## **Built Data Warehouse Using GCP Services**

### **Introduction:**

This project aims to build a DWH using Python and GCP services. Later, create a custom dashboard for data analysis.

### **Architecture:**

### **Graphical user interface, application Description automatically generated**

### **Technology Used**

* Programming Language - Python

Google Cloud Platform

1. Google Storage
2. Compute Instance
3. BigQuery
4. Looker Studio

Modern Data Pipeline Tool - <https://www.mage.ai/>

### **Dataset Used**

TLC Trip Record Data Yellow and green taxi trip records include fields capturing pick-up and drop-off dates/times, pick-up and drop-off locations, trip distances, itemized fares, rate types, payment types, and driver-reported passenger counts.

More info about the dataset can be found here:

1. Website - <https://www.nyc.gov/site/tlc/about/tlc-trip-record-data.page>
2. Data Dictionary - <https://www.nyc.gov/assets/tlc/downloads/pdf/data_dictionary_trip_records_yellow.pdf>

### **Dataset Model**

Diagram

Description automatically generated