Dataprep by Trifacta Pricing

Dataprep is an interactive web application in which users define the data preparation rules by interacting with a sample of their data. Use of the application is free. Once a data preparation flow has been defined, the sample can be exported for free or the flow can be executed as a Dataprep job (using Dataflow (/dataflow)) over the original dataset.

Pricing

When you submit a job to Dataprep, it is executed by Dataflow (/dataflow) workers. Dataprep is billed according to the number of Dataflow worker virtual CPUs (vCPUs) that are needed to process a job and the time that the vCPUs are used multiplied by the Dataprep service rate of $0.60 per hour.

The Dataprep service is available in each of the Dataflow regions (/dataflow/docs/concepts/regional-endpoints).

Example: A Dataprep job runs for 1 hour and requires 5 Dataflow virtual CPUs.

Dataprep job cost = 1 hour * $0.60 * 5 vCPUs = $3.00

ımber of workers and job duration depend on how Dataflow is configured.

While the rate for pricing is per-hour, Dataprep service usage is billed in per-second increments on a per-job ba

The use of Dataprep to define flows is free of charge.

The use of other resources in a job, such as BigQuery or Cloud Storage,are billed separately at their pricing.

Refer to the Dataflow documentation (/dataflow/docs) and the Google Cloud Platform Pricing Calculator (/products/calculator) to estimate costs for Dataprep jobs.

Except as otherwise noted, the content of this page is licensed under the Creative Commons Attribution 4.0 License (https://creativecommons.org/licenses/by/4.0/), and code samples are licensed under the Apache 2.0 License

(https://www.apache.org/licenses/LICENSE-2.0). For details, see the Google Developers Site Policies (https://developers.google.com/site-policies). Java is a registered trademark of Oracle and/or its affiliates.

Last updated 2020-06-22 UTC.