Overview of Data Export

Contents:

This section provides an overview of exporting data from the Trifacta® application to your preferred destinations, such as file-based storage, connected datastores, or your desktop. In addition to exporting of job results, other types of exports are covered in this section.

> **Tip:** In most cases, the source of your data does not limit the type of output that you can generate. You can create a file-based imported dataset and generate results to a database table. Some exceptions may apply.

How to Export

Job results are generated based on the specifications of an output object. An **output object** is a reference object for one or more types of outputs. This reference information includes full path to the output location, file or table name, and other settings. For more information, see *Create Outputs*.

In the Run Job page, you can specify additional settings and overrides. See *Run Job Page*.

Export Job Results

After you have executed a job, the application writes a set of results to the designated output locations. These results are the application of the recipe's transformation steps to the imported dataset, written to the location or locations specified in the output object in the specified output format.

You can export the results directly from the designated output destination. For more information, see *Job Details Page*.

> **Tip:** Job results for your latest job may be exportable from Flow View. For more information, see *View for Outputs*.

Writing to Files

As a result of job execution, you can publish your outputs to a file-based system.

> **NOTE:** You must have write permissions to the location where you are writing your output files. These permissions should be set up during initial configuration of the product. For more information, please contact your administrator.

Defaults for file-based outputs:

- Files are written to your designated output directory on the backend datastore. As needed, you can modify your default output directory. For more information, see *Storage Config Page*.
- Files are written in CSV format to the designated location.

You can modify the publishing action and generate results in your preferred formats.

- For more information on changing file output settings, see *File Settings*.
- For more information on supported file formats, see *Supported File Formats*.

Writing to Tables

You can export generated results directly to a connected relational database.

> **Tip:** Some relational connection types support read-only or write-only connections.

The Trifacta application writes results to a database through an object called a connection. A **connection** is a configuration object that defines the interface between the application and the database. Among its properties are a set of credentials that provide access.

> **NOTE:** You must have write permissions to the database where you are writing your output tables. These permissions must be enabled by a database administrator outside of the product.

> **NOTE:** Connections can be shared among users. When a user chooses to share a connection, the user can also choose to share credentials. If credentials are not shared, other users must provide their own credentials if they wish to use the connection. For more information, see *Share a Connection*.

For relational databases, the Trifacta application passes the information in the connection definition to a third-party driver that performs the actual connection. Thereafter, the Trifacta application maintains the open connection as long as it is needed to write results. After the results are written, the connection is closed.

When you choose to write results to a table:

- Through the connection, you browse and select the database to which to write the results.
- You can choose to write to an existing table or to a new one.
- You can specify one of the following publishing actions on the table you selected:
    - **New**: Each run generates a new table with a timestamp appended to the name. For example, `myexample_test_1.csv`.
    - **Update**: Each run adds any new results to the end of the table.
    - **Truncate**: With each run, all rows columns of data in a table is removed and retain the empty table as an object.
    - **Load**: With each run, the table is dropped (deleted), and all data is deleted. A new table with the same name is created, and any new results are added to it.
    - **Merge:** Some databases may support merge (upsert) operations.

Additional options may be available, depending on the connection. For more information, see *Relational Table Settings*.

Parameterized Outputs

For file-or table-based publishing actions, you can parameterize elements of the output path. You can create parameters for your outputs of the following types:

- **Timestamp:** You can insert the timestamp of when the output was written as part of the path to the output location.
- **Variable:** Variable parameters allow you to insert values that you define as part of the output object.

> **Tip:** You can optionally override the values of your variable parameters as part of your job definition.

For more information on parameters, see *Overview of Parameterization*.

Ad-hoc Publishing

After a job has successfully completed, you can review and download the set of generated outputs and export results. Optionally, you may be able to publish the generated results to a secondary datastore through the Job Details page.

> **NOTE:** Additional configuration may be required.

For more information on ad-hoc publishing, see *Publishing Dialog*.

Exporting Metadata

In addition to the job results, you can export aspects of the flow definition and other objects that you have created in the Trifacta application. These exports can be useful for:

- Migrating flows to other workspaces
- Archiving data
- Taking snapshots of work in progress

Export flows

You can export a flow from application. An exported flow is stored in a ZIP file that contains references to all objects needed to use the flow in another workspace or project. Exported flows can be imported into the same workspace/project or a different one.

> **NOTE:** Users of the imported flow must have access to the datasources and specified output locations. If not, these objects must be remapped in the new environment.

For more information, see *Export Flow*.

Export recipes

You can download a recipe in text form and reuse it in another flows.

Reuse recipes in a different environment

If you need to reuse a recipe in a different instance of Trifacta , you can do the following:

- Export the entire flow and import it into the new environment. Open the flow in the new environment.
- Convert all steps of a recipe into a macro. Export the macro and then import it into the new environment. For more information, see *Export Macro*.

Download recipes

You can download recipe in a text form of Wrangle (a domain-specific language for data transformation). For more information, see *Recipe Panel* .

Export sample data

From the recipe panel, you can download the current state of the data grid, which includes the current sample plus any recipe steps that have been applied to it.

> **Tip:** When a sample is taken, it is tied to the current recipe step. All steps later in the recipe than the current recipe step are computed in memory using the sample as the baseline. For more information, see *Overview of Sampling*.

F or example, if the sample was generated when the recipe cursor was displaying step 7 and you download the data from the recipe when the recipe cursor is on step 10, then you are downloading the state of the recipe at step 10.

> **NOTE**: When a flow is shared, its samples are shared with other users. However, if shared users do not have access to the underlying sources that back a sample, they do not have access to the sample. These samples are invalid for the other users, who must create their own.

For more information, see *Samples Panel*.

Export via API

Job results

After a job has run, you can acquire the path to the results when you query for the job. For more information, see Trifacta API Reference docs: *Enterprise | Professional | Premium*