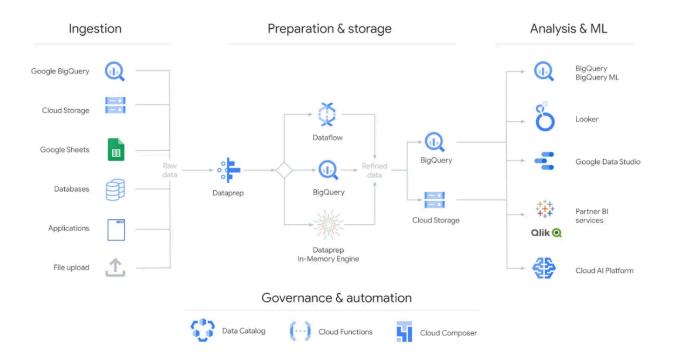# 01 Dataprep intro

1. Introduction to Cloud Dataprep. A quick introduction of Google Cloud... | by datadice | Data School | Medium 🔗

2. ~~Medium parser - Exploratory Data Analysis Made Easy with Dataprep | by daython3 | Medium (googleusercontent.com)~~ 🔗

3. ~~sfu-db/dataprep: Open-source low code data preparation library in python. Collect, clean and visualization your data in python with a few lines of code. (github.com)~~ 🔗

- ◆ Dataprep is a service that lives on GCP but is actually created by another company called Trifacta (part of Alteryx).
- ◆ Dataprep enables analysts, data specialists, and other domain experts to quickly cleanse and transform datasets of varying sizes for use throughout the enterprise.
- ◆ Dataprep key capabilities:
    1. Import from flat file, databases, or distributed storage systems.
    2. Locate and remove or modify missing or mismatched data .
    3. Unnest complex data structures.
    4. Identify statistical outliers in your data for review and management.
    5. Perform lookups from one dataset into another reference dataset.
    6. Aggregate columnar data using a variety of aggregation functions.
    7. Normalize column values for more consistent usage and statistical modeling.
    8. Append one dataset to another through union operations.

Offered by Trifacta (a privately owned software company, San Francisco)
This tool is very useful for data analysts and data scientists to visually explore, clean, and transform a dataset using a simple workflow of drag and drop transformations in a user-friendly interface instead of writing complicated code.

Serverless, fully Managed