

Workflow Basics

Contents:

- *Overview*
 - *Pre-requisites*
 - *Basic Workflow*
-

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

Overview

Trifacta® Wrangler 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
- 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.

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. **Review object overview:** Before you begin, you should review the overview of the objects that are created and maintained in Trifacta Wrangler. See *Object Overview*.
2. **Import data:** Integrate data from a variety of sources of data. See *Import Basics*.
3. **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*.
4. **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*.
5. **Sample your data:** In Trifacta Wrangler, 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*.
6. **Run job:** Launch a job to run your recipe on the full dataset. Review results and iterate as needed. See *Running Job Basics*.
7. **Export results:** Export the generated results data for use outside of Trifacta Wrangler. See *Export Basics*.