx.

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

Informatica Product Availability Matrixes

Product Availability Matrixes (PAMs) indicate the versions of operating systems, databases, and other types of data sources and targets that a product release supports. If you are an Informatica Network member, you can access 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. Developed from the real-world experience of hundreds of data management projects, Informatica Velocity represents the collective knowledge of our consultants who have worked with organizations from around the world to plan, develop, deploy, and maintain successful data management solutions.

If you are an Informatica Network member, you can access 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 augment, extend, or enhance your Informatica implementations. By leveraging any of the hundreds of solutions from Informatica developers and partners, you can improve your productivity and speed up time to implementation on your projects. You can access Informatica Marketplace at https://marketplace.informatica.com.

Informatica Global Customer Support

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

To find your local Informatica Global Customer Support telephone number, visit the Informatica website at the following link:
http://www.informatica.com/us/services-and-training/support-services/global-support-centers.

If you are an Informatica Network member, you can use Online Support at http://network.informatica.com.

Part I: Test Data Management 10.1.0

This part contains the following chapters:

CHAPTER 1

New Features and Enhancements (10.1.0)

This chapter includes the following topic:

New Features and Enhancements for Version 10.1.0

This section describes the new features and enhancements for version 10.1.0.

Data Generation Rules

To generate test data, you can create the following types of generation rules:

Advanced Generation Rule

An advanced generation rule is a combination of generation techniques that generates test data in target columns based on the configuration of one or more than one input port, variable port, and output port. You can generate test data for string, numeric, and date data types. When you create an advanced rule, create input ports, variable ports, and output ports to generate data as per the requirements.

Credit Card Generation Rule

Use a credit card generation rule to generate credit card numbers based on the issuing credit card network. You can select the credit card issuer type that you want to generate, specify the starting numbers for the card, and the distribution percentage.

You can generate test data for the following issuing networks: American Express, Discover, JCB, Mastercard, and Visa.

You can generate credit card numbers for string data types.

You cannot generate credit card numbers for numeric, date, and binary data types.

Conditional Generation Rule

A conditional generation rule is an ad hoc rule in which you can specify a conditional expression and a generation rule to generate test data. You can use the string, numeric, and date data types in a conditional generation rule.

You cannot apply conditions on effective dates and sequence generation rules.

For information about how to create and use the data generation rules, see the *Test Data Management 10.1.0 User Guide*.

Blaze Execution Engine

You can select an execution engine to run the mappings for a Hadoop connection.

Select a Blaze execution engine at the plan level or from the **Administrator** | **Preferences** view. TDM prioritizes the plan-level settings when you run the plan in the Hadoop environment.

If you select the Blaze execution engine, Blaze uses an internal workflow compiler to run the mapping. Use a Blaze engine to improve the speed and performance of the task.

For information about the Blaze execution engine, see the *Test Data Management 10.1.0 User Guide*.

Avro and Parquet Data Sources

Avro and Parquet are semi-structured data sources.

When you select an HDFS target connection, use Avro or Parquet resource formats to mask data and to move data in groups. You can select the resource format at the plan level.

For information about Avro and Parquet data sources, see the *Test Data Management 10.1.0 User Guide*.

Single Masking Rule Assignment

You can assign a single masking rule, a policy, or both to multiple columns of similar data type.

You can search for column names with similar data type. Assign a single rule or a policy to all the selected columns at once instead of individual assignment to each column.

For information about how to assign a single rule to multiple columns, see the *Test Data Management 10.1.0 User Guide*.

Advanced Search Text Filter

In Test Data Manager, you can enter a text instead of the wild characters in the filter criteria to search for columns.

Based on the type of filter operators, TDM searches and displays the results.

You can use the following types of filters: exact match, starts with, ends with, and contains

For information about how to use advanced search text filter, see the *Test Data Management 10.1.0 User Guide*.

Oracle Data Source Import

If you import data sources from an Oracle database, you can choose to ignore the tables that do not contain any data.

Test Data Manager excludes all the empty tables in the Oracle database and lists the tables that contain data.

For information about how to import Oracle data source without empty tables, see the *Test Data Management 10.1.0 User Guide*.

Data Coverage Analysis

You can create data coverage tasks to analyze test data in a data set.

Create a data coverage task to analyze values in a combination of two columns. You can add up to two filter columns to configure the scope of the analysis. Change the combinations of data to verify that a data set contains data required for specific test cases. You can update data across combinations of columns to add data where there might not be sufficient data density.

For information about how to perform data coverage analysis, see the *Test Data Management 10.1.0 User Guide*.

View and Edit Data in a Data Set

You can view and edit the data in a data set from Test Data Manager.

View and manage data from the data set page. You can add, delete, or update data in the data set. You can use SQL queries to perform these tasks or directly update the data in the grid. Add row-level tags to the data set to track changes or mark specific rows. You can perform a keyword search based on row-level tags.

For information about how to view and edit data in a data set, see the *Test Data Management 10.1.0 User Guide*.

Related Data Sets

You can create a related data set from an existing data set.

The related data set contains a subset of the data in the original data set. Create a related data set when you want to separately store and use a specific chunk of the data in a data set.

For information about how to create a related data set, see the *Test Data Management 10.1.0 User Guide*.

Subset Data Set Commands

To create a subset of a data set, you can use the following ilmcmd commands:

- ListPlans. Lists all the plans within a data set version.
- TDWPlanGenerate. Generates a workflow to create a subset of a data set.
- TDWPlanExecute. Runs a plan to create a subset of a data set.
- TDWPlanGenExe. Generates and runs a data set subset plan from a data set version.

For information about how to run the commands, see the *Test Data Management 10.1.0 User Guide*.

Pre Workflow and Post Workflow Plan Settings

In a TDM plan, you can select workflows that you want to run before and after the current workflow.

Configure the pre workflow parameter if you want to run a specific workflow before you run the workflow from the current plan. Configure the post workflow parameter if you want to run a specific workflow after you run the workflow from the current plan.

For information about how to configure and use pre workflow and post workflow parameters, see the *Test Data Management 10.1.0 User Guide*.

CHAPTER 2

Changes (10.1.0)

This chapter includes the following topic:

Changes for Version 10.1.0

This section describes the changes for version 10.1.0.

Configure Parent Child Ratio

Effective in 10.1.0, you can configure the number of child records individually for entities and for XML elements to generate test data.

You can select a relationship in an entity and specify the minimum and maximum number of child records to generate test data. To generate test data from XSD metadata, select a link between the XML elements and attributes and specify the minimum and maximum number of child records for each parent.

When you run a data generation plan, TDM prioritizes the parent child ratio configured at the individual relationship level over the plan-level configuration.

Previously, you could specify the minimum and maximum number of child records only at the plan level and at the global preferences level to generate test data.

For information about how to configure the parent child ratio, see the *Test Data Management 10.1.0 User Guide*.

TDM Repository Configuration

Effective in version 10.1.0, you cannot use the Custom database type to configure the TDM repository. You can use the Oracle, Microsoft SQL SERVER, or IBM DB2 database types to configure the TDM repository.

Previously, you could use the Custom database type to configure the TDM repository with custom database drivers. If you upgrade from a previous version with a TDM repository that uses custom JDBC drivers, you must delete the Test Data Manager Service and create the TDM repository with native drivers after you upgrade.

For information about how to configure the TDM repository, see the *Test Data Management 10.1.0 Installation Guide*.

Test Data Warehouse Service

Effective in version 10.1.0, test data warehouse is available as a managed service in the Informatica domain.

Create and administer the Test Data Warehouse Service from the Administrator tool to create the test data warehouse repository and repository content. Alternatively, use the *infacmd tdw* command line program to create and manage the Test Data Warehouse Service.

Previously, the Test Data Manager Service managed the test data warehouse. If you upgrade from a previous version, you must use the same test data warehouse repository and test data warehouse information when you create the Test Data Warehouse Service.

For information about how to create a Test Data Warehouse Service, see the *Test Data Management 10.1.0 Installation Guide*.

Roles and Privileges for Test Data Warehouse

Effective in version 10.1.0, you require Test Data Warehouse Service privileges to perform tasks in the test data warehouse. You can assign a Test Engineer role specific to test data warehouse.

Test Data Warehouse Service privileges include the following data set privileges:

- View data set
- Manage data set
- Reset data set
- View data in data set
- Manage data in data set

The Test Engineer role includes all the data set privileges.

Previously, you required specific Test Data Manager Service privileges to perform test data warehouse tasks.

Repository and Data Mart Tab

Effective in version 10.1.0, the **Administrator | Repository and Data Mart** tab is not available in Test Data Manager.

You configure a test data warehouse repository and test data warehouse when you create the Test Data Warehouse Service in Informatica Administrator.

Previously, you created a test data warehouse repository and test data warehouse from the **Administrator | Repository and Data Mart** tab in Test Data Manager.

infacmd Command Line Program

Effective in version 10.1.0, the *infacmd* command line program contains the following changes:

You run the *infacmd tdw* command line program to create and administer the Test Data Warehouse Service.

Previously, you used the *tdwcmd* command line program to create the test data warehouse repository and the repository content.

The *infacmd tdm* CreateService parameters are changed.

The following parameters are removed:

- customDriver

- httpProtocolType
- EnableTDWService

The following parameters are added:

- TDWServiceName
- hadoopDistDir
- hadoopKerbSPN
- hadoopKerbKeytab
- SSLProtocol

For information about how to run the *infacmd* command line program, see the *Test Data Management 10.1.0 Installation Guide*.

Data Set Reset Configuration

Effective in version 10.1.0, you can configure a reset operation to perform an insert in master tables instead of an upsert. You can use this configuration if all tables in the data set are master tables.

The reset operation might take less time to run. This is because the reset directly inserts all records from the data set instead of first checking if any records exist in the target.

Previously, you could not change the default reset configuration. The default reset configuration is to perform an upsert on master tables.

For information about how to configure a reset operation, see the *Test Data Management 10.1.0 User Guide*.

tdwquery

Effective in version 10.1.0, the tdwquery properties file contains the properties TDWSERVICENAME and TDW Service URL. This is the name of the Test Data Warehouse Service that manages the test data warehouse and the URL of the Test Data Warehouse Service.

Previously, you entered the name of the Test Data Manager Service and the Test Data Manager Service URL.

If you upgrade from a previous version, update the `userConfig.ilm` properties file with the name and URL of the Test Data Warehouse Service before you use tdwquery in 10.1.0.

For information about how to configure the `userConfig.ilm` file and use tdwquery, see the *Test Data Management 10.1 .0User Guide*.

tdwcmd

Effective in 10.1.0, the CreateTDWREPO and CreateTDWMART commands are removed.

The following table lists the tdwcmd commands and parameters that are changed from TDM 9.7.0 to TDM 10.1.0:

TDM 9.7.0 and TDM 9.7.1	TDM 9.7.1 HotFix 1	TDM 10.1.0
TDMServiceName (tsn)	TDMServiceName (tsn)	TDWServiceName (tsn)
MartName (mn)	TestDataWarehouseName (twn)	TestDataWarehouseName (twn)

TDM 9.7.0 and TDM 9.7.1	TDM 9.7.1 HotFix 1	TDM 10.1.0
DataMart	TestDataWarehouse	TestDataWarehouse
MartName (mn)	TestDataWarehouseName (twn)	TestDataWarehouseName (twn)
VersionMartName (vmn)	TestDataWarehouseName (twn)	TestDataWarehouseName (twn)
TableMartName (tmn)	TestDataWarehouseName (twn)	TestDataWarehouseName (twn)

Effective in 10.1.0, the HostName, HTTPPort, HTTPSPort, and DomainName parameters connect to Test Data Warehouse hosts, ports, and domains respectively.

Previously, the HostName, HTTPPort, HTTPSPort, and DomainName parameters connected to Test Data Manager hosts, ports, and domains respectively.

For information on tdwcmd commands and parameters, see the *Test Data Management 10.1.0 User Guide*.

Templates for Data Subset Objects

Effective in version 10.1.0, you cannot create a template to add entities and groups. You can use and edit existing templates but you cannot create new templates.

Previously, you could create a template to add multiple entities and groups and use the template in a plan.

Part II: Test Data Management 9.7.1 HotFix 1

This part contains the following chapters:

CHAPTER 3

New Features and Enhancements (9.7.1 HotFix 1)

This chapter includes the following topic:

New Features and Enhancements for Version 9.7.1 HotFix 1

This section describes the new features and enhancements for version 9.7.1 HotFix 1.

Support for Cassandra and MongoDB Connections

You can use Cassandra and MongoDB connections to perform data domain discovery and data masking operations.

In Test Data Manager, use ODBC connections to connect to Cassandra and MongoDB. You can use value cascades for Cassandra and MongoDB data sources.

You cannot perform data subset and data generation operations. You cannot use auto cascades for Cassandra and MongoDB data sources.

For information about Cassandra and MongoDB data types, see the *Test Data Management 9.7.1 HotFix 1 User Guide*.

Bulk Load in Netezza Connections

You can use the bulk reader and bulk writer to read and write data from a Netezza source and target.

For information about how to configure the bulk reader properties and bulk writer properties for a Netezza connection, see the *Test Data Management 9.7.1 HotFix 1 Administrator Guide*.

CHAPTER 4

Changes (9.7.1 HotFix 1)

This chapter includes the following topic:

Changes for Version 9.7.1 HotFix 1

This section describes the changes for version 9.7.1 HotFix 1.

Change to Support in Version 9.7.1 HotFix 1

Effective in version 9.7.1 HotFix 1, TDM deferred support for Hadoop connections. Support will be resumed with version 10.1.0.

Test Data Warehouse

Effective in 9.7.1 HotFix 1, the naming convention used for the test data warehouse databases is changed.

The database that stores the source table metadata and the project metadata generated when you create a data set is called the test data warehouse repository. Previously, it was called the test data repository.

The database that stores the source data that you include in a data set is called the test data warehouse. Previously, it was called the test data mart.

tdwcmd

Effective in 9.7.1 HotFix 1, you must use the command line scripts with the updated commands and parameters.

The following tdwcmd commands and parameters are changed:

TDM 9.7.1	TDM 9.7.1 HotFix 1
CreateTDWMART	CreateTestDataWarehouse
MartName (mn)	TestDataWarehouseName (twn)
DataMart	TestDataWarehouse

TDM 9.7.1	TDM 9.7.1 HotFix 1
MartName (mn)	TestDataWarehouseName (twn)
VersionMartName (vmn)	TestDataWarehouseName (twn)
TableMartName (tmn)	TestDataWarehouseName (twn)

For information on tdwcmd commands and parameters, see the *Test Data Management 9.7.1 HotFix 1 Administrator Guide* and *Test Data Management 9.7.1 HotFix 1 User Guide*.

tdwquery

Effective in 9.7.1 HotFix 1, the tdwquery properties file contains a property *TESTDATAWAREHOUSENAME*. This is the name of the test data warehouse where you store the data set.

Previously, the property was called *MARTNAME*.

Update the `userConfig.ilm` properties file before you use tdwquery in 9.7.1 HotFix 1.

Part III: Test Data Management 9.7.1

This part contains the following chapters:

CHAPTER 5

New Features and Enhancements (9.7.1)

This chapter includes the following topic:

New Features and Enhancements for Version 9.7.1

This section describes the new features and enhancements for version 9.7.1.

XML Sources

You can perform data subset, data masking, and data generation operations on XML sources.

Import the source definition from the PowerCenter® repository. Apply subset criteria, data masking rules, or data generation rules to perform the required operation.

For information about how to perform TDM operations on XML sources, see the *Test Data Management 9.7.1 User Guide*.

Session and Workflow Logs

You can configure the workflow and session log properties when you create a plan.

Edit the **Log Settings** section when you configure a plan. Choose to create workflow logs by runs or by time stamp. Choose to create a log file in addition to binary logs for log events. Configure properties for session and partial session log files.

For information about how to configure session and workflow logs in a plan, see the *Test Data Management 9.7.1 User Guide*.

Naming Convention for Workflows, Mappings, and Sessions

You can customize the naming convention for workflows, mappings, and sessions that you create in TDM.

Create a naming format to standardize the names of workflows, mappings, and sessions. When you customize the naming format, all workflows, mappings, and sessions that TDM creates use the same format.

For information about how to customize the naming format, see the *Test Data Management 9.7.1 Administrator Guide*.

Import Data Domains From the Model Repository

You can import data domains from the Model repository into the TDM repository.

The TDM repository must contain data domains listed in profiles that you import from the Model repository. Import the data domains from the Model repository before you import the profile into TDM.

For information about how to import a data domain from Test Data Manager, see the *Test Data Management 9.7.1 User Guide*.

Seed Parameterization

You can enter the seed value as a parameter when you create a masking rule that requires repeatable output.

Add a variable for the seed value to the parameter file and enter the variable when you create the masking rule. If you require a different seed value, you can edit the seed value in the parameter file and then run the plan again. You do not need to edit the masking rule.

For information about how to use seed values in data masking rules, see the *Test Data Management 9.7.1 User Guide*.

Change Propagation

When you update or delete policies, data domains, masking rules, generation rules, and data subset components, a warning message appears with the list of affected objects that contain assignments.

You can either choose to cancel the updates or continue with the changes. You can download the list of affected objects and save the .csv file.

If you update a rule, you change the properties within the rule. The changes do not take effect in the plans that contain the rules. You must fix the changes in the plan, and generate and run the plan again. If you make changes in the name or type of the field of a custom rule, the assignments do not contain the changes. You must import the mapplet again.

For information about how to use seed values in data masking rules, see the *Test Data Management 9.7.1 User Guide*.

JDBC Connection

You can use the connection type JDBC to create a JDBC target connection.

Use this connection when your target connection is a MySQL database and you want to disable and enable constraints in a plan. You need to select a target JDBC connection. You can use the JDBC connection type to create a connection with a MYSQL database.

Hadoop Data Sources

You can enable Kerberos authentication for Hadoop data sources.

You can use auto cascades and value cascades for Hadoop data sources.

You can apply shuffle, substitution, mapplet, and advanced masking rules to mask Hadoop data sources. You cannot use advanced email masking and unique substitution rules to mask Hadoop data.

For information about the data masking techniques, see the *Test Data Management 9.7.1 User Guide*.

Flat Files

You can configure the flat file properties such as file encoding, file format, optional quotes, and include headers.

If the source is relational, you can specify the file encoding type for the target flat file. If the source is a flat file, the default file encoding type of the target flat file is the same as that of the source flat file. You can choose the fixed width or delimited format for a flat file.

You can select an optional quote to separate columns in a .csv file based on the data. You can select single quote, double quotes, or no quotes. You can choose to include headers in the target flat file.

For information about the flat file properties, see the *Test Data Management 9.7.1 User Guide*.

Data Generation Rules

The data generation rules include the following enhancements:

Decimal Values in Percentages

In a data generation rule, you can enter decimal values up to two places in distribution percentages for data patterns, list of values, null values, and values that are not valid.

Generate Numbers in a String Column

In an ad hoc generation rule, you can generate numbers in a column with string data type. To generate numbers in the column, select Override Data Type to Numeric when you create an ad hoc generation rule.

Start Sequence from Last Value

In a sequence generation rule, you can choose to start the data sequence from the last output value. When you run a plan again, TDM generates the sequence of data from the previous output value. You can choose to start the sequence from the last value when you want to generate primary and unique keys.

For information about the data masking techniques, see the *Test Data Management 9.7.1 User Guide*.

Add Data Generation Rules to a Data Domain

Assign one or more data generation rules to a data domain. When you assign a generation rule to a data domain, the columns in the domain receive the data generation rule when you configure data generation.

When you add data generation rules to a data domain, you can enable one of the rules to be the default rule. The default rule is applicable to all columns in the data domain.

For information about how to add data generation rules to a data domain, see the *Test Data Management 9.7.1 User Guide*.

Column Profile

Create and run a column profile to determine the data source column characteristics, such as a list of values, data patterns, and value ranges.

When you create a column profile, you can select the columns on which you want to run a profile, and configure data sampling options.

After you configure and run a column profile, TDM analyzes column data and infers rules based on the data type. Column data types include all the inferred data types for each column in the profile results. Use the column profile results to generate test data.

Auto rules are the inferred rules that you assign after you run a column profile. When the status of inferred rule is Yes, you can assign the auto rule to target columns to generate test data.

For information about how to create a column profile, see the *Test Data Management 9.7.1 User Guide*.

Unique Key Constraints

Add a column or a set of columns as a constraint to define a unique key in a table.

A unique key column can contain null values. You can create more than one unique key in table. You can perform multiple data generation assignments when you filter the columns based on unique keys.

For information about how to create a unique key constraint, see the *Test Data Management 9.7.1 User Guide*.

Default Settings

Use default settings to generate test data for the columns that you have not assigned specific generation rules to.

When you create a project and import target metadata, the project contains the default generation rules for all data types that a TDM administrator configures. In a project, you can view the default data generation rules, default primary key generation rules, and default unique key generation rules for all data types.

For information about how to configure default settings to generate test data, see the *Test Data Management 9.7.1 User Guide*.

CHAPTER 6

Changes (9.7.1)

This chapter includes the following topic:

Changes for Version 9.7.1

This section describes the changes for version 9.7.1.

Import Profiles from the Model Repository

Effective in 9.7.1, you can import profiles with curation information directly from the Model repository in Test Data Manager.

Previously, you used Informatica Developer to export the profile as an XML file and then imported the XML file in Test Data Manager. You could not import the curation information.

For information about how to import a profile in Test Data Manager, see the *Test Data Management 9.7.1 User Guide*.

Rule Simulation With Pre- and Post-Processing of Rules

Effective in 9.7.1, you can perform rule simulation with pre- and post-processing of rules for columns with Double and Date data types also.

Previously, you could perform rule simulation with pre and post processing of rules for columns with the String data type.

Disable and Enable Constraints in Plans That Use Connection Parameters

Effective in 9.7.1, you can choose to disable constraints when you run a plan that uses connection parameters.

Previously, you could not disable constraints and indexes when you used connection parameters in a plan.

Default Settings

Effective in 9.7.1, you can view the default data generation rules, default primary key generation rules, and default unique key generation rules for all data types in the **Administrator** view.

Previously, you could only view the default data generation rules in the **Administrator** view.

Create a Primary Key Constraint

Effective in 9.7.1, you can create a primary key constraint from the **Tables | Constraints** view.

Previously, you created a logical primary key from the **Tables | Columns** view.

Part IV: Test Data Management 9.7.0

This part contains the following chapters:

CHAPTER 7

New Features and Enhancements (9.7.0)

This chapter includes the following topic:

New Features and Enhancements for Version 9.7.0

This section describes the new features and enhancements for version 9.7.0.

Test Data Warehouse

Configure a test data warehouse to store and manage multiple versions of test data in a central location. Store versions of test data from relational sources as versions of a data set.

A test data warehouse consists of a test data repository and a test data mart. The test data repository stores the metadata related to the data sets. The test data mart stores the data. You can use a single test data repository and a single test data mart for multiple machines in the domain that host a Test Data Manager Service. You can share and reuse test data across teams and projects. To return a test environment to a specific state, you can reset a data set version from the test data mart to a target. To search the test data repository, you can add tags to the data set versions.

For information about how to configure a test data repository and test data mart, see the *Test Data Management 9.7.0 Administrator Guide*.

For information about how to create a data set in the test data mart, see the *Test Data Management 9.7.0 User Guide*.

tdwquery

tdwquery is a command line program that you can use to analyze data in a data set version.

Store multiple versions of test data in the test data mart and then use tdwquery to run queries to analyze the data. You can use the Select and Select Distinct clauses with multiple functions and joins to filter and search for data when you run queries.

For information on how to configure and use tdwquery, see the *Test Data Management 9.7.0 User Guide*.

TDM Support for Hadoop

You can perform data movement, data masking, and data domain profiling operations on Big Data Edition Hadoop clusters. In Test Data Manager, you can create Hive and Hadoop Distributed File System (HDFS) connections and use them as source and target connections.

You can add groups and data masking components to a Hadoop plan. You can run data domain profiles on Hive and HDFS data sources to identify sensitive data.

You can create a Hadoop plan to move data from flat files, and relational databases such as Oracle and DB2, ODBC-Sybase, and ODBC-Microsoft SQL Server into Hive and HDFS targets. You can also create a Hadoop plan when you want to move data between Hadoop sources and targets.

In a Hadoop plan, you can perform data movement and data masking operations. You can run a plan audit report and plan detail report for a Hadoop plan. You cannot perform data subset and data generation operations in a Hadoop plan.

Connection Permissions

To restrict access to connections that you create in Test Data Manager, you can assign connection permissions.

Allow or deny access, or provide a certain level of access to users and user groups by assigning connection permissions. The level of tasks that you can perform depend on the permission assigned. Read, write, and execute are the three levels of permission. Assign and edit connection permissions from Test Data Manager.

For more information, see the *Test Data Management 9.7.0 Administrator Guide*.

CHAPTER 8

Changes (9.7.0)

This chapter includes the following topic:

Changes for Version 9.7.0

This section describes the changes for version 9.7.0.

Default Value in Reference Lookup Generation

Effective in 9.7.0, you can specify a default value that you want to generate in the target column instead of null values.

The default value can be one of the values present in the parent table if the tables are in a parent-child relationship.

Previously, you could not specify a default value and TDM generated null values in the target.

Override Entity Relations at Run Time

Effective in 9.7.0, you can choose to override constraints and maintain only referential integrity at run time for entities of any size. The changes apply to the data at run time and do not affect the entity.

Choose to maintain referential integrity only or maintain transactional integrity in the subset operation. When you select referential integrity only, the subset operation fetches minimal data to maintain referential integrity, but might break transactional integrity. When you select transactional integrity, the subset operation maintains relational and transactional integrity, but might fetch additional data.

Previously, you could choose to maintain only referential integrity in entities with up to 25 tables.