

Create Dataset with Parameters

Contents:

- *From File System*
 - *Parameterize bucket names*
- *From Relational Sources*
- *Edit Parameter*
- *Apply Parameter Overrides*
 - *Apply parameter overrides for your flow*
 - *Apply parameter overrides for your job*
- *Delete Parameter*

This section provides an overview on how to parameterize relational sources and files while importing data to your flow.

From File System

When browsing for data on your default storage layer, you can choose to parameterize elements of the path. Through the Import Data page, you can select elements of the path, apply one of the supported parameter types and then create the dataset with parameters.

NOTE: When you import a file, the data is not stored in Trifacta Wrangler Pro . What you create is an imported dataset that is simply a reference to the source of the data. Trifacta Wrangler Pro never stores or modifies source data.

When you create a dataset with parameters in Trifacta Wrangler Pro, you can replace segments of the input path with parameters. Suppose you have the following files that you'd like to capture through a parameterized dataset:

```
//source/user/me/datasets/month01/2017-01-31-file.csv
//source/user/me/datasets/month02/2017-02-28-file.csv
//source/user/me/datasets/month03/2017-03-31-file.csv
//source/user/me/datasets/month04/2017-04-30-file.csv
//source/user/me/datasets/month05/2017-05-31-file.csv
//source/user/me/datasets/month06/2017-06-30-file.csv
//source/user/me/datasets/month07/2017-07-31-file.csv
//source/user/me/datasets/month08/2017-08-31-file.csv
//source/user/me/datasets/month09/2017-09-30-file.csv
//source/user/me/datasets/month10/2017-10-31-file.csv
//source/user/me/datasets/month11/2017-11-30-file.csv
//source/user/me/datasets/month12/2017-12-31-file.csv
```

A parameterized reference to all of these files would look something like:

```
//source/user/me/datasets/month##/YYYY-MM-DD-file.csv
```

Through the application, you can specify the parameters to match all values for:

- ## - You can use a wildcard or (better) a pattern to replace these values.
- YYYY-MM-DD - A formatted Datetime parameter can replace these values.

For more information, see *Parameterize Files for Import*.

Parameterize bucket names

You can create environment parameters for your bucket names. For more information, see *Parameterize Files for Import*.

From Relational Sources

You can create datasets from a relational source by applying parameters to the custom SQL that pulls the data from the source. During import of database tables through relational connections, you can apply parameters to the SQL query that you use to define the imported dataset. In some scenarios, you may need to define the table to import using a variable parameter or to parameterize the time value associated with a table name. Using parameters, you can define the tables, columns, and conditions of the query that you use to bring in data from a relational database.

For more information, see *Parameterize Tables for Import*.

Edit Parameter

After you have created your dataset with parameters, you can edit the parameter as needed.

Steps:

1. In the left nav bar, select **Library**.
2. In the Library page, locate the dataset. From its context menu, select either of the following:
 - a. **Files:** Select **Edit parameters**. In the Edit Dataset with Parameters, click the parameter to modify its definition. For more information, see *Parameterize Files for Import*.
 - b. **Tables:** Click **Edit Custom SQL**. In the Custom SQL window, you can modify the SQL statement, including any parameters in it. For more information, see *Parameterize Tables for Import*.
 - i. For more information, see *Create Dataset with SQL*.

Apply Parameter Overrides

After you have created a parameterized dataset, you can apply overrides to the default value. These override values can be applied in the following cases.

Case	Precedence	Scenario
Job	1	When you choose to execute a job, you can set a new value for the parameter, which is applied for the specified job only.
Flow	2	If your imported dataset containing a parameter is added to a flow, you can define an override value for the dataset's parameter through Flow View. Whenever a job is executed on the imported dataset within the flow, the override value is applied to the dataset. <div style="border: 1px solid black; padding: 5px; margin: 10px 0;">NOTE: If a job-level override is applied on top of a flow-level override, the job override value is applied to the job.</div>
Default	3	The default value for the parameter is used if no override is applied.

Apply parameter overrides for your flow

Steps:

1. In Flow View, select the dataset with parameters icon.
2. From the context menu, select **Parameter**.
3. In the Manager Parameter dialog, click the Overrides tab.
4. Edit the required values, click **Save**.

For more information, see *Manage Parameters Dialog*.

Apply parameter overrides for your job

You can apply job-level parameter overrides through the Trifacta application or through the APIs.

via Trifacta application:

1. In Flow View, select the output that you wish to generate.
2. In the right context panel, click **Run Job**.
3. In the Run Job page, you can specify job-level overrides at the bottom of the screen.

For more information, see *Run Job Page*.

via APIs:

For more information, see *API Workflow - Run Job*.

Delete Parameter

Steps:

1. In the Edit Dataset with Parameters screen, select the parameter that you wish to remove.

NOTE: Before you remove parameter, you may want to take note of the default value, which may need to be applied to the path or query after you remove the parameter.

2. In the popup, click **Delete**.
3. Save your changes.
4. The parameter is removed from the imported dataset definition.