Get started with Google-provided templatesbookmark_border

cloud.google.com/dataflow/docs/guides/templates/provided-templates

Google provides a set of open source Dataflow templates. For general information about templates, see the Overview page. To get started, use the `WordCount` template. See other Google-provided templates:

Streaming templates - Templates for processing data continuously: hi sc

Batch templates - Templates for processing data in bulk:

- BigQuery to Cloud Storage TFRecords
- BigQuery export to Parquet (via Storage API)
- Bigtable to Cloud Storage Avro
- Bigtable to Cloud Storage Parquet
- Bigtable to Cloud Storage SequenceFiles
- Datastore to Cloud Storage Text
- Cloud Spanner to Cloud Storage Avro
- Cloud Spanner to Cloud Storage Text
- Cloud Storage Avro to Bigtable
- Cloud Storage Avro to Cloud Spanner
- Cloud Storage Parquet to Bigtable
- Cloud Storage SequenceFiles to Bigtable
- Cloud Storage Text to BigQuery
- Cloud Storage Text to Datastore
- Cloud Storage Text to Pub/Sub (Batch)
- Cloud Storage Text to Cloud Spanner
- Java Database Connectivity (JDBC) to BigQuery
- Apache Cassandra to Bigtable
- Apache Hive to BigQuery
- File Format Conversion

Utility templates: Mis

WordCount

The WordCount template is a batch pipeline that reads text from Cloud Storage, tokenizes the text lines into individual words, and performs a frequency count on each of the words. For more information about WordCount, see WordCount Example Pipeline.

Template parameters

Parameter	Description

`inputFile`	The Cloud Storage input file's path.
`output`	The Cloud Storage output file's path and prefix.

Running the WordCount template

Console gcloud API

Run using Google Cloud Console.

1. Go to the Dataflow page in the Cloud Console.
   Go to the Dataflow page
2. Click *add*Create job from template.

3. Select the WordCount template from the **Dataflow template** drop-down menu.
4. Enter a job name in the **Job Name** field.
5. Enter your parameter values in the provided parameter fields.
6. Click **Run Job**.

Template source code

Rate and review