

Workflow Basics

Contents:

- *Overview*
 - *Pre-requisites*
 - *Basic Workflow*
 - *Example Flows*
-

Learn the basics of how to import, wrangle, execute jobs, profile, and export your data from Trifacta Wrangler Pro.

Overview

Trifacta Wrangler Pro enables analysts, data specialists, and other domain experts to quickly cleanse and transform datasets of varying sizes for use throughout the enterprise. Using an innovative set of web-based tools, you can import complex datasets and wrangle them for use in virtually any target system. Key capabilities include:

- **Import** from flat file, databases, or distributed storage systems
- Locate and remove or modify **missing or mismatched** data
- **Unnest complex** data structures
- Identify statistical **outliers** in your data for review and management
- Perform **lookups** from one dataset into another reference dataset
- Aggregate columnar data using a variety of **aggregation functions**
- **Normalize** column values for more consistent usage and statistical modeling
- Merge datasets with **joins**
- Append one dataset to another through **union** operations

Most of these operations can be executed with a few mouse clicks. This section provides a basic overview of common workflows through Trifacta Wrangler Pro.

Pre-requisites

Before you begin, please verify the following:

- **Trifacta account:** You have a Trifacta account and can login.
- **Example data:** You should use a sample set of data during this workflow.

Basic Workflow

1. **Import data:** Integrate data from a variety of sources of data.

Tip: When you login for the first time, you can immediately upload a dataset to begin transforming it.

See *Import Basics*.

2. **Profile your data:** Before, during, and after you transform your data, you can use the visual profiling tools to quickly analyze and make decisions about your data. See *Profiling Basics*.
3. **Build transform recipes:** Use the various views in the Transformer Page to build your transform recipes and preview the results on sampled data. See *Transform Basics*.

4. **Sample your data:** In Trifacta Wrangler Pro, you create your recipes while working with a sample of your overall dataset. As needed, you can take new samples, which can provide new perspectives and enhance performance in complex flows. See *Sampling Basics*.
5. **Run job:** Launch a job to run your recipe on the full dataset. Review results and iterate as needed. See *Running Job Basics*.
6. **Export results:** Export the generated results data for use outside of Trifacta Wrangler Pro. See *Export Basics*.

Object overview: You should review the overview of the objects that are created and maintained in Trifacta Wrangler Pro. See *Object Overview*.

Example Flows

New users of Trifacta Wrangler Pro can learn by example! When a new workspace is created, the first user can access example flows through the product. You can learn more about how to build a flow and perhaps acquire some tips on how to apply to your uses.

Steps:

1. Log in to the Trifacta application.
2. In the left navigation bar, click the Flows icon.
3. Select any of the `Example Flow` flows.
4. Each flow contains detailed notes on the various objects in the flow, as well as recommended practices for building your own flows.

NOTE: These flows are provided only to the first user of a new workspace. This user must invite other users and can then share the flow.

- To invite other users, select **User menu > Admin console > Users**. Click **Invite Users**. For more information, see *Workspace Users Page*.
- After users have been invited, you can share these example flows with them. Open one of the flows. From the flow context menu, select **Share**. Specify the usernames of the users with whom you wish to share the flow. For more information, see *Share Flow Dialog*.