Your task in this assignment is to aggregate the data found in the Citi Bike Trip History Logs to build a data dashboard, story, or report. You may work with a timespan of your choosing. Optionally, you may merge multiple datasets from different periods. The following are some questions you may wish to tackle, especially if you are working with merged datasets. Do not limit yourself to these questions; they are suggestions for a starting point. Be creative!

How many trips have been recorded total during the chosen period?

500,828 Trips

By what percentage has total ridership grown?

Ridership grew by 55% in January, 23% in February, 38% in March, 40% in April, and 5% in May YoY.

How has the proportion of short-term customers and annual subscribers changed?

The proportion of Short-term riders to Annual Subscribers increased from 6.26% in 2018 to 6.52% in 2019.

What are the peak hours in which bikes are used during summer months?

8 am and 6 pm are the most popular times during the summer months.

What are the peak hours in which bikes are used during winter months?

Same hours as summer.

Today, what are the top 10 stations in the city for starting a journey? (Based on data, why do you hypothesize these are the top locations?)

See visual. “Start Journey T”. They are probably in the downtown areas of town where most commuting occurs.

Today, what are the top 10 stations in the city for ending a journey? (Based on data, why?)

See visual. “End Journey T”. They are probably in the downtown areas of town where most commuting occurs.

Today, what are the bottom 10 stations in the city for starting a journey? (Based on data, why?)

See visual. “Start Journey B”. They are probably in the rural suburban areas of town and mostly for leisure.

Today, what are the bottom 10 stations in the city for ending a journey (Based on data, why?)

See visual. “End Journey B”. They are probably in the rural suburban areas of town and mostly for leisure.

Today, what is the gender breakdown of active participants (Male v. Female)?

74% M to 21% F

How effective has gender outreach been in increasing female ridership over the timespan?

MoM comparing 2018 to 2019 there has been a marked increase in Female Riders. It clearly has had an impact.

How does the average trip duration change by age?

Seems like Trip Duration is highest from ages 16 to 21 and declines thereafter. This is likely because younger people have less money.

What is the average distance in miles that a bike is ridden?

Average distance in miles per day by all riders is 104,826 miles.

Which bikes (by ID) are most likely due for repair or inspection in the timespan?

See chart for top 10.

How variable is the utilization by bike ID?

Not sure what this is asking….

Additionally, city officials would like to see the following visualizations:

A static map that plots all bike stations with a visual indication of the most popular l ocations to start and end a journey with zip code data overlaid on top.

Unable to locate zips. See map.

If you're working with a merged dataset: a dynamic map that shows how each station's popularity changes over time (by month and year) -- with commentary pointing to any interesting events that may be behind these phenomena.

N/A

Lastly, as a chronic over-achiever:

Find at least two unexpected phenomena in the data and provide a visualization and analysis to document their presence.

N/A