# **DATA ENGINEERING BIG QUERY**

This GitHub repository contains solutions for data engineering tasks using Google BigQuery. The dataset used is bigquery-public-data.chicago\_taxi\_trips.taxi\_trips. The repository provides solutions to three specific tasks:

1. **Calculating Statistics for Trip Durations on Monday and Saturday:**
   * Calculates the average, median, and standard deviation of trip durations (in seconds) for trips taken on Monday and Saturday. The results are then compared between these two days.
2. **Finding Top Five Routes with the Most Trips in 2023:**
   * Identifies the top five routes (from the initial community area to the destination community area) with the highest number of trips in the year 2023.
3. **Comparing Average Taxi Trip Costs Based on Payment Methods in 2019:**
   * Compares the average cost of taxi trips (fare, tips, and taxes) based on payment methods in the year 2019.

Each solution is implemented using SQL queries executed on the BigQuery dataset bigquery-public-data.chicago\_taxi\_trips.taxi\_trips. The expected results are provided along with the queries.

**Task Solutions:**

1. **Statistics for Trip Durations on Monday and Saturday:**
   * Calculates the statistics for trip durations on Monday and Saturday and compares the results.
2. **Top Five Routes with the Most Trips in 2023:**
   * Identifies the top five routes with the highest number of trips in 2023.
3. **Comparison of Average Taxi Trip Costs Based on Payment Methods in 2019:**
   * Compares the average costs of taxi trips based on payment methods in 2019.

**Dataset:**

* The dataset bigquery-public-data.chicago\_taxi\_trips.taxi\_trips from Google BigQuery is used for all the tasks.

**How to Use:**

* Clone or download the repository.
* Run the provided SQL queries on Google BigQuery with the specified dataset to get the results for each task.

This repository provides clear and concise solutions to the specified data engineering tasks using Google BigQuery. Each solution is accompanied by its expected results for verification purposes.