**Uber Data Analytics | End-to-End Data Engineering GCP Project**

Introduction

The objective of this project is to conduct data analytics on Uber datasets, employing a diverse range of tools and technologies such as GCP Storage, Python, Compute Instance, Mage Data Pipeline Tool, BigQuery, and Looker Studio.  
  
Architecture:

A diagram of a company

Description automatically generated

Technologies Used:

* Programming Language - Python

Google Cloud Platform

* Google Storage
* Compute Instance
* BigQuery
* Looker Studio

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

Dataset Used:

The TLC trip records for both yellow and green taxis encompass details such as timestamps for pick-up and drop-off, location information for both points, travel distances, fare breakdowns, rate classifications, payment methods, and passenger counts reported by drivers.

More info about dataset can be found here:

Website - https://www.nyc.gov/site/tlc/about/tlc-trip-record-data.page

Data Dictionary - <https://www.nyc.gov/assets/tlc/downloads/pdf/data_dictionary_trip_records_yellow.pdf>

Data Model:

A diagram of a computer

Description automatically generated