Importing and Exporting Oracle Data Assets in Intelligent Data Lake
Abstract
You can import data assets from Oracle to Intelligent Data Lake and export data assets to Oracle. The article lists the steps required for importing and exporting data assets.

Supported Versions
• Intelligent Data Lake 10.1.1, 10.2

Table of Contents
Overview .......................................................... 2
Creating a JDBC Connection to Access Oracle ........................................... 2
Importing a Data Asset ........................................................ 4
Exporting a Data Asset to External Targets ............................................... 5
Limitations for Oracle Import and Export Operations ............................... 6

Overview
You can import data from Oracle tables into the Hadoop data lake and export data from the Hadoop data lake to Oracle.
You can add the imported data assets to your projects as worksheets. You can perform data preparation on the worksheets and publish the prepared worksheets. For the data assets that you export to Oracle, you can perform further operations in Oracle.

Creating a JDBC Connection to Access Oracle
Create a JDBC connection to connect to the Oracle database to import and export data assets, preview data, and profile data.
Note that exporting data to an Oracle database from Intelligent Data Lake on a Cloudera CDH cluster requires the use of upper case table names and column names. If you create a connection to an Oracle database from Intelligent Data Lake on a Cloudera CDH cluster, you must add the CustomOptions.sqoopSchemaUpperCase custom property to the Intelligent Data Lake Service.
You do not need to add a custom property to connect to an Oracle database from any other Hadoop distribution.

1. In the Administrator tool, click the Manage tab.
2. Click the Connections view.
3. In the Navigator, select the domain.
4. In the Navigator, click Actions > New > Connection.
   The New Connection dialog box appears.
5. In the New Connection dialog box, select Database > JDBC.
6. Click OK.
   The New Connection wizard appears.
7. Enter the connection properties.
8. Click Next.
When you finish entering connection properties, you can click **Test Connection** to test the connection.

9. Enter the following JDBC connection properties for Oracle.

<table>
<thead>
<tr>
<th>Property</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Database Type</td>
<td>The database type.</td>
</tr>
<tr>
<td>Name</td>
<td>Name of the connection. The name is not case sensitive and must be unique within the domain. The name cannot exceed 128 characters, contain spaces, or contain the following special characters: `~ ! $ % ^ * ( ) - + = { ] }</td>
</tr>
<tr>
<td>ID</td>
<td>String that the Data Integration Service uses to identify the connection. The ID is not case sensitive. It must be 255 characters or less and must be unique in the domain. You cannot change this property after you create the connection. Default value is the connection name.</td>
</tr>
<tr>
<td>Description</td>
<td>The description of the connection. The description cannot exceed 765 characters.</td>
</tr>
<tr>
<td>User Name</td>
<td>The database user name.</td>
</tr>
<tr>
<td>Password</td>
<td>The password for the database user name.</td>
</tr>
<tr>
<td>JDBC Driver Class Name</td>
<td>The full name of the class of the JDBC driver class.</td>
</tr>
<tr>
<td></td>
<td>com.informatica.jdbc.oracle.OracleDriver</td>
</tr>
<tr>
<td></td>
<td>For more information about which driver class to use with specific databases, see the vendor documentation.</td>
</tr>
<tr>
<td>Connection String</td>
<td>Connection string to connect to the database. Use the following connection string: jdbc:informatica:oracle://&lt;host_name&gt;:&lt;port&gt;;SID=&lt;database name&gt; Provide your own connection string if you download the JDBC driver and want to use it for the connection.</td>
</tr>
<tr>
<td>Environment SQL</td>
<td>Optional. Enter SQL commands to set the database environment when you connect to the database. The Data Integration Service executes the connection environment SQL each time it connects to the database.</td>
</tr>
<tr>
<td>Transaction SQL</td>
<td>Optional. Enter SQL commands to set the database environment when you connect to the database. The Data Integration Service executes the transaction environment SQL at the beginning of each transaction.</td>
</tr>
<tr>
<td>SQL Identifier Character</td>
<td>Select &quot; &quot; from the drop down list to connect to the database from a Hortonworks HDP cluster.</td>
</tr>
<tr>
<td>Support Mixed-case Identifiers</td>
<td>Select Yes from the drop down list.</td>
</tr>
<tr>
<td>Use Sqoop Connector</td>
<td>The Sqoop connector to connect the data lake to the database.</td>
</tr>
<tr>
<td>Sqoop Arguments</td>
<td>Enter the arguments that Sqoop must use to connect to the database. Separate multiple arguments with a space. Use the following argument string: --connect jdbc:oracle:thin:@&lt;host_name&gt;:&lt;port&gt;/&lt;database name&gt;`</td>
</tr>
</tbody>
</table>
If you create a connection to an Oracle database from Intelligent Data Lake on a Cloudera CDH cluster, you must add the CustomOptions.sqoopSchemaUpperCase custom property to the Intelligent Data Lake Service. The property configures the connection to use upper case table names and column names.

Click **Services and Nodes** in the Administrator tool, and then select the Intelligent Data Lake Service.

1. Click the pencil icon in the Custom Options section of the page.
2. Enter CustomOptions.sqoopSchemaUpperCase as the custom property name.
3. Enter the name of the JDBC connection as the custom property value.

The following image shows the New Custom Property dialog box:

![New Custom Property dialog box](image)

4. Click **OK**.

**Importing a Data Asset**

You can import a data asset from Oracle into the data lake.

You need to have access privileges to a connection and **Read** and **Execute** permissions for each connection to import a data asset using that connection.

1. Log in to Intelligent Data Lake.
2. On the Asset Overview of any data asset from Oracle, click **Import to Lake**. The **Import** dialog box appears.
3. If the asset has duplicates in the lake, a warning message appears. Click **Import Anyway** if you'd like to import the asset.
4. Select a connection for the asset from the specified resource. Click **Next**.
   
   **Note:** Contact your administrator if you do not see any connection in the list or if you do not see a desired connection in the list.
5. Enter the target location details to import the asset using the connection.
   a. Select a schema from the drop down list.
b. Enter the name of the table. The name of the table must be in lower case. The name can include letters, numbers, and the underscore character. The first character of the name cannot be an underscore or a number.

c. If the table already exists, select Append or Overwrite.
   - Append. Select this to append the data to an existing table. Make sure the column names, the data types of the columns, precision and scale of columns, and sequence of columns are same for the importing table and the existing table. The import will fail if any of these do not match.
   - Overwrite. Select this to overwrite the existing table. It drops the table and re-creates the table with the data from the importing table.

6. Click Import to import the table into the data lake.
   Check the progress of the import activity on My Activities page. You can see if the import activity is in progress, completed, or if it has failed.

   For more information about uploading data assets from local folder, see the Upload Data chapter.

**Exporting a Data Asset to External Targets**

You can export a data asset or a publication to Oracle.

You need to have access privileges to a connection and Read and Execute permissions for each connection to export a data asset using that connection.

1. Log in to Intelligent Data Lake.
2. On the Asset Overview of any data asset, click the Manage Data Assets icon (assets), and then click Export.
   The Export dialog box appears.
3. Select a connection for the asset. Click Next.
   **Note:** Contact your administrator if you do not see any connection in the list.
4. Enter the target location details to export the asset using the selected connection.
   a. Enter the name of the schema.
   b. Enter the name of the table. The name can include letters, numbers, and the underscore character. The first character of the name cannot be an underscore or a number.
5. Click Export.
6. If a table does not exist with the same name, a new table is created. If a table exists with the name already, the table schema will be validated and the following options appear:
   - If the source table has the same schema structure as the target table, the following options appear:
     - Append. Select this option to append the data to an existing table.
     - Overwrite. Select this option to overwrite the existing table. It drops the table and re-creates the table with the data from the exporting table.
   - If the source table has a different schema structure, select to overwrite the table. This replaces the existing table with the source data.

   Check the progress of the export activity on My Activities page. You can see if the export activity is in progress, completed, or if it has failed.
Limitations for Oracle Import and Export Operations

Your import and export operations can fail due to certain limitations.

Export operation limitations

Export operation to Oracle fails for the following reasons:

- A float data type mismatch. When you try to export a data asset that has a column with float data type, the export operation fails due to third-party limitations.
- A varchar data type mismatch. When you try to export a data asset that has a column with varchar data type, the export operation fails due to third-party limitations.
- A double data type mismatch. When you try to export a data asset that has a column with double data type, the export operation fails due to third-party limitations.

Import operation limitations

Import operation from Oracle fails for the following reasons:

- A timestamp data type mismatch. When you try to import a data asset that has a column of timestamp data type with time zone, the import operation fails due to third-party limitations.
- An encoded table name. If the table name is encoded or contains non-English characters, lineage information is not added to the catalog during the import operation.
- Negative Scale Number. If the table has a column with negative scale in Oracle database, the import operation fails due to third-party limitations.

Authors

Manoj Rajan
Principal QA Engineer

Swathi N.R.
Senior QA Engineer

Arun Porwal
Senior QA Engineer

Chakravarthy Tenneti
Lead Technical Writer