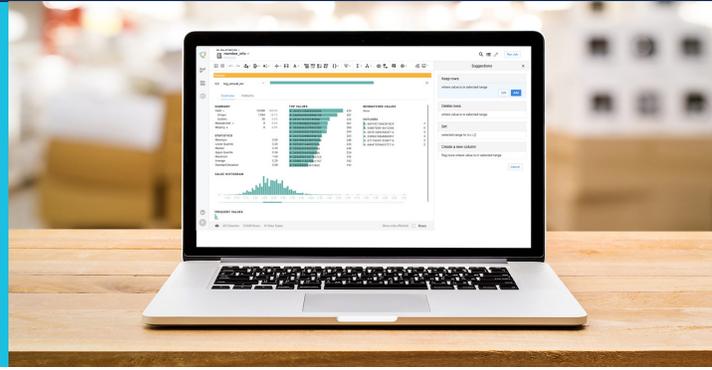


Getting Started with Self-Service Analytics on BigQuery with Cloud Dataprep



What is self-service analytics?

The objective of self-service analytics is to empower data-driven professionals to create their own end-to-end solution for their analytics, respond to their decision-making requirements, and research for insight. Self-service means that there are no necessary dependencies on others to get the job done and that the technology in use is simple yet powerful and scalable enough to be easily adopted by anyone to deliver results quickly.

What are the ingredients for self-service analytics on the Google Cloud Platform (GCP)?

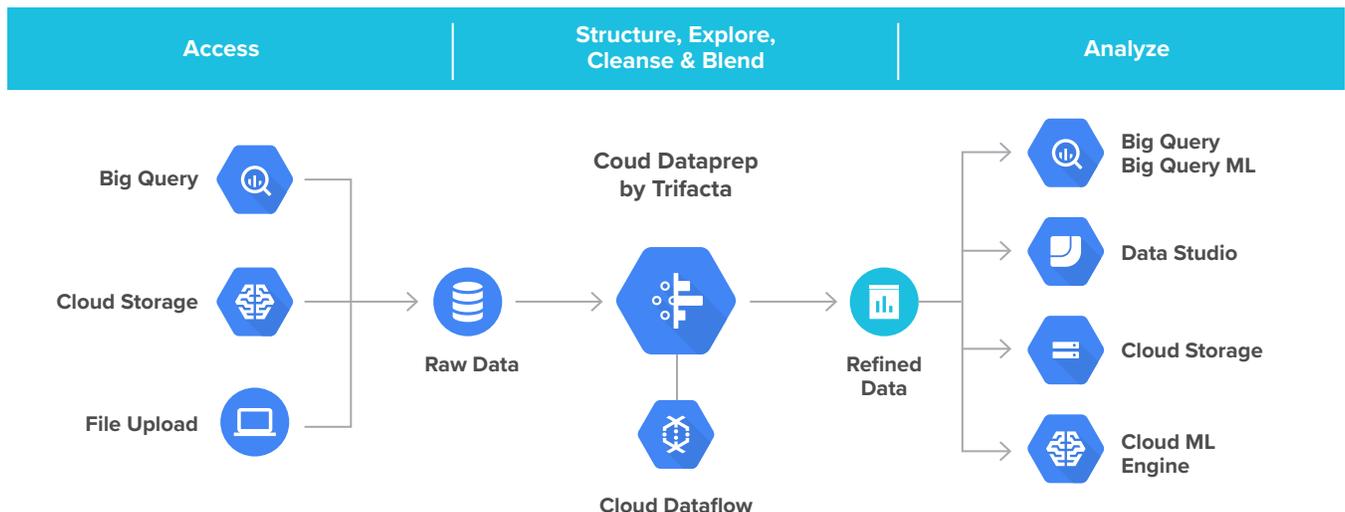
To establish a self-service analytics practice, you need to rely on three fundamental constituents: data storage, data preparation, and data visualization.

Data Storage. You need a space that your data can live and grow within. This data storage is the foundation to construct your analytics on top of. Within the GCP, the easiest way to store and retrieve your data for analytics at scale is BigQuery.

Data Preparation. From the data storage layer, you can prepare your data for reporting and dashboarding. You'll be using Cloud Dataprep to clean, standardize, combine, and create various calculations and metrics stored in BigQuery so that you can design your visualization layer on top of it.

Data Visualization. The visualization layer shows your data in the form of various reports and dashboards containing graphs and tables to visually represent data for your insight and to help your data-driven decision process. You'll be using Data Studio consuming BigQuery data to create your reports and dashboards.

With these fundamental elements in place, you're ready to implement a scalable and comprehensive self-service analytics solution.



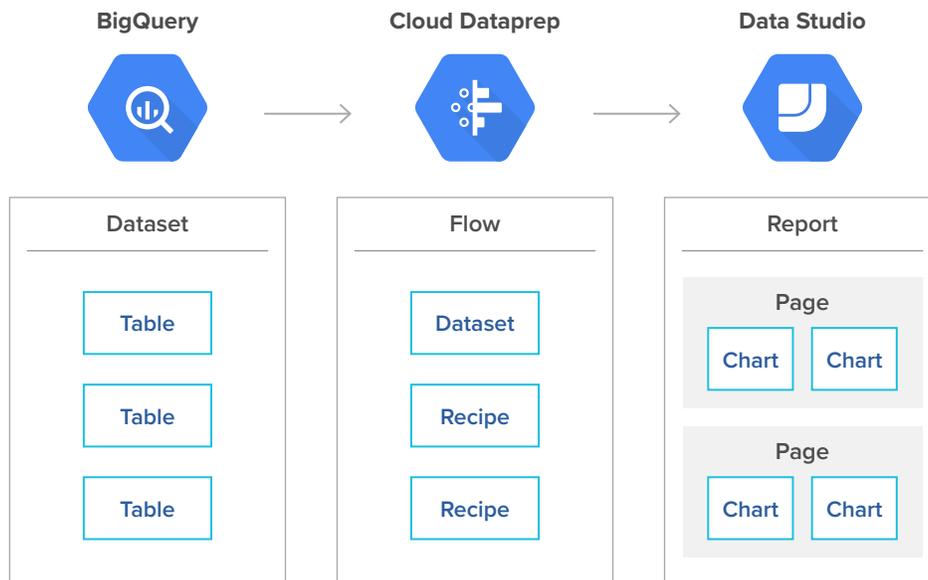
Understand GCP self-service analytics concepts

Here are the basic concepts and their relations for an end-to-end self-service analytics solution built on the GCP. These concepts are the foundation you need to understand and iterate on your self-service analytics solution.



Concept	Service	Description
GCP Project	GCP	A project organizes all your GCP resources. You'll need to use an existing project or create a new one.
BigQuery	Dataset	Datasets are top-level containers to organize your tables and views within BigQuery. You need to create at least one dataset before loading data into BigQuery.
BigQuery	Table	A BigQuery table contains individual records organized in rows. Each record is composed of columns (also called fields). A table belongs to a BigQuery Dataset.
Cloud Dataprep	Dataset	A Cloud Dataprep dataset is a reference to a data source such as a BigQuery table or a file in Google Cloud Storage. This is not to be confused with BigQuery tables.
Cloud Dataprep	Recipe	A recipe is a sequence of steps applied to one or multiple Cloud Dataprep datasets to transform the data into the desired output.
Cloud Dataprep	Flow	A flow contains Cloud Dataprep datasets and associated recipes to define a logical order for transforming the data. A flow can be scheduled to prepare data on a regular basis.
Data Studio	Report	A report with Data Studio contains various pages, each combining charts (tables, bars, pies, etc.) to create interactive dashboards.
Data Studio	Dimension	A dimension is a particular field within a table that describes the characteristics of the data (e.g. product name, region, or customer).
Data Studio	Metric	A metric is a numeric value that quantifies something relative to a dimension. For example, a quantity or a discount are metrics relative to a product name. Metrics are often expressed with an aggregation function (e.g. sum of quantities, max price, or average discount).

Here is a visual representation of these concepts and their relations to create a self-service analytics solution within the GCP.



You too can benefit from self-service analytics on GCP

Still a bit too theoretical? No problem! You can experience it for yourself following this [step-by-step guide](#) to build your first end-to-end analytics solution.

You will establish a scalable self-service analytics solution that leverages BigQuery to store and retrieve data; Cloud Dataprep to clean, combine, and create metrics; and Google Data Studio to visually report on top of your data. By exploring each GCP service deeper and iterating on these principles, you will be able to solve any requirements for your analytics in a self-service manner.

About Trifacta

Trifacta is the global leader in data wrangling. Trifacta leverages decades of innovative research in human-computer interaction, scalable data management and machine learning to make the process of preparing data faster and more intuitive. Around the globe, tens of thousands of users at more than 10,000 companies, including leading brands like Deutsche Börse, Kaiser Permanente, New York Life and PepsiCo, are unlocking the potential of their data with Trifacta's market-leading data wrangling solutions. Learn more at trifacta.com.

For Additional Questions, Contact Trifacta

www.trifacta.com | 844.332.2821

Experience the Power of Data Wrangling Today

<https://www.trifacta.com/start-wrangling/>