A picture containing logo

Description automatically generated

Data Warehousing for Analytics

SQL Assignment 2

For this individual assignment, you need to use BigQuery, and 2020 Yellow Taxi trip data ﬁles available on the [NYC TLC Trip Record Data](https://www1.nyc.gov/site/tlc/about/tlc-trip-record-data.page) web site. Each month's data is stored separately on the site.

* Create a new data set in your google Big Query using the instruction previously posted and discussed in class (Google how to import data into Big query if you missed it).
* When you import make sure the schema is consistent with what is documented [here](https://www1.nyc.gov/assets/tlc/downloads/pdf/data_dictionary_trip_records_yellow.pdf).

1. For the first four months of 2020, ﬁnd the total fare amount for yellow taxi trips taken, charged at the standard rate, with 2 or more passengers who paid with a credit card or cash.
2. What was the average cost (fare amount only) per mile of a yellow taxi standard rate ride in 2020 traveled more than 8 miles but less than 20 miles (excluding trips made to airports i.e. RateCodeID 2, 3 and 4)? (Result should be 1 record)
3. Which day of the week in 2018 has the lowest number of single rider trips on average ? (Result should be one record)
4. ETL: Create new dataset from **Jan 2020** data which includes the following columns
   * Date of pickup (extracted from tpep\_pickup\_datetime)
   * Time of pickup (extracted from tpep\_pickup\_datetime)
   * Time of drop-off (extracted from tpep\_dropoff\_datetime)
   * Time of pickup (extracted from tpep\_pickup\_datetime)
   * trip\_distance
   * RatecodeID
   * fare\_amount
   * total\_amount / trip distance
   * the table should include only trips that had:
   * Zero tolls
   * Less than 30 miles

\*If you need to clean or filter any other value (e.g. unexpected or null values as we discussed in ETL session) do it before you run the final queries. Make sure you explain what you did and why you did it.

\*\*Follow one of the methods to write your query results onto a new table in BigQuery.

<https://chartio.com/resources/tutorials/how-to-create-a-table-from-a-query-in-google-bigquery/>

1. Write a paragraph to answer to explain:
   * What you learned from the assignment.
   * What did you find particularly challenging
   * How long it took from you to complete it
   * What would you change about the assignment if you could.

**Additional Notes:**

 This is an individual assignment.

 Assemble all of the responses to the questions in one MS Word.

 For the SQL Queries, format your query in Courier New font (at least 12 point size). Paste in the results of the query into MS Word. So, I need to have both the code and the results.