

API JobGroups Create v3

Contents:

- *Required Permissions*
- *Request*
- *Response*
- *Reference*

Create a jobGroup, which launches the specified job as the authenticated user.

The request specification depends on one of the following conditions:

- Dataset has already had a job run against it and just needs to be re-run.
- Dataset has not had a job run, or the job definition needs to be re-specified.

NOTE: In this release, you cannot execute jobs sourced from datasets in Redshift or SQL DW or publish to these locations via the API. This known issue will be fixed in a future release.

Version: v3

Required Permissions

NOTE: Each request to the Trifacta® platform must include authentication credentials. See *API Authentication*.

Request

Request Type: POST

Endpoint:

```
/v3/jobGroups
```

Request Body - Re-run job:

If you are re-running a job that has already executed and do not need to modify any job settings, you can use the following simplified body to launch it:

```
{
  "wrangledDataset": {
    "id": 7
  }
}
```

Request Body - Specify job:

If you are specifying a new job or must re-run a job with new settings, you must include a version of the following request body. Required parameters are listed below:

```

{
  "wrangledDataset": {
    "id": 1
  },
  "overrides": {
    "execution": "photon",
    "profiler": false,
    "writesettings": [
      {
        "path": "hdfs://hadoop:50070/trifacta/queryResults/admin@trifacta.local/cdr_txt.csv",
        "action": "create",
        "format": "csv",
        "compression": "none",
        "header": false,
        "asSingleFile": false
      }
    ]
  },
  "ranfrom": "cli"
}

```

Response

Response Status Code - Success: 201 - Created

Response Body Example:

```

{
  "jobgroupId": 3,
  "jobIds": [
    5,
    6
  ],
  "reason": "JobStarted",
  "sessionId": "9c2c6220-ef2d-11e6-b644-6dbff703bdfc"
}

```

Reference

Request Reference:

Property	Description
wrangledDataset	(required) Internal identifier for the object whose results you wish to generate. The recipes of all preceding datasets on which this dataset depends are executed as part of the job.
overrides.execution	(required, if first time running the job) Indicates the running environment on which the job is executed. Accepted values: <ul style="list-style-type: none"> • photon • spark For more information, see <i>Running Environment Options</i> .
overrides.profiler	(required, if first time running the job) When set to <code>true</code> , a visual profile of the job is generated as specified by the profiling options for the platform. See <i>Profiling Options</i> .
overrides.writesettings	(required, if first time running the job) These settings define the publishing options for the job. See below.

ranfrom	(optional) If this value is set to <code>null</code> , then the job does not show up in the Job Results page. If set to <code>cli</code> , the job appears as a CLI job. See <i>Job Results Page</i> .
---------	--

writesettings Reference:

The `writesettings` values allow you to specify aspects of the publication of results to the specified `path` location.

NOTE: `writesettings` values are required if you are running this specified job for the dataset for the first time.

NOTE: To specify multiple outputs, you can include additional `writesettings` objects in the request. For example, if you want to generate output to `csv` and `json`, you can duplicate the `writesettings` object for `csv` and change the `format` value in the second one to `json`.

These settings correspond to values that you can apply through the UI or through the command line interface.

- For UI information, see *Run Job Page*.
- For CLI information, see *CLI for Jobs*.

Property	Description
<code>path</code>	(required) The fully qualified path to the output location where to write the results
<code>action</code>	(required) If the output file or directory exists, you can specify one of the following actions: <ul style="list-style-type: none"> • <code>create</code> - Create a new, parallel location, preserving the old results. • <code>append</code> - Add the new results to the old results. • <code>overwrite</code> - Replace the old results with the new results.
<code>format</code>	(required) Output format for the results. Specify one of the following values: <ul style="list-style-type: none"> • <code>csv</code> • <code>json</code> • <code>avro</code> • <code>pqt</code> <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <p>NOTE: Parquet format requires execution on a Hadoop running environment (<code>overrides.execution</code> must be <code>spark</code>).</p> </div> <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <p>NOTE: To specify multiple output formats, create additional <code>writesettings</code> object for each output format.</p> </div>
<code>compression</code>	(optional) For <code>csv</code> and <code>json</code> results, you can optionally compress them using <code>bzip2</code> or <code>gzip</code> compression. Default is <code>none</code> .
<code>header</code>	(optional) For <code>csv</code> results with <code>action</code> set to <code>create</code> or <code>append</code> , this value determines if a header row with column names is inserted at the top of the results. Default is <code>false</code> .
<code>asSingleFile</code>	(optional) For <code>csv</code> and <code>json</code> results, this value determines if the results are concatenated into a single file or stored as multiple files. Default is <code>false</code> .