



Informatica® Data Archive  
6.4.3 HotFix 1

# Archiving Data from a Custom Application to the Informatica Data Vault

Informatica Data Archive Archiving Data from a Custom Application to the Informatica Data Vault  
6.4.3 HotFix 1

© Copyright Informatica LLC 2014, 2018

Publication Date: 2018-06-25

# Table of Contents

|   |          |
|---|----------|
| Abstract. . . . .   | iv       |
| <b>Chapter 1: Archiving Data from a Custom Application to the Informatica Data Vault. . . . .</b> | <b>5</b> |
| Overview. . . . .   | 6        |
| Airline Example. . . . .  | 6        |
| Archive to the Data Vault Process. . . . .  | 7        |
| Prerequisites. . . . .  | 7        |
| Step 1. Start Enterprise Data Manager. . . . .  | 8        |
| Step 2. Create an Application Version. . . . .  | 8        |
| Step 3. Import Metadata from the Source Database. . . . .   | 9        |
| Step 4. Define the Primary Key. . . . .   | 12       |
| Step 5. Define the Foreign Keys. . . . .  | 14       |
| Step 6. View the Association Tree. . . . .  | 17       |
| Step 7. Create Archive Entities . . . . .   | 19       |
| Create an Entity. . . . .   | 19       |
| Step 8. Define Table Types. . . . .   | 21       |
| Step 9. Create a Parameter. . . . .   | 22       |
| Step 10. Create a Business Rule. . . . .  | 24       |
| Create a Business Rule . . . . .  | 24       |
| Step 11. Update the Default Entity Steps. . . . .   | 26       |
| Step 12. Create a Default Index. . . . .  | 26       |
| Step 13. Create a Source Connection. . . . .  | 27       |
| Step 14. Create a Target Connection. . . . .  | 30       |
| Step 15. Create the Archive Project. . . . .  | 32       |
| Step 16. Review the Row Count Report. . . . .   | 35       |
| Step 17. Validate Data through the Data Discovery Portal. . . . .                                 | 36       |

# Abstract

You can use Data Archive to archive data from custom applications to the Informatica Data Vault. This article explains how to create and define entities for archiving. It also explains how to configure and schedule the archive project.

# CHAPTER 1

## Archiving Data from a Custom Application to the Informatica Data Vault

This chapter includes the following topics:

- [Overview, 6](#)
- [Airline Example, 6](#)
- [Archive to the Data Vault Process, 7](#)
- [Prerequisites, 7](#)
- [Step 1. Start Enterprise Data Manager, 8](#)
- [Step 2. Create an Application Version, 8](#)
- [Step 3. Import Metadata from the Source Database, 9](#)
- [Step 4. Define the Primary Key, 12](#)
- [Step 5. Define the Foreign Keys, 14](#)
- [Step 6. View the Association Tree, 17](#)
- [Step 7. Create Archive Entities , 19](#)
- [Step 8. Define Table Types, 21](#)
- [Step 9. Create a Parameter, 22](#)
- [Step 10. Create a Business Rule, 24](#)
- [Step 11. Update the Default Entity Steps, 26](#)
- [Step 12. Create a Default Index, 26](#)
- [Step 13. Create a Source Connection, 27](#)
- [Step 14. Create a Target Connection, 30](#)
- [Step 15. Create the Archive Project, 32](#)
- [Step 16. Review the Row Count Report, 35](#)
- [Step 17. Validate Data through the Data Discovery Portal, 36](#)

# Overview

You can archive data from active custom applications to the Informatica Data Vault. A custom application is an application that Informatica does not provide pre-configured entities for.

To archive an active custom application to Data Vault, you first import metadata from the source database with Enterprise Data Manager. Then you define constraints between the application tables and create archive entities. To configure the archive entities, you create parameters to define the data that you want to archive and business rules to identify the data that is not eligible for archiving.

After you create and configure an archive entity in Enterprise Data Manager, you can schedule the archive project to run in Data Archive.

This article uses an example to explain the process of archiving data to the Data Vault. If the application you want to archive is no longer active, see the Informatica How-to Library article:

[Retiring an Application to Data Vault with On-Premise Data Archive.](#)

## Airline Example

You use an application called Airline for airline ticket details. The application contains 11 tables and resides on an Oracle source database. To improve source database performance and reduce costs to support the application, you want to archive some of the application data.

Four of the Airline application tables contain transactional data related to ticket purchases and coupons. This data has a time line of usefulness in the production database. You want to archive application data older than one year. Because you might need the archived data in the future to create reports about ticket purchases or coupon use, you want to maintain access to it through the Data Discovery portal. You can use the Data Discovery portal to search and view data that is archived to Data Vault.

The following table lists the application tables and columns that contain transactional data that you want to archive on a regular basis:

| Table         | Columns  |
|---------------|--|
| TICKET        | <ul style="list-style-type: none"><li>- TICKET_NUMBER</li><li>- ISSUE_DATE</li><li>- CUST_ID</li><li>- PASSENGER_NAME</li><li>- LOCATION</li></ul> |
| TICKET_COUPON | <ul style="list-style-type: none"><li>- TICKET_NUMBER</li><li>- COUPON</li><li>- FROM_CITY</li><li>- TO_CITY</li><li>- IS_USED</li></ul>           |

| Table              | Columns  |
|--------------------|--|
| TICKET_SAVE        | <ul style="list-style-type: none"> <li>- TICKET_NUMBER</li> <li>- ISSUE_DATE</li> <li>- CUST_ID</li> <li>- PASSENGER_NAME</li> </ul>           |
| TICKET_COUPON_SAVE | <ul style="list-style-type: none"> <li>- TICKET_NUMBER</li> <li>- COUPON</li> <li>- FROM_CITY</li> <li>- TO_CITY</li> <li>- IS_USED</li> </ul> |

## Archive to the Data Vault Process

To archive to the Data Vault, you perform the following tasks through the Enterprise Data Manager and Data Archive user interfaces:

1. Start the Enterprise Data Manager.
2. Create an application version.
3. Import metadata from the source database.
4. Define the primary key.
5. Define the foreign keys.
6. View the association tree.
7. Create the archive entities.
8. Define entity table types.
9. Create parameters and modify the entity WHERE clause.
10. Create business rules.
11. Update the default entity steps.
12. Create a default index.
13. Create a source connection in Data Archive.
14. Create a target connection in Data Archive.
15. Create and schedule the archive project.
16. Review the row count report.
17. Verify archive data in Data Vault.

## Prerequisites

Before you begin, verify the following prerequisites:

1. Create database users and assign privileges. For information about required database users and privileges, see Chapter 3 of the *Informatica Data Archive Administrator Guide*.

2. Configure properties in the `conf.properties` file. For information about the `conf.properties` file, see Chapter 2 of the *Informatica Data Archive Administrator Guide*.
3. Start the ILM application server. For information about starting the ILM application server, see Chapter 1 of the *Informatica Data Archive Administrator Guide*.
4. Configure system profile properties on the **Configuration Settings** page in Data Archive. For information about system profile properties, see Chapter 2 of the *Informatica Data Archive Administrator Guide*.

## Step 1. Start Enterprise Data Manager

Open Enterprise Data Manager from Data Archive.

1. In Data Archive, select **Accelerators > Enterprise Data Manager**.  
The **Save As** window opens.
2. Save `RunEdm.jnlp` on the local machine.
3. To open Enterprise Data Manager, run `RunEdm.jnlp`.

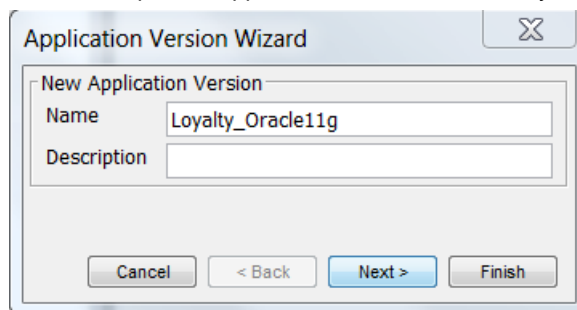
**Note:** You must have Java 7 installed on the machine where you run Enterprise Data Manager.

## Step 2. Create an Application Version

Create an application version for the application that you plan to archive.

1. In Enterprise Data Manager, right-click **Custom Apps Only** on the **Explorer** pane.
2. Select **New Application Version**.
3. Enter the application version name.

In this example, the application version name is `Loyalty_Oracle11g`.

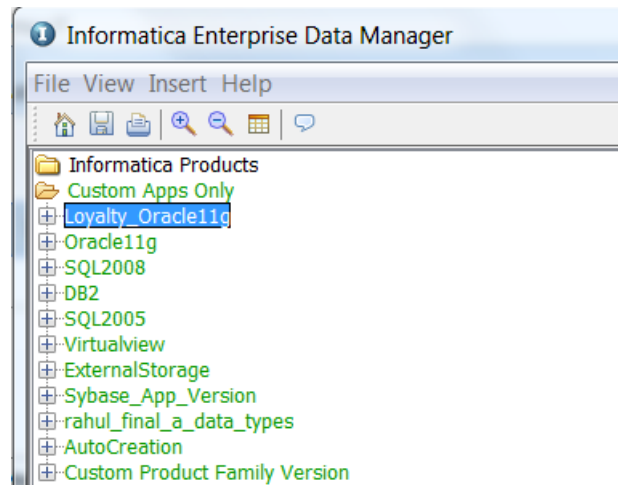


4. Click **Finish**.



5. Click the **Save** icon.

The Loyalty\_Oracle11g application version appears below the Custom Apps Only application.



## Step 3. Import Metadata from the Source Database

To import table metadata for the application, first define a connection to the source database and then specify the application module or schema to import.

When you import metadata, you import a list of tables and other metadata from the source database.

1. In Enterprise Data Manager, expand the application **Custom Apps Only** in the **Explorer** pane.

The list of application versions appear.

2. Click the application version that you created.

3. Click **File > Import Metadata from Database**.

The **Connect to Import Metadata from Database** window appears.

4. In the **Connect to Import Metadata from Database** window, create a connection to the Airline source database. Configure the following parameters to create a connection:

### Database Type

The type of database that you want to connect to.

### Database Host

The name of the machine that hosts the source data.

### Database Port

The port number that the source database uses.

### Service Name

Unique identifier or system identifier for the source database server.

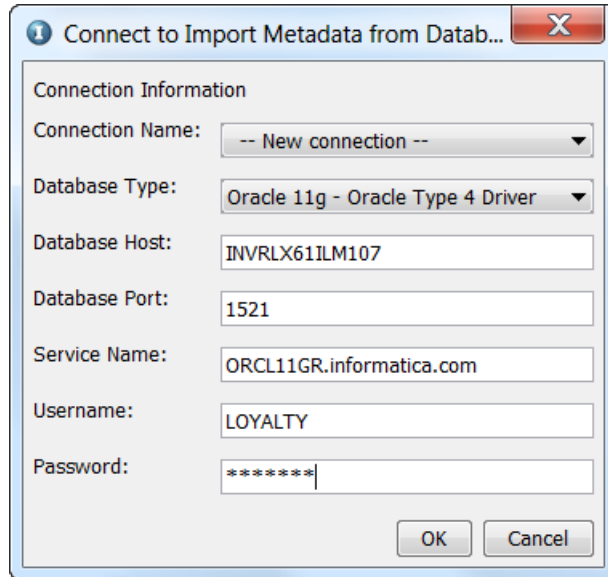
### User Name

The name of the user that connects to the source database. The user is the owner of the application module and tables that you want to import the metadata for.

For Oracle, the user name is the same as the name of the application module.

### Password

The password for the database user name.



A dialog box titled "Connect to Import Metadata from Datab..." with a close button (X) in the top right corner. The dialog contains the following fields and controls:

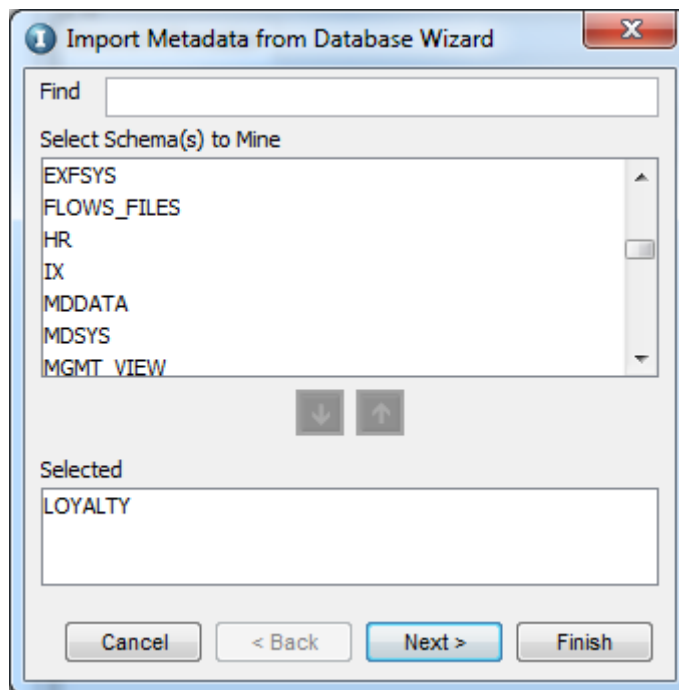
- Connection Information** (Section Header)
- Connection Name:** A dropdown menu showing "-- New connection --".
- Database Type:** A dropdown menu showing "Oracle 11g - Oracle Type 4 Driver".
- Database Host:** A text field containing "INVRLX61ILM107".
- Database Port:** A text field containing "1521".
- Service Name:** A text field containing "ORCL11GR.informatica.com".
- Username:** A text field containing "LOYALTY".
- Password:** A text field containing "\*\*\*\*\*".
- Buttons:** "OK" and "Cancel" buttons at the bottom right.

5. Click **OK**.

The **Import Metadata from Database Wizard** opens.

6. Select the schemas that you want to import metadata from. Double-click a schema to move it to the **Selected** box. Alternatively, select the schema and use the down arrow.

In this example, the airline application contains a schema called LOYALTY that includes the tables you want to archive.



A wizard dialog box titled "Import Metadata from Database Wizard" with a close button (X) in the top right corner. The wizard contains the following elements:

- Find:** A text field for searching.
- Select Schema(s) to Mine:** A list box containing the following schemas: EXFSYS, FLOWS\_FILES, HR, IX, MDDATA, MDSYS, and MGMT\_VIEW.
- Buttons:** Two arrows (down and up) between the list boxes.
- Selected:** A text field containing "LOYALTY".
- Buttons:** "Cancel", "< Back", "Next >", and "Finish" buttons at the bottom.

7. Click **Next**.

8. Choose one of the following options:

- **Submit Import Metadata as a Background Job.**

Runs the job in the background. If you run the job in the background, you can continue to perform other tasks.

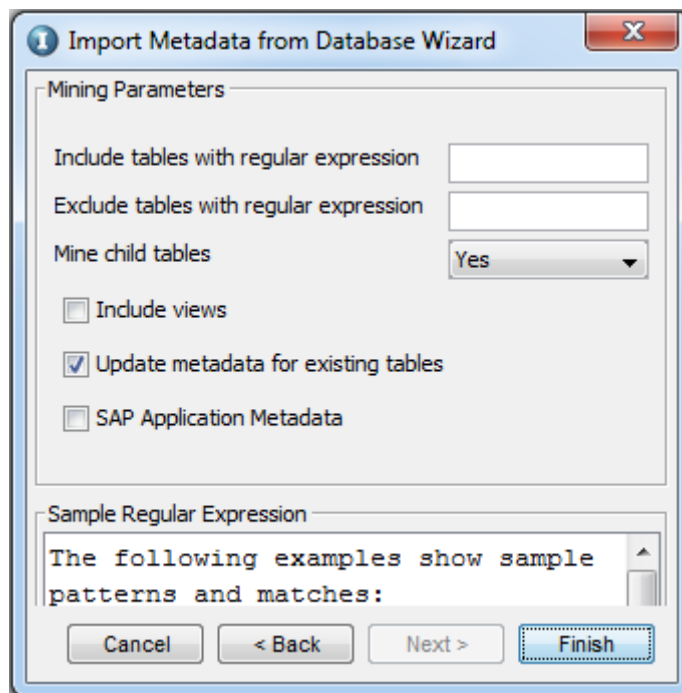
- **Continue Import Metadata through EDM.**

Runs the job in the foreground. If you run the job in the foreground, you must wait until the job completes to perform another task. Additionally, if you run the job in the foreground and you have a large volume of metadata to import, you might receive memory errors. You might want to run the job in the foreground if you have a low volume of metadata to import.

In this example, you submit the metadata as a background job so that you can continue to perform other tasks.

9. Click **Next**.

10. Select **Update metadata for existing tables**.



11. Click **Finish**.

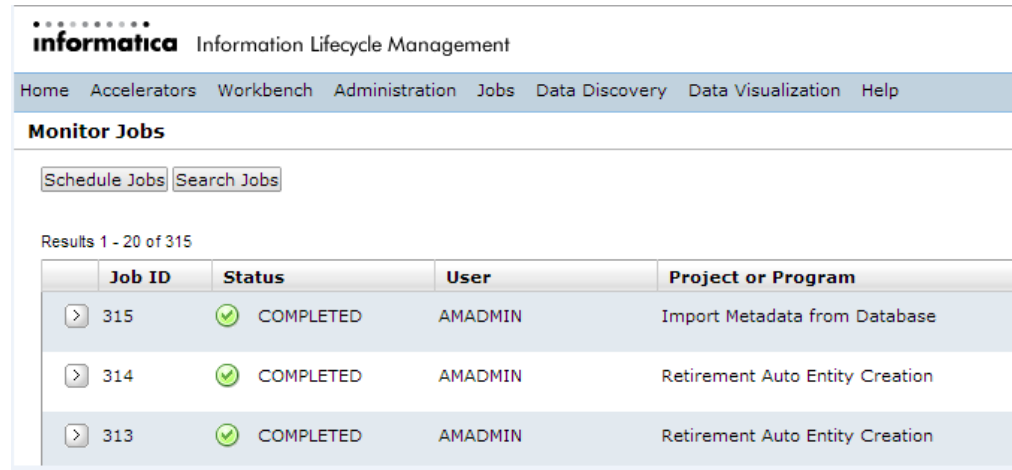
You receive a message that the wizard started the Import Metadata from Database job. You can use the job ID in the message to monitor the job status in Data Archive.

12. Click **OK** to close the message window.

13. To view the status of the job, go to the Data Archive user interface.

14. Click **Jobs > Monitor Jobs**.

When the job is complete, the job status is COMPLETED. You can expand the job icon on the left to view job details.



informatics Information Lifecycle Management

Home Accelerators Workbench Administration Jobs Data Discovery Data Visualization Help

**Monitor Jobs**

[Schedule Jobs](#) [Search Jobs](#)

Results 1 - 20 of 315

|   | Job ID | Status      | User    | Project or Program              |
|---|--------|-------------|---------|---------------------------------|
| > | 315    | ✓ COMPLETED | AMADMIN | Import Metadata from Database   |
| > | 314    | ✓ COMPLETED | AMADMIN | Retirement Auto Entity Creation |
| > | 313    | ✓ COMPLETED | AMADMIN | Retirement Auto Entity Creation |

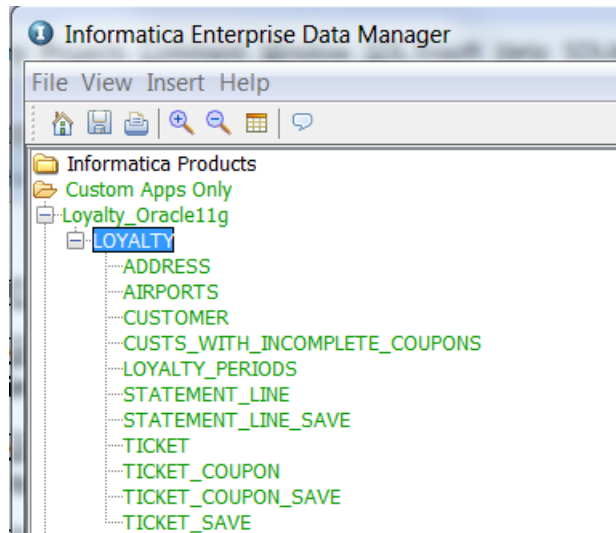
## Step 4. Define the Primary Key

To create a primary constraint, define a primary key on the table that you want to designate as the parent table.

After you import metadata from the source, determine which transactions you want to archive. A typical archive project contains three to five entities and the entity tables represent about 10% of the total application tables. Before you create a primary constraint, you must identify the parent tables of the transactions that you want to archive. Then define a primary key on the parent tables.

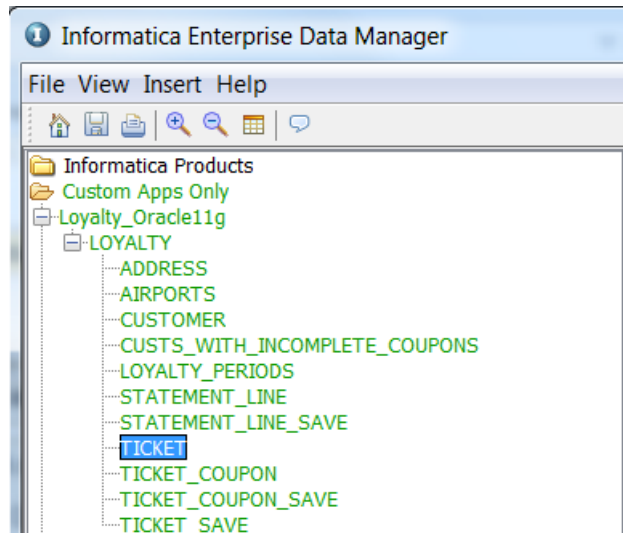
1. Click **View > Refresh**.
2. Click **View > Constraints**.
3. Expand the application version that contains the parent table.

4. To view the tables, expand the schema or application module.



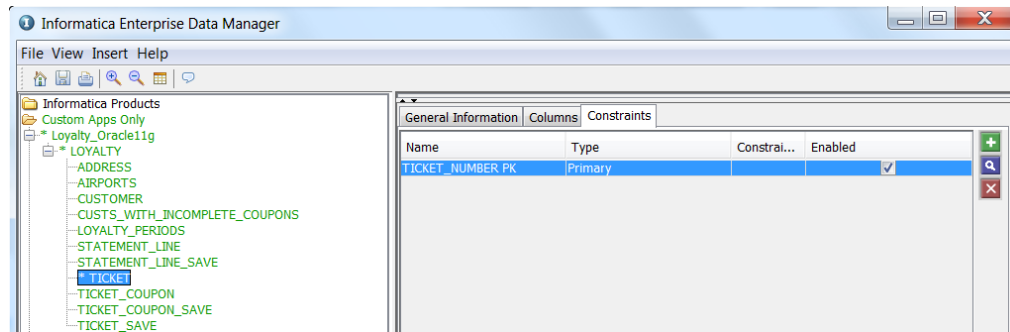
5. Select the parent table.

In this example, the parent table is TICKET.

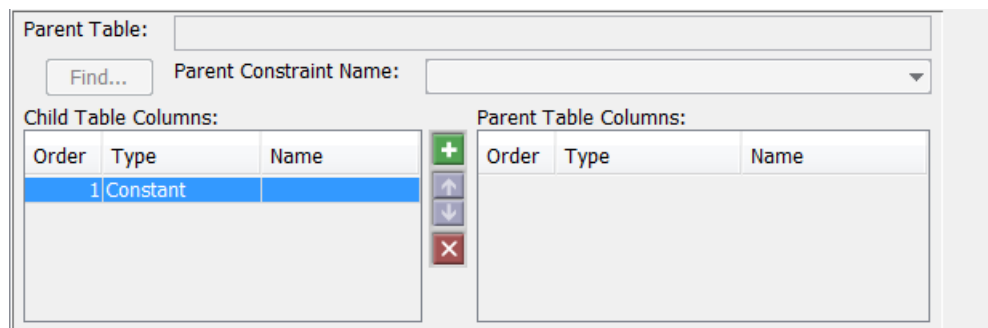


6. In the **Canvas** pane, click the **Constraints** tab.
7. Click the **Add** icon.  
A new row appears in the constraints table.
8. Double-click the **Name** column and enter a name for the constraint.  
In this example, the constraint is named TICKET\_NUMBER PK.
9. Click the **Type** column and select **Primary**.

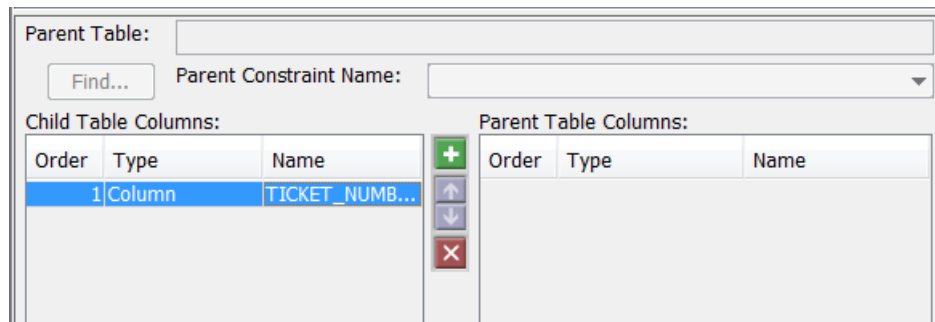
10. Click **Enabled**.



11. In the **Details** pane, click the **Add** icon.  
A row appears in the **Child Table Columns** table.



12. Click the **Type** column and select column as the type of primary key.
13. Click the **Name** column and select the column to use as the primary key.  
In this example, the column is TICKET\_NUMBER.



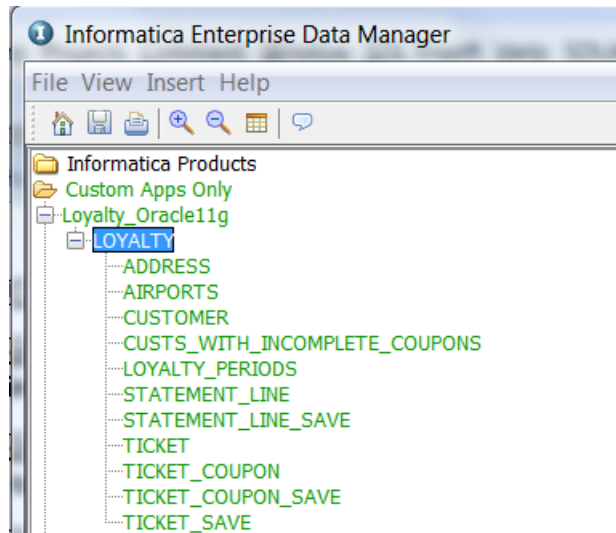
14. Click **Save**.

## Step 5. Define the Foreign Keys

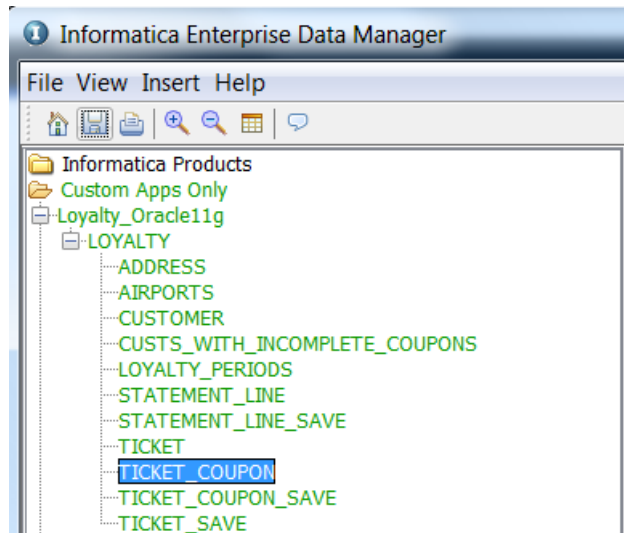
To create referential constraints, define foreign keys on the tables that you want to designate as child tables.

1. Click **View > Constraints**.
2. Expand the application version that contains the child table.

3. To view the tables, expand the application module.

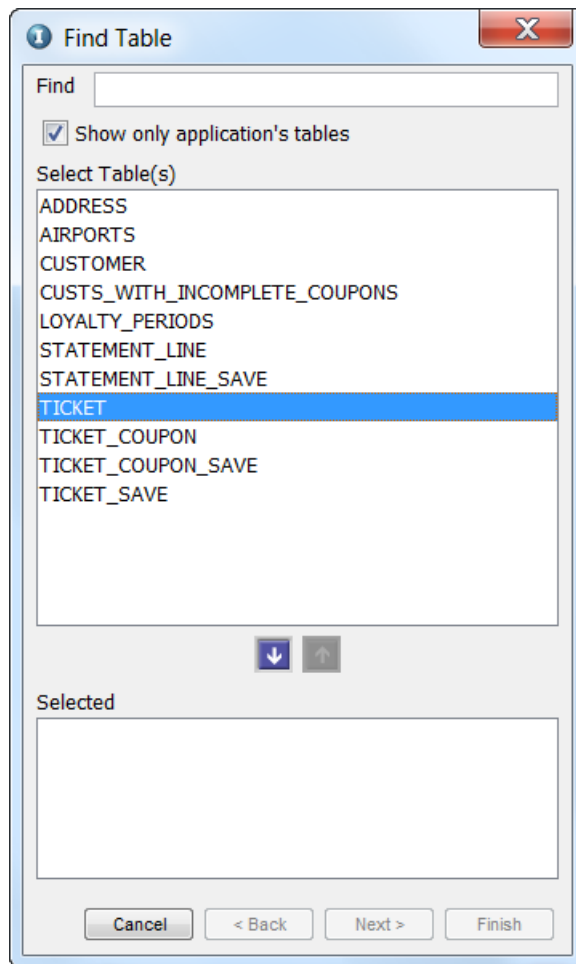


4. Select the child table.  
In this example, you select the table TICKET\_COUPON.

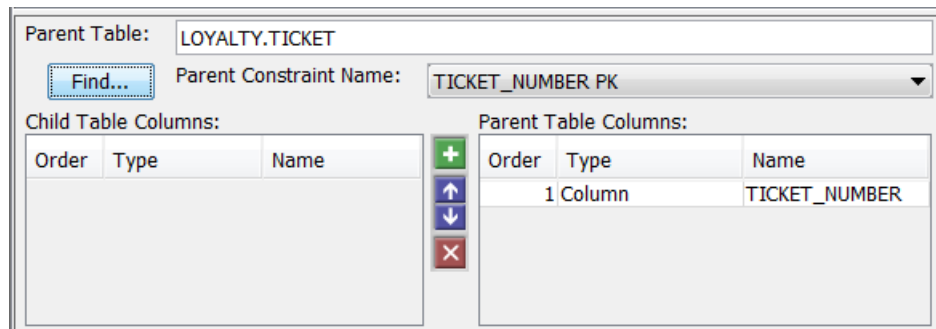


5. In the **Canvas** pane, click the **Constraints** tab.
6. Click the **Add** icon.  
A row appears in the constraints table.
7. Double-click the **Name** column and enter a name for the constraint.  
In this example, the constraint is named "TICKET\_NUMBER FK."
8. Click the **Type** column and select **Referential**.
9. Click **Enabled**.
10. In the **Details** pane, click **Find** to select the parent table with the primary key that you created.

The **Find Table** window appears.



11. To move the table to the Selected box, select the parent table and click the **Down** arrow.
  12. Click **Finish**.
- The selected parent table appears in the **Details** pane.
13. Next to **Parent Constraint Name**, select the primary constraint that you created.



14. Click the **Add** icon.



A row appears in the Child Table Columns table.

Parent Table:

Find... Parent Constraint Name:

Child Table Columns:

| Order | Type     | Name |
|-------|----------|------|
| 1     | Constant |      |

Parent Table Columns:

| Order | Type   | Name          |
|-------|--------|---------------|
| 1     | Column | TICKET_NUMBER |

15. Click the **Type** column and select column as the type of foreign key.
16. Click the **Name** column and select the column to use as the foreign key.

Parent Table:

Find... Parent Constraint Name:

Child Table Columns:

| Order | Type   | Name           |
|-------|--------|----------------|
| 1     | Column | TICKET_NUMB... |

Parent Table Columns:

| Order | Type   | Name          |
|-------|--------|---------------|
| 1     | Column | TICKET_NUMBER |

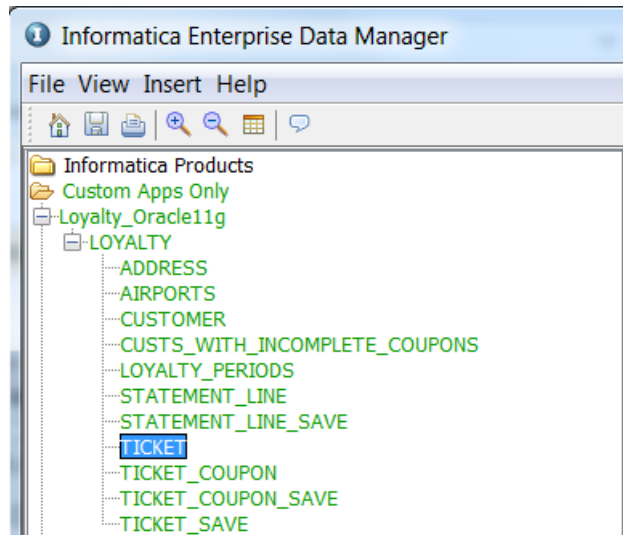
17. Click **Save**.
18. Repeat the process for each child table that you want to include in the entity.  
In this example, you create referential constraints on the TICKET\_SAVE, TICKET\_COUPON, and TICKET\_COUPON\_SAVE tables.

## Step 6. View the Association Tree

After you create constraints, you can view a graphical representation of the table relationships.

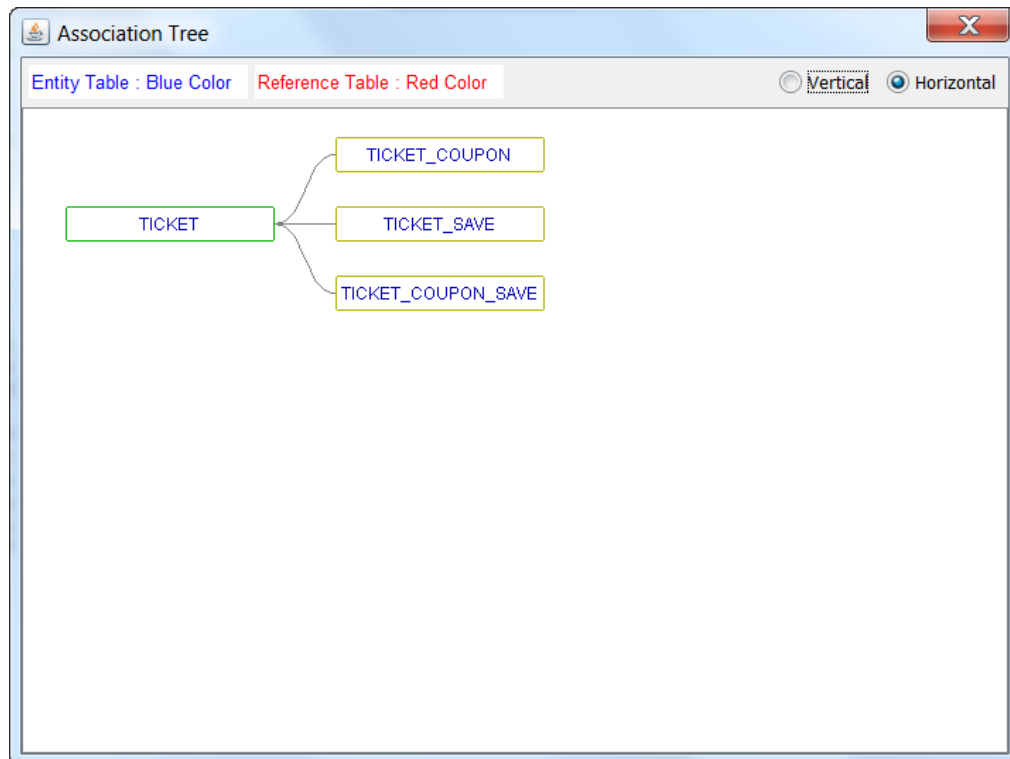
1. Click **View > Refresh**.
2. Click **View > Constraints**.

3. In the **Explorer** pane, select the parent table that you want to view table relationships for.



4. Click **View > Association Tree**.

The **Association Tree** window appears.



5. Verify that the constraints you created between the tables are correct.

## Step 7. Create Archive Entities

Before you can create and run an archive project in Data Archive, you must create an entity for archiving.

An archive entity consists of related tables that you want to archive together. Entities also include an interim table that the Enterprise Data Manager generates and names when you create the entity. For more information about entities, see the *Informatica Data Archive Enterprise Data Manager Guide*.

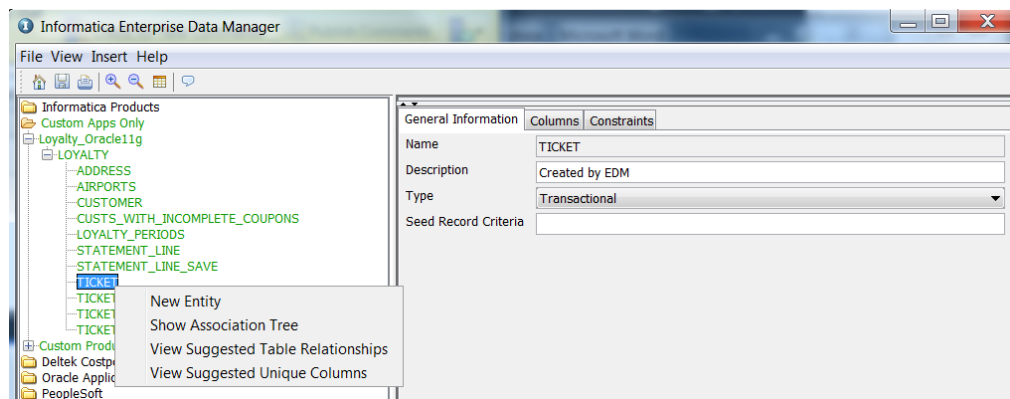
In this example, you create an entity from four of the tables you imported from the LOYALTY schema.

### Create an Entity

After you define constraints, create an entity that includes the tables you want to archive.

1. Click **View > Constraints**.
2. In the **Explorer** pane, select the application version that contains the entity tables.
3. Select and expand the application module that contains the entity tables.
4. Right-click the parent table that you created a primary constraint on.

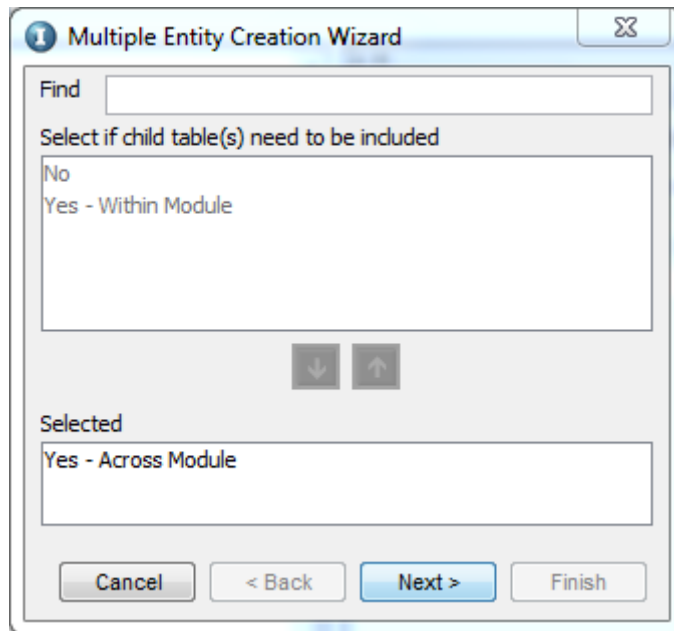
This table will become the entity driving table.



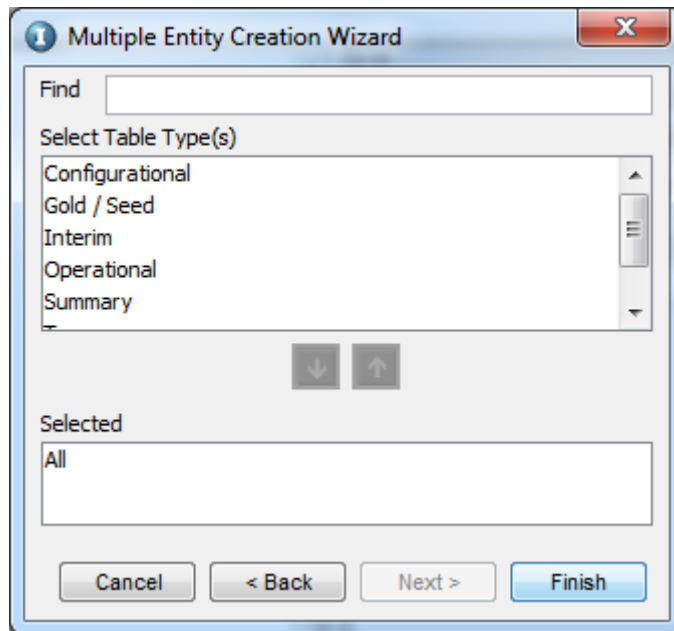
5. Select **New Entity**.

The **Multiple Entity Creation Wizard** appears.

6. To ensure that the child tables you created are added to the entity, select **Yes - Within Module** and click the **Down** arrow.

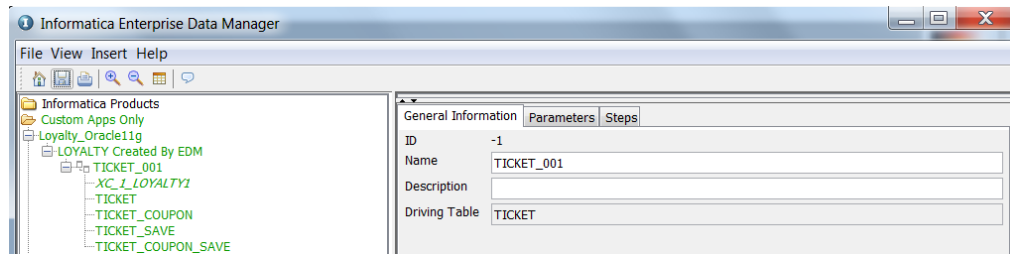


7. Click **Next**.
8. Under Tables Types, select **All** and click the **Down** arrow.



9. Click **Finish**.  
A message appears to inform you that the entity was successfully created.
10. Click **OK**.

11. To view the entity you created, expand the application version and module on the **Database Archive** tab.



12. Click **Save**.

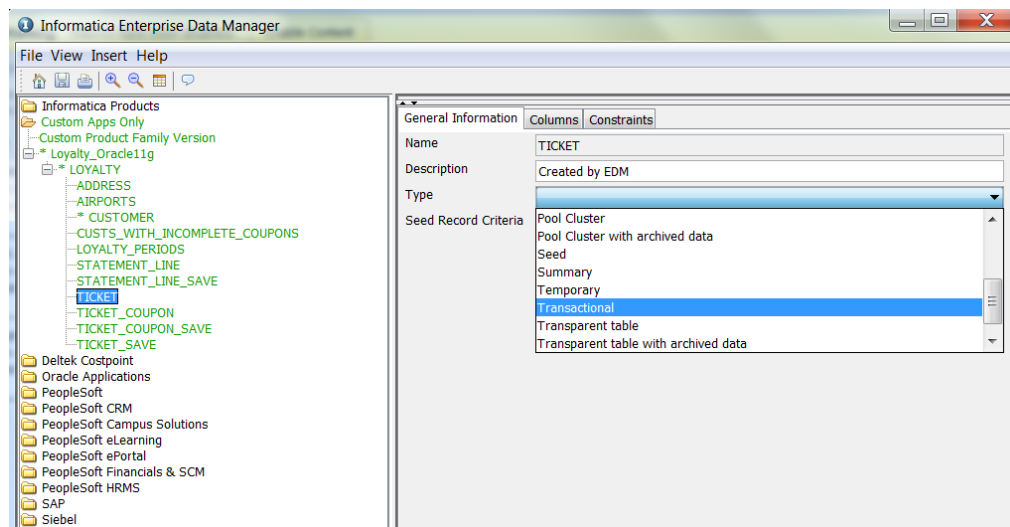
## Step 8. Define Table Types

After you create an archive entity, define the types of tables included in the entity. Table types determine how Data Archive processes the tables.

Someone familiar with the application data model must identify which tables are transactional tables. Transactional tables are tables that contain old or seldom-used rows that can be archived. Transactional tables often grow quickly because of typical day-to-day table inserts and updates.

In this example, you mark the TICKET table as a transactional table. The TICKET table contains many rows of data about ticket purchases. The data in the TICKET table has a time line of usefulness and you can safely archive the data when it is no longer needed.

1. Click **View > Constraints**.
2. Expand the application version that contains the table you want to define.
3. To view the tables, expand the schema or application module.
4. Select the table that you want to define.  
In this example, you want to mark the TICKET table as transactional.
5. In the **Canvas** pane, click the **General Information** tab.
6. In the **Type** menu, select **Transactional**.



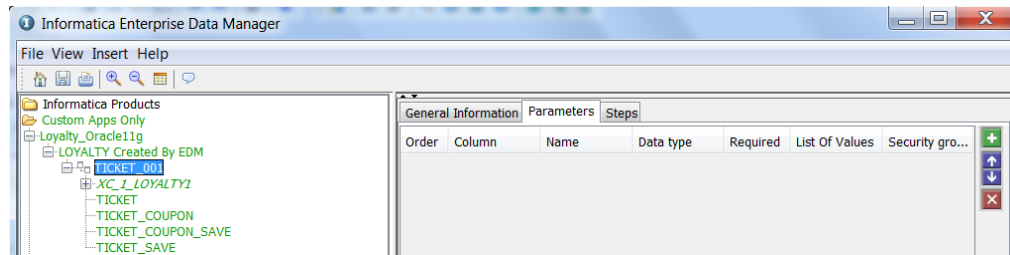
7. Click **Save**.

## Step 9. Create a Parameter

Create an archive parameter to define the data that you want to archive. When the archive job runs, parameters filter the data that meets the parameter criteria and the job creates a list called the archive candidates list. Parameters are typically date-driven.

In this example, the parameter is based on the issue date of the ticket. You specify the value for the parameter in the Data Archive user interface before you run an archive project.

1. On the **Database Archive** tab, expand the application version and module that contains the entity you created.
2. Select the entity that you want to create a parameter for.
3. In the **Canvas** pane, click the **Parameters** tab.



4. In the **Canvas** pane, click the **Add** icon.  
A row appears in the parameters table.
5. Define the following properties:

### Column

Column used to define the parameter. Add "p\_" to the beginning of the column name.  
In this example, the column is ISSUE\_DATE, so you enter "p\_issue\_date" in the **Column** field.

### Name

Name for the parameter.  
In this example, the parameter name is ISSUE\_DATE.

### Data type

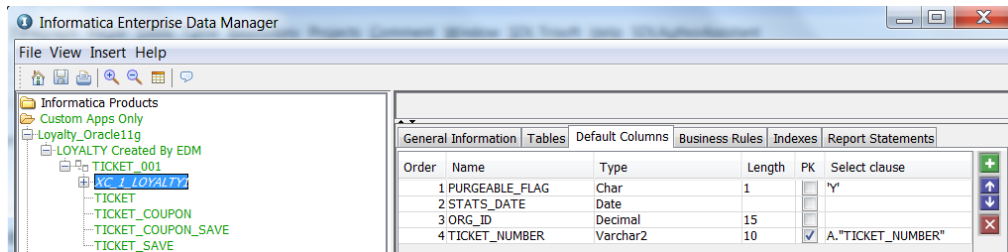
Type of column data.

### Required

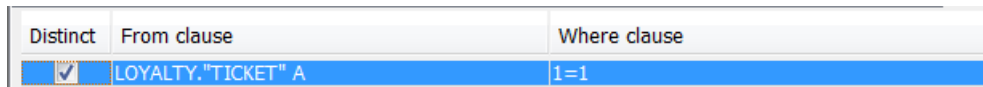
Indicates whether each archive job run requires a parameter column value.

6. Click **Save**.
7. In the **Explorer** pane, select the entity interim table.
8. In the **Canvas** pane, click the **Default Columns** tab.

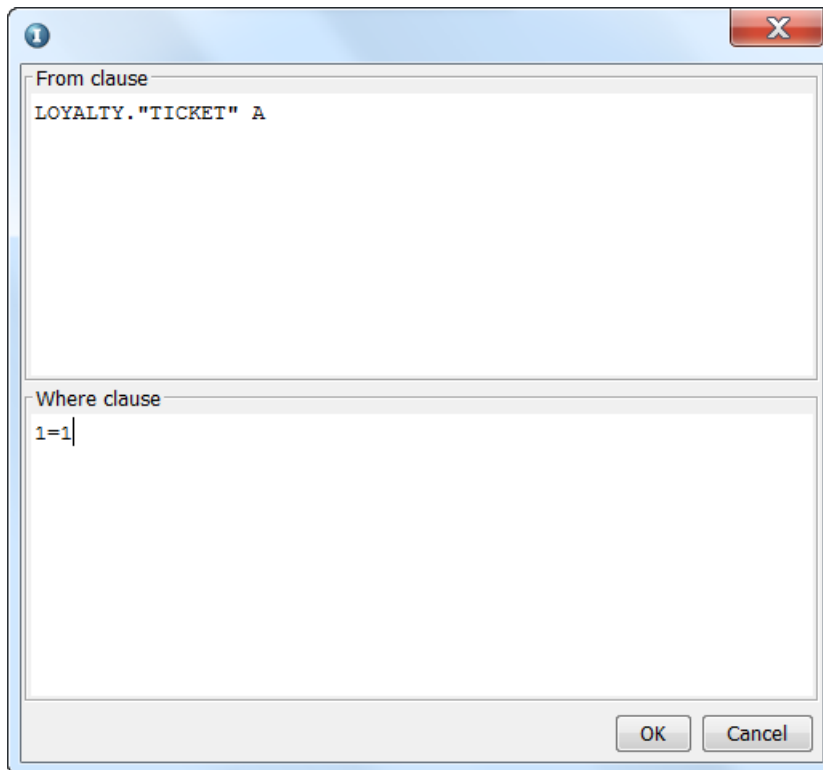
The **Default Columns** tab lists the default columns in the interim table.



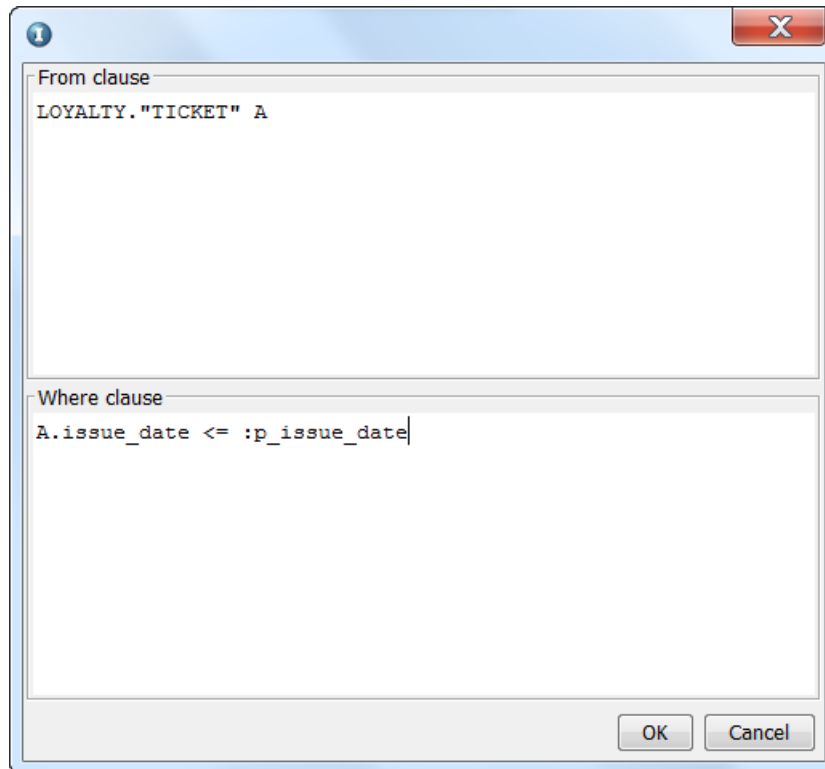
- To edit the entity WHERE clause, click the **Where clause** column in the **Details** pane.



A pop-up window appears.



10. In the Where clause box, replace "1=1" with the syntax to implement the parameter you created.



The screenshot shows a dialog box with two main sections: 'From clause' and 'Where clause'. The 'From clause' section contains the text 'LOYALTY.\"TICKET\" A'. The 'Where clause' section contains the text 'A.issue\_date <= :p\_issue\_date'. The dialog box has a title bar with a question mark icon and a close button (X). At the bottom right, there are 'OK' and 'Cancel' buttons.

11. Click **Save**.

## Step 10. Create a Business Rule

After you create parameters, create business rules to identify the data that is not eligible to be archived and purged. Business rules are SQL statements that check settings, status codes, and column values to determine whether a row should stay in the production system.

For example, some data might pass the issue date parameter you created, but the customer did not use their flight coupon when they purchased the ticket. You decide that you do not want to archive transactions where the coupon has not been used. To identify this data, you create a business rule called COUPON\_NOT\_USED.

Note that the rule identifies transactions where the coupon is not used instead of transactions where the coupon is used. Business rules filter out transactions that cannot be archived and purged, so they are written to identify transactions that fail the rule. For more information about business rules, see the Informatica How-to Library article: [How to Create Business Rules to Archive and Purge Transactional Data](#).

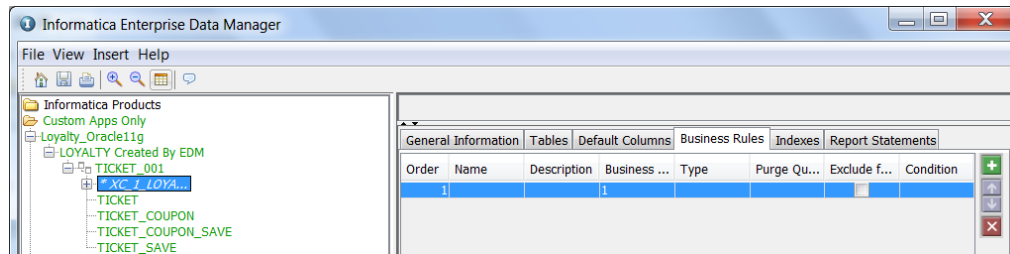
### Create a Business Rule

Create business rules to identify the data that is not eligible to be archived and purged.

1. In the **Explorer** pane, select the entity interim table.
2. Click the **Business Rules** tab.
3. Click the **Add** icon.



A row appears in the business rules table.



- Define the following properties:

#### Name

Name for the business rule.

In this example, the business rule name is COUPON\_NOT\_USED.

#### Description

Description for the parameter.

In this example, the description is "Flight coupon is not used."

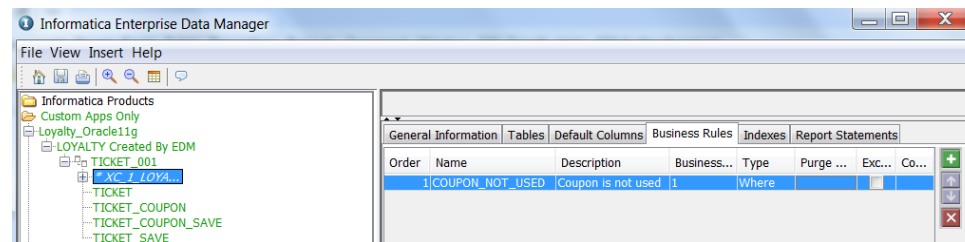
#### Business Rule Order

Order in which the archive job applies the business rule. If you have multiple business rules, you can designate the order in which the archive job applies them.

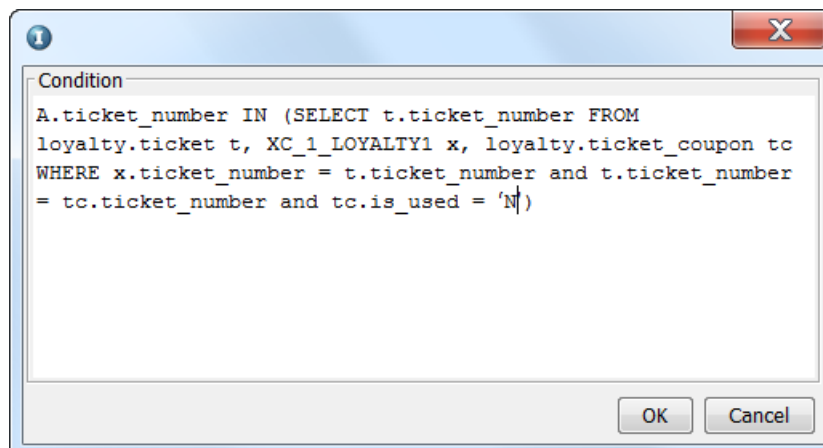
#### Type

Type of business rule.

In this example, the type is WHERE clause.



- Click the empty box under the **Condition** column.  
A pop-up window appears.
- Enter the syntax to implement the business rule you want to create.

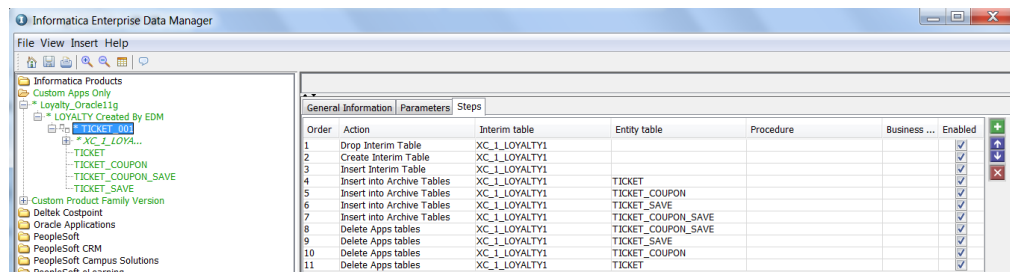


7. Click **OK**.
8. Click **Save**.

## Step 11. Update the Default Entity Steps

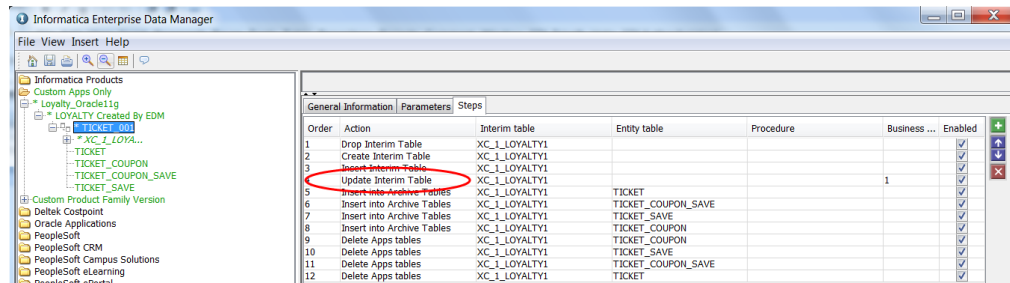
After you create a business rule, update the default entity steps to insert a step in the archive job that runs the business rule.

1. In the **Explorer** pane, select the entity.
2. Click the **Steps** tab.



3. In the **Details** pane, click **Update default steps**.

A row appears in the steps table. In preparation to run the archive job, this step updates the entity interim table with the transactions that pass the business rule you created.



4. Click **Save**.

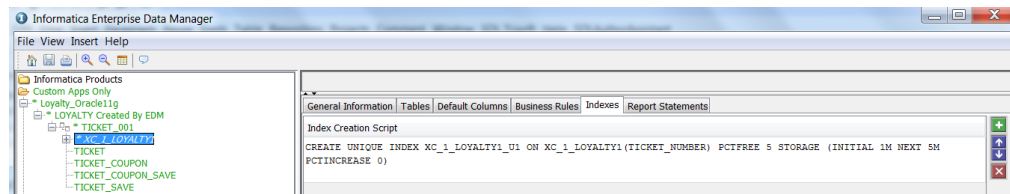
## Step 12. Create a Default Index

Before you run the archive job, create an index on the interim table to help the archive job run efficiently.

1. In the **Explorer** pane, select the entity interim table.
2. In the **Canvas** pane, click the **Indexes** tab.

3. In the **Details** pane, click the **Insert default index** button.

Enterprise Data Manager generates an index creation script that creates the index when you run the archive job.



## Step 13. Create a Source Connection

Create a connection to the source database that stores the application data that you want to archive.

When you create a source connection, you define general properties and database properties. The ILM engine job uses the attributes that you specify to connect to the source database. Use the same source connection information you entered to import the metadata in Enterprise Data Manager.

1. Go to the Data Archive user interface.
2. Choose one of the following options:
  - Click **Administration > New Source Connection**.
  - If you have an existing source connection, click **Administrator > Manage Connections**. Click the **Source** tab. Click the copy icon for the source connection you want to copy.

The **Create or Edit an Archive Source** page appears.

3. Define the following general properties:

### Connection Name

Name of the source connection. When you create an archive project and you assign a source, you select the connection name that you define here.

### Connection Description

Long text description for the connection name. When you manage connections, you can use this field to filter the list of connections.

### Connection Type

Database connection type that determines how you connect to the source database.

### Application Version

Business application and application versions that the database connection type supports. The possible values depend on the database connection type.

4. Define the following database properties:

### Host

IP address of the source application database server. When the source application server is in an Oracle Real Application Cluster, enter the IP address of node1 and the IP address of node 2, separated by a comma.

**Port**

Port of the source application database server. When the source application server is in an Oracle RAC, enter the port number of node 1 and port number of node 2, separated by a comma.

**Service Name**

Unique identifier or system identifier for the source application database server.

When you define an Oracle RAC source database server, enter the service name.

**Admin Schema Name**

Default administration database user for the source database server, such as SYSTEM.

The administration database user has DBA rights to the database, including the ability to run DDL and access system-level objects.

**Admin Login Name**

Admin Login Name is the same as the schema name.

**Password**

Password for the administration login name.

**Apps Schema Name**

Database user that owns the tables with the data that you want archive. For example, APPS for Oracle applications.

**Application Login Name**

Login name that connects to the source database that contains the data that you want to archive. This user does not require special permissions as this value is used for the database connection. You can provide any user name, such as a database connection user or a read-only user.

**Password**

Password for the administration login name.

**Staging Schema Name**

Staging database user for the source database.

**Staging Login Name**

Login name for the staging database.

**Password**

Password for the staging login name.

**Staging Tablespace**

Tablespace in the staging database. Enter the name of the tablespace you created when you installed Data Archive. For example, AM\_DATA.

**Note:** You do not need to enter or update values for the other parameters on the **Create or Edit an Archive Source** page.

\*\*\*\*\*  
**informatica** Information Lifecycle Management

Administrator Log Off | Help

Home Accelerators Workbench Administration Jobs Data Discovery Data Visualization Help

Create or Edit an Archive Source

Current Location: Administration > Create or Edit an Archive Source

★ Mandatory fields

★ Connection Name

Airline Loyalty Source

Description

Connection to the Loyalty source

★ Connection Type

ORACLE 11g

★ Application Version

Custom Apps Only Custom Product Family Version

★ Host

ec2-54-214-72-66.us-we

★ Port

1521

★ Service Name

ORA11G

★ Admin Schema Name

LOYALTY

★ Admin Login Name

loyalty

★ Password

\*\*\*\*\*

★ Confirm Password

\*\*\*\*\*

★ Apps Schema Name

LOYALTY

★ Application Login Name

loyalty

★ Password

\*\*\*\*\*

★ Confirm Password

\*\*\*\*\*

★ Staging Schema Name

LOYALTY

★ Staging Login Name

loyalty

★ Password

\*\*\*\*\*

★ Confirm Password

\*\*\*\*\*

★ Staging Tablespace

STAGE\_SRC\_DATA

Use Copy to Staging

☒

Use Row ID for Delete

☒

JDBC Fetch Size

1000

Use Row Id For File Gen

☒

Use Oracle Parallel DML For Delete

☐

Database Link to ILM Repository

5. Click **Save**.

The system validates the database connection. The **Manage Connections** page appears with the source connection listed on the **Source** tab.

\*\*\*\*\*  
**informatica** Information Lifecycle Management

Home Accelerators Workbench Administration Jobs Data Discovery Data Visualization Help

Manage Connections

Results 1-10 of 42

Source Target

View Source Connections: Select the filters and click on the filter to fetch the Source details.

| Connection Name                            | Description                      | Application Version                            |
|--|----------------------------------|--|
| <a href="#">Airline Loyalty Source</a>     | Connection to the Loyalty source | Custom Apps Only AIRLINE                       |
| <a href="#">ArchivingDigitalAssets_404</a> | ArchivingDigitalAssets           | Custom Apps Only Custom Product Family Version |
| <a href="#">Buspart</a>                    | Src_Oral11g_StgOn_Desc           | Custom Apps Only Custom Product Family Version |

## Step 14. Create a Target Connection

Create a connection to the Data Vault, the target destination for the application data that you want to archive.

When you create a target connection, you define general properties and database properties. The Data Vault Loader job uses the attributes that you specify to connect to the target.

1. Go to the Data Archive user interface.
2. Choose one of the following options:
  - Click **Administration > New Target Connection**.
  - If you have an existing target connection, click **Administrator > Manage Connections**. Click the **Target** tab. Click the copy icon for the target connection you want to copy.

The **Create or Edit an Archive Target** page appears.

3. Define the following general properties:

### Connection Name

Name of the target connection. When you create a retirement project and you assign a target, you select the connection name that you define here.

### Connection Description

Long text description for the connection name. When you manage connections, you can use this field to filter the list of connections.

### Connection Type

Database connection type that determines how you connect to the target database.

### Application Version

Business application and application versions that the database connection type supports. The possible values depend on the database connection type.

4. Define the following database properties:

### Staging Directory

Directory in which the Data Vault Loader job temporarily stores data as it completes the archive process. Enter the absolute path for the directory.  
The directory must be accessible to the ILM application server.

### Number of Rows Per File

Maximum number of rows that the Data Vault Loader stores in a file in the Data Vault. Default is 1 million rows.

### Data Vault Data Directory

Directory in which the Data Vault Loader job creates the archive folder. Enter the absolute path for the directory.

### Data Vault Archive Folder Name

Name of the archive folder in Data Vault where the application data is stored. The Data Vault archive folder corresponds to the database in the source.

### Data Vault Host

Host name or IP address of the machine that hosts the Data Vault Service.

### Data Vault Port

Port number used by the ssasql command line program and other clients such as the Data Vault SQL Tool and ODBC applications to connect to the Data Vault. Default is 8500.

### Data Vault Administration Port

Port number used by the Data Vault Agent and the Data Vault Administration Tool to connect to the Data Vault. Default is 8600.

### Data Vault User

Name of the administrator user account to connect to the Data Vault Service.

You can use the default administrator user account created during the Data Vault installation. The user name for the default administrator user account is *dba*.

### Data Vault User Password

Password for the administrator user account.

### Confirm Password

Verification of the password for the administrator user account.

**Note:** You do not need to enter or update values for the other parameters on the **Create or Edit an Archive Target** page.

The screenshot displays the 'Create or Edit an Archive Target' page in the Informatica Information Lifecycle Management interface. The page header includes the Informatica logo and navigation links. The main content area is titled 'Create or Edit an Archive Target' and contains a form with the following fields:

- Connection Name:** Airline Loyalty Target
- Connection Type:** Optimized File Archive
- Description:** Connection to Airline Loyalty application in Data Vault
- Application Version:** Custom Apps Only AIRLINE
- Staging Directory:** /data/Airline/STAGE
- Number of Rows Per File:** 1000000
- File Archive Data Directory:** /data/Airline/SCT
- File Archive Folder Name:** AirlineLoyalty
- File Archive Host:** INVRXLX62ILM86
- File Archive Port:** 8521
- File Archive Administration Port:** 8621
- File Archive User:** dba
- File Archive User Password:** (masked with asterisks)
- Confirm Password:** (masked with asterisks)
- Add On URL:** (empty)
- Maintain Imported Schema Name:** ☒
- Application Owner:** (empty)
- Application Owner email-id:** (empty)
- Archive Store Type:** (dropdown menu)

At the bottom right of the form, there are 'Save' and 'Cancel' buttons.

5. Click **Save**.

The system validates the database connection. The **Manage Connections** page appears with the target connection listed on the **Target** tab.

Results 1-10 of 160

Source Target

View Target Connections: Select the filters and click on the filter to fetch the Target details.

| Connection Name        | Description   | Application Version                            |
|------------------------|---|--|
| Airline Loyalty Target | Connection to Airline Loyalty application in Data Vault | Custom Apps Only AIRLINE                       |
| cas1                   |   | Custom Apps Only Custom Product Family Version |

## Step 15. Create the Archive Project

To archive data, you must create an archive project.

When you create an archive project, you specify whether you want to retain the data in the source database after archiving. You specify the source and target connections and the entities to include. Then, you configure the different phases in the archive process and schedule the job.

1. In Data Archive, click **Workbench > Manage Archive Projects**.  
The **Manage Archive Projects** page appears.
2. Click **New Archive Project**.  
The **Create or Edit an Archive Project** page appears.
3. Enter the name of the archive project in **Project Name**.
4. Optionally, enter a description.
5. Click **Action** and select one of the following options:
  - **Archive Only**. Data Archive extracts data from the source repository and loads it into the target repository.
  - **Purge Only**. Data Archive removes data from the source repository.
  - **Archive and Purge**. Data Archive extracts data from the source repository, loads the data into the target repository, and removes the data from the source repository.
6. Select the **Source** connection you created in Step 13.  
A list of source connection properties appear with default values.
7. Update the following source connection properties or keep the default values:
  - **Analyze Interim**. Determines when the interim table is analyzed for table structure and data insertion.
  - **Delete Commit Interval**. Number of rows per thread that the archive job deletes before the job issues a commit to the database.
  - **Insert Commit Interval**. Number of database insert transactions after which a commit point should be created. Commit interval is applicable when the archive job uses JDBC for data movement.
  - **Delete Degree of Parallelism**. Number of concurrent threads that are available for parallel deletion. The number of rows to delete for each table is equally distributed across the threads.



- **Insert Degree of Parallelism.** The number of concurrent threads for parallel insertions.
  - **Update Degree of Parallelism.** The number of concurrent threads for parallel updates.
  - **Reference Data Store Connection.** Required when you use the history database as the source and you only want to purge from the history database.
8. Select the **Target** connection you created in Step 14.  
The target connection properties appear.
  9. Update the following target connection properties:
    - **Include Reference Data.** Select this option if you want the archive project to include transactional and reference data from ERP tables.
    - **Reference Data Store Connection.** Required when you use a history database as the source.
- In the airline example, you keep the default values for the source and target connection parameters.

The screenshot shows the 'Create or Edit an Archive Project' wizard in Informatica Information Lifecycle Management. The 'General Information' tab is selected, displaying the following fields and values:

- Project Name:** Airline Loyalty
- Description:** Ticket and coupon info
- Action:** Archive And Purge
- Source:** AirlineSource
  - Analyze Interim: After Insert and Update
  - Delete Commit Interval: 30000
  - Insert Commit Interval: 10000
  - Delete Degree of Parallelism: 1
  - Insert Degree of Parallelism: 1
  - Update Degree of Parallelism: 1
- Target:** IDV Test
  - Include Reference Data: ☐
  - Seamless Access Required: ☐

At the bottom of the wizard, there are navigation buttons: Previous, Next, Save Draft, Publish, and Cancel.

10. Click **Next**.
11. Click **Add Entity**.  
The list of entities appear.
12. Select the archive entity and click **Select**.
13. Specify values for the following entity parameters:
  - **Candidate Generation Report Type.** Determines the report that the archive project generates during the candidate generation step. You can select **Summary**, **Detail** or **None**.
  - **Enabled.** Determines whether to exclude the entity from archiving. Clear the checkbox to exclude the entity from archiving. Default is enabled.
  - **Policy.** Displays a list of retention policies that you can select for the entity. If you do not select a retention policy, the records in the entity do not expire.
  - **Role.** The Data Vault access role for the archive project. The access role determines who can view data from Data Discovery searches.  
Select the Data Vault access role you created in Step 15.
14. Select an operator and enter a value for the archive parameter you created in Enterprise Data Manager.
15. Click **Next**.  
The archive steps appear.

16. Select the **Row Count Report** check box for the step **Copy to Destination**.

| Step                    | Pause After              | Row Count Report                    | Skip                     | Run Before           | Run After            | Notify                   |
|-------------------------|--------------------------|-------------------------------------|--------------------------|----------------------|----------------------|--------------------------|
| 1. Generate Candidates  | <input type="checkbox"/> | <input type="checkbox"/>            | <input type="checkbox"/> | <input type="text"/> | <input type="text"/> | <input type="checkbox"/> |
| 2. Build Staging        | <input type="checkbox"/> | <input type="checkbox"/>            | <input type="checkbox"/> | <input type="text"/> | <input type="text"/> | <input type="checkbox"/> |
| 3. Copy To Staging      | <input type="checkbox"/> | <input type="checkbox"/>            | <input type="checkbox"/> | <input type="text"/> | <input type="text"/> | <input type="checkbox"/> |
| 4. Validate Destination | <input type="checkbox"/> | <input type="checkbox"/>            | <input type="checkbox"/> | <input type="text"/> | <input type="text"/> | <input type="checkbox"/> |
| 5. Copy To Destination  | <input type="checkbox"/> | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="text"/> | <input type="text"/> | <input type="checkbox"/> |
| 6. Purge Staging        | <input type="checkbox"/> | <input type="checkbox"/>            | <input type="checkbox"/> | <input type="text"/> | <input type="text"/> | <input type="checkbox"/> |

17. Select **Publish and Schedule**.

The **Schedule Job** page appears.

18. To start the archive process immediately, select **Immediately** in the **Schedule** section.

**Projects or Programs to Run**

☐ Projects ☒ Standalone Programs

**Airline Loyalty1** Order

☒ **File Archive Loader** Order

**Schedule**

**Start**

☒ Immediately ☐ On

Time Zone:

**Notification**

**When** ☐ Completed ☐ Terminated ☐ Error **To** Email:

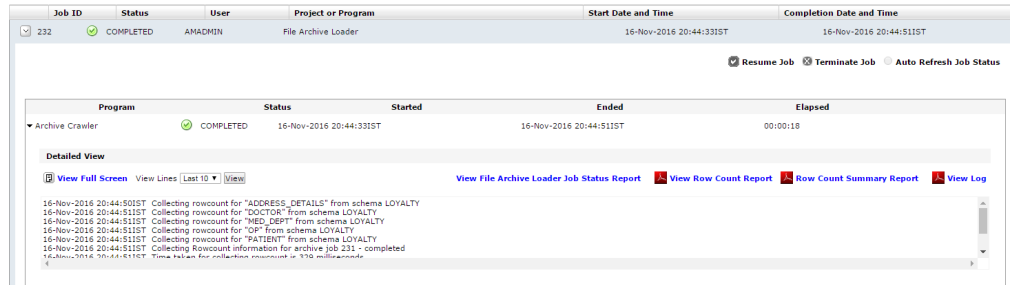
19. Click **Schedule**.

The **Monitor Jobs** page appears with the status of the extract job and the Data Vault Loader job. When the Data Vault Loader job is complete, the job status is **COMPLETED**.

## Step 16. Review the Row Count Report

Review the row count report to verify that the Data Vault Loader job moved the data to Data Vault.

1. In Data Archive, click **Jobs > Monitor Jobs**.
2. Wait for the Data Vault Loader job to complete.
3. Expand the view for the job.
4. Expand the view for the Archive Crawler program.



5. Click **View Row Count Report**.

The row count report opens in PDF format.

| <b>informatica</b>                    |                         | Page 1 of 2              |                          |
|---------------------------------------|-------------------------|--------------------------|--------------------------|
| Row Count Validation Report           |                         | Generated                | 16-Nov-2016 20:45:03 IST |
| Loader Job Id 232                     |                         | Started Date             | 16-Nov-2016 20:44:33IST  |
|                                       |                         | Completion Date          | 16-Nov-2016 20:44:51IST  |
| Definition                            | AirLine Medical Loyalty |                          |                          |
| Source Connection                     | Loyalty Source          |                          |                          |
| Target Connection                     | Loyalty Target          |                          |                          |
| Archive Job Id                        | 230                     | Created Date             | 16-Nov-2016 20:44:21IST  |
| Entity Name                           | AIRLINE_1_V1            | Scheduled Date           | 16-Nov-2016 20:44:22IST  |
|                                       |                         | Completion Date          | 16-Nov-2016 20:44:33IST  |
| Table Name                            | Source Row Count        | Archive Folder Row Count | Discrepancy              |
| LOYALTY.AIRPORTS                      | 16                      | 16                       | 0                        |
| LOYALTY.STATEMENT_LINE                | 12492                   | 12492                    | 0                        |
| LOYALTY.TICKET_COUPON                 | 74916                   | 74916                    | 0                        |
| LOYALTY.STATEMENT_LINE_SAVE           | 74234                   | 74234                    | 0                        |
| LOYALTY.CUSTS_WITH_INCOMPLETE_COUPONS | 645                     | 645                      | 0                        |
| LOYALTY.TICKET_SAVE                   | 30000                   | 30000                    | 0                        |
| LOYALTY.TICKET                        | 29990                   | 29990                    | 0                        |
| LOYALTY.LOYALTY_PERIODS               | 78                      | 78                       | 0                        |
| LOYALTY.TICKET_COUPON_SAVE            | 74939                   | 74939                    | 0                        |
| LOYALTY.CUSTOMER                      | 4364                    | 4364                     | 0                        |

Alternatively, you can click the **Row Count Summary** link to review the row count summary report.

6. For each table name, check the **Discrepancy** column.

If the **Discrepancy** column is 0, then the Data Vault Loader job moved the data from the source to the Data Vault.

## Step 17. Validate Data through the Data Discovery Portal

Use Browse Data from the Data Discovery portal to search and examine application data from specific tables in Data Vault.

When you use Browse Data, you specify the archive folder, entity, schema, and table you want to browse. You specify the table columns to display in the search results. You can use a `WHERE` clause to limit the search results. You can also specify the sort order of the search results. Before you run the search, you can preview the SQL statement that the Data Discovery portal uses to access the data.

1. Click **Data Discovery > Browse Data**.
2. Specify the following search parameters:

### **Archive Folder**

Archive folder that contains the entity of the archived data that you want to search.

### **No. of Rows**

Maximum amount of rows in the search results. Default is 100.

### **Entity**

Entity that contains the table that you want to view archived data from.

Optional. Select the entity to limit the amount of tables shown in the list of values.

### **Schema**

Schema that contains the table that you want to view.

### **Table**

Table that includes the archived data that you want to view.

After you select the table, Data Archive populates the **Available Columns** list.

3. Use the arrows to move the columns that you want to display to the **Display Columns** list.
4. Optionally, enter a `WHERE` clause in the **Where Clause** box to filter the search results.
5. Optionally, specify the sort order of the search results in the **Order By** box.
6. Optionally, click **Preview SQL** to view the generated SQL query.
7. Click **Search**.

The records that match the search criteria appear.

\*\*\*\*\*  
**informatica** Information Lifecycle Management Administrator Log Off | Help

Home Accelerators Workbench Administration Jobs Data Discovery Data Visualization Help

**Browse Data**

★ Required field.

**Criteria**

★ Archive Folder    
Entity

**Data Columns**

| Available Columns | Display Columns  | Where Clause | Order By |
|-------------------|--|--------------|----------|
|                   | <div><div>&gt;&gt;</div><div>&gt;</div><div>&lt;</div><div>&lt;&lt;</div></div> <div>CUST_ID<br/>MEMBERSHIP_STATUS<br/>MEMBER_SINCE_DATE<br/>LAST_ACTIVITY_DATE<br/>MILEAGE_BALANCE<br/>LAST_NAME<br/>FIRST_NAME</div> |              |          |

**Results 1-10 of 100**

| CUST_ID | MEMBERSHIP_STATUS | MEMBER_SINCE_DATE    | LAST_ACTIVITY_DATE   | MILEAGE_BALANCE | LAST_NAME | FIRST_NAME  | Hold | Retention | Expiration Date | Effect |
|---------|-------------------|----------------------|----------------------|-----------------|-----------|-------------|------|-----------|-----------------|--------|
| 833     | GENERAL           | 17-Mar-1998 00:00:00 | 01-Jun-2005 00:00:00 | 11000           | Williams  | Rex         |      |           |                 |        |
| 834     | GENERAL           | 01-Jan-1987 00:00:00 | 08-Dec-2001 00:00:00 | 10000           | Stoddard  | George      |      |           |                 |        |
| 853     | GENERAL           | 15-Jan-1997 00:00:00 | 02-Dec-2005 00:00:00 | 10500           | Benson    | Julianne    |      |           |                 |        |
| 854     | GENERAL           | 15-Jan-1997 00:00:00 | 07-Jul-2005 00:00:00 | 9500            | Cambria   | Sharon Kaye |      |           |                 |        |
| 872     | GENERAL           | 15-Feb-1997 00:00:00 | 02-Dec-2005 00:00:00 | 4500            | William   | Paul        |      |           |                 |        |

8. Verify that the values for the records in Data Vault match the values for the same records on the source database.