

Diagnose Failed Jobs

Contents:

- *Job Types*
- *Identify Job Failures*
 - *Jobs that Hang*
- *Try Other Job Options*
- *Download Error Report*

Job Types

The following types of jobs can be executed in Dataprep by Trifacta®:

- **Transform job:** This type of job executes the steps in your recipe against the dataset to generate results in the specified format. When you configure your job, any set of selected output formats causes a transform job to execute according to the job settings.
- **Profile job:** This type of job builds a visual profile of the generated results. When you configure your job, select **Profile Results** to generate a profile job.
- **Publish job:** This job publishes results generated by the platform to a different location or datastore.
- **Ingest job:** This job manages the import of data from a JDBC source into the default datastore for purposes of running a transform or sampling job.

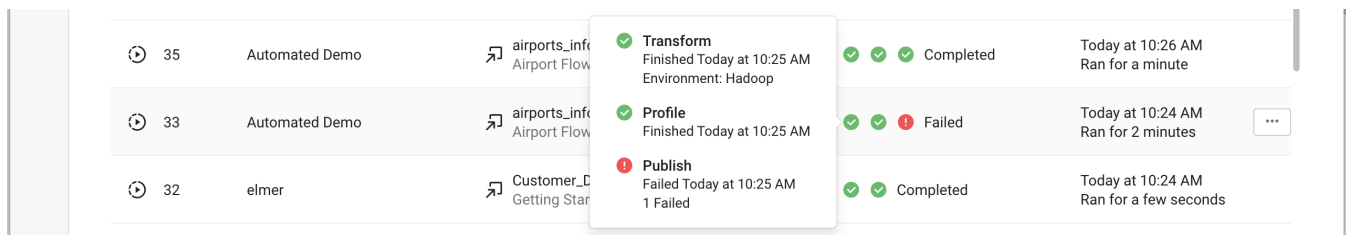
For more information, see *Run Job Page*.

Identify Job Failures

When a job fails to execute, a failure message appears in following locations:

- Jobs tab in Flow View. See *Flow View Page*.
- Listing in the Jobs page. See *Jobs Page*.

The following is an example from the Jobs page:



35	Automated Demo	airports_info Airport Flow	Transform Finished Today at 10:25 AM Environment: Hadoop	Completed	Today at 10:26 AM Ran for a minute
33	Automated Demo	airports_info Airport Flow	Profile Finished Today at 10:25 AM	Failed	Today at 10:24 AM Ran for 2 minutes
32	elmer	Customer_D Getting Star	Publish Failed Today at 10:25 AM 1 Failed	Completed	Today at 10:24 AM Ran for a few seconds

Figure: Publish job failed

In the above example, the Transform and Profile jobs completed, but the Publish job failed. In this case, the results exist and, if the source of the problem is diagnosed, they can be published separately.

Jobs that Hang

In some cases, a job may stay in a pending state indefinitely. Typically, these errors are related to a failure of the job tracking service. You can try to the following:

- Resubmit the job.

Try Other Job Options

You can try to re-execute the job using different options.

Tips:

- **Look to cut data volume.** Some job failures occur due to high data volumes. For jobs that execute across a large dataset, you can re-examine your data to remove unneeded rows and columns of data. Use the Deduplicate transformation to remove duplicate rows. See *Remove Data*.
- **Gather a new sample.** In some cases, jobs can fail when run at scale because the sample displayed in the Transformer page did not include problematic data. If you have modified the number of rows or columns in your dataset, you can generate a new sample, which might illuminate the problematic data. However, gathering a new sample may fail as well, which can indicate a broader problem. See *Samples Panel*.

Download Error Report

When a job fails, you can download an error report, which can be useful for diagnosing issues:

Tip: If you are filing a request with *Trifacta Support*, please be sure to include this error report. Additionally, please provide the information available in the **About > Report issue** dialog.

1. Login to the Trifacta application as an admin or the user who created the job.
2. Click the Jobs link in the left navigation bar.
3. In the Jobs page, click the identifier for the job that failed.
4. In the Overview tab, review the Failed Stages entry or entries.
5. To acquire the error report, click **Download error report**. Review and forward the report as needed.

For each job in the failing job group, the error report contains:

Item	Description
jobGroupId	The internal job identifier within the workspace where the error occurred. This value matches the job ID that is displayed in the Trifacta application
workspaceId	The identifier of the workspace where the error occurred

Tip: The workspace-person combination of identifiers provides a unique ID for the account where the error occurred.

jobId	<p>The ID for the specific job in the job group. These jobs can include sampling, profiling, transformation, writing results, ingesting data, and more.</p> <div style="border: 1px solid #c8e6c9; padding: 10px; margin: 10px 0;"> <p>Tip: Individual job identifiers can be helpful in identifying specific problem areas.</p> </div>
personId	The internal identifier of the user within the workspace
status	The status of the job, such as "Failed", "Completed", "Pending".
errorMessageId	The ID of the error message that is shown in the UI.
errorTimestamp	Timestamp of the job failure. Not included if no error occurs for the job.
uiErrorMessage	Error message that is shown in the job UI. Not included if no error occurs for the job.

Contact Support

If you are unable to diagnose your job failure, please contact *Trifacta Support*.

NOTE: When you contact support about a job failure, please be sure to download and include the entire zip file, your recipe, and (if possible) your dataset.