

Cyclistic Bike Share Capstone Project (Excel & SQL)

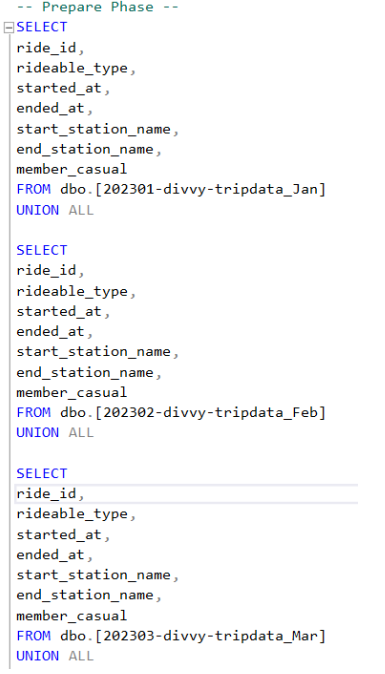
Thanks to the Google Data Analytics Professional Certificate program for giving me the opportunity to showcase my skills in this capstone project.

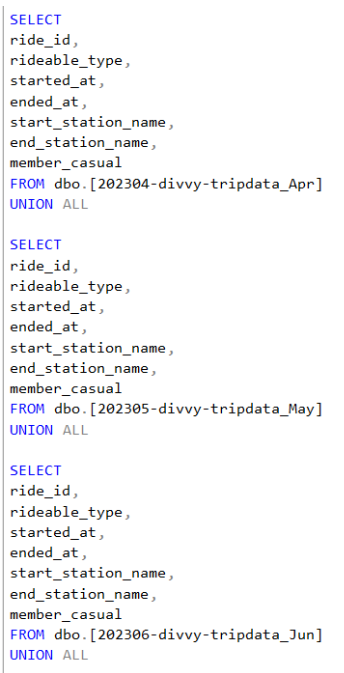
Second Version (SQL)

**Documentation of any cleaning or manipulation of data**

**Prepare Phase**

In this phase, I uploaded datasets of cyclistic bike-share for 2023 year into Microsoft SQL Server Management Studio (SSMS) in the type of csv files. The database called as “cyclistic\_bike\_share”. After that I wrote down a query for joining all datasets of 12 months into a new table.





A screenshot of a computer code

AI-generated content may be incorrect.

A screenshot of a computer code

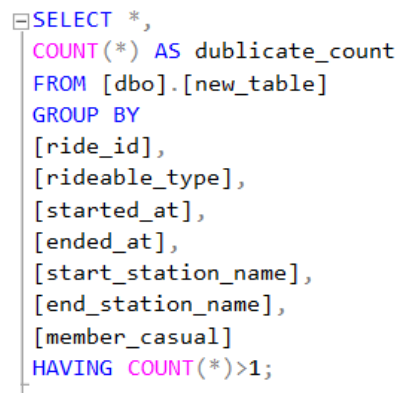
AI-generated content may be incorrect.

**Process Phase**

First, I identified null values in a new dataset that contains all 12 months’ data by using SQL query. The result showed that SQL identified about 1 million rows that include null values. Then I decided to write a SQL query for deleting those null values. After deleting null values in the new dataset, I identified duplicated values by using SQL query. Fortunately, SQL showed that there were not any duplicated values.

A screenshot of a computer program

AI-generated content may be incorrect.



Analyze phase

First, I identified the number of individual riders and their membership type by SQL query.

A screenshot of a computer program

AI-generated content may be incorrect.

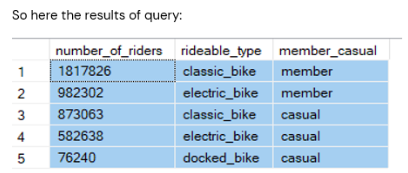
A blue and black text

AI-generated content may be incorrect.

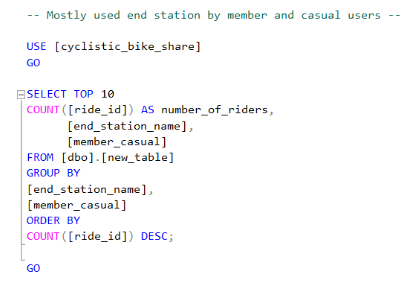
Secondly, I queried to determine the number of riders for using different types of bicycle and for what type of membership they were registered.

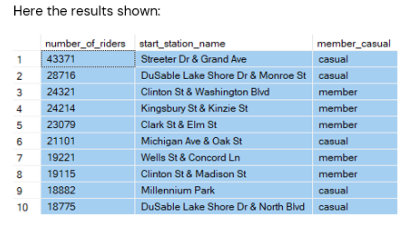
A screenshot of a computer code

AI-generated content may be incorrect.



Thirdly, I identified top 10 station names that riders started their bicycle trip by SQL query.

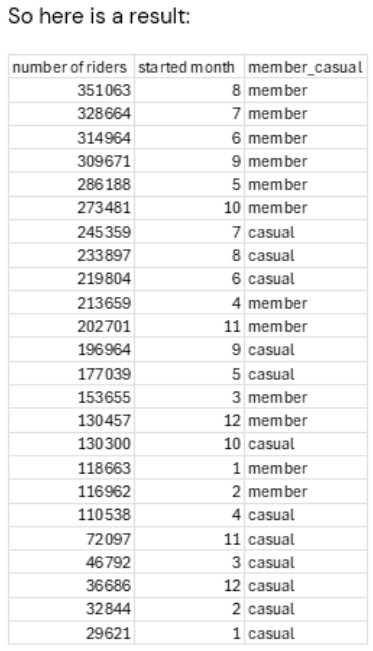


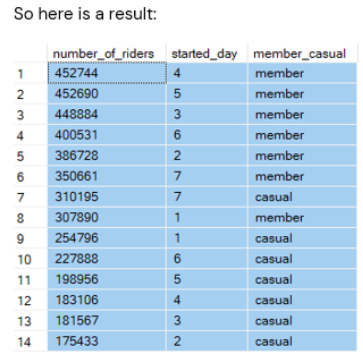


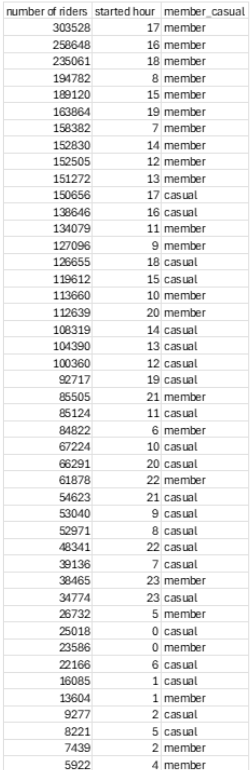
Then I implemented time analysis for started month of bicycle trips by membership type.

A screenshot of a computer code

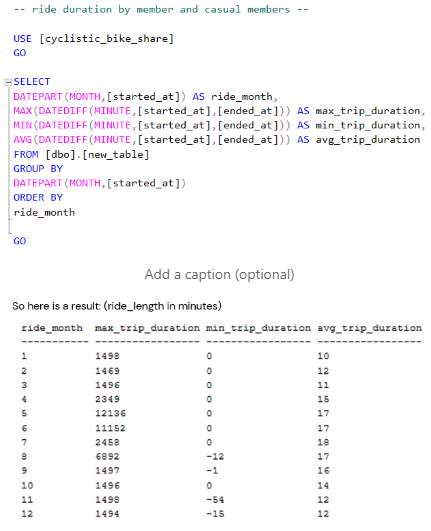
AI-generated content may be incorrect.







At the end, I determined a new column that shows the ride length of every bicycle trip. For this column I calculated maximum, minimum and average values based on each month.



Key findings

 The table above shows the overall trend of ride\_length (ride duration) of casual riders and annual members for all months of 2023 year. As you can see all riders had reached its peak in May, June, July, August and September.