Change Log:

1. Downloaded/Extracted and imported data from November of 2021 to April of 2022. Encountered issues importing the most current data, likely due to size. I continued the project using three months of data. I acknowledge this limits the scope and time relevance of the data.

2. Added column for total ride time formatted in HH:MM:SS

3. Added column for the day of the week the ride happened in a number format from 1 to 7, with Sunday equalling 1 and Saturday equalling 7

4. Trimmed whitespace and removed duplicated

5. Removed any entries with less than 30 seconds of ride time.

    a. These rentals are likely due to client entry mistakes or opted to not rent.

6. Calculated average ride length, maximum ride length and the most common day of the week for a rental.

    a. I noted that some rides spanned over many days after calculating the max ride time. I considered eliminating them but I feel they are also considered valid data as opposed to rides less than 30 seconds.

7. Created a pivot table to show the average ride time per rider type (member versus casual), for each day of the week.

    a. Used to get a general idea of trends, to be analyzed further in SQL

8. Created a pivot table for unique to show number of rides per rider type, including what kind of bike they rode

    a. Used to get a general idea of trends, to be analyzed further in SQL

9. Added column labeled “trip_time_sec” and used “trip_time” to convert that column from a TIME datatype to and INTEGER datatype to be more usable in SQL.

10. Created a new table in BigQuery to include all files with columns “ride_id”, “rideable_type”, “started_at”, “ended_at”, ride_time_sec”, “day_of_week” and “member_casual”. (See “SQL Queries” Document)

11. Queried SQL to return ride count by day of the week, for each rideable type per member type, for every month in the data range. (See “SQL Queries” Document)

    a. Results were copied into Sheets for further analysis

12. Performed quick calculations to confirm data found in spreadsheets