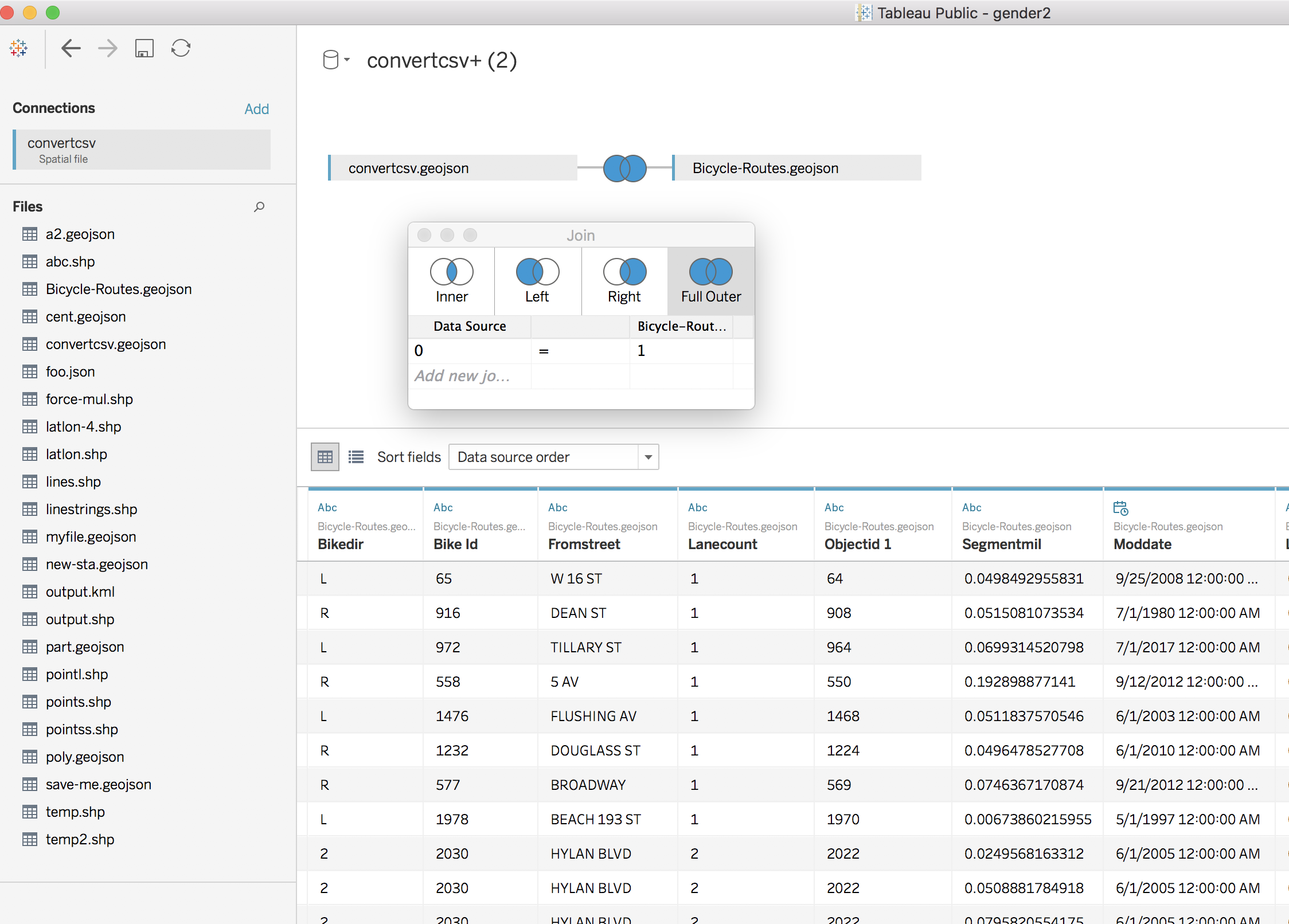
# Tableau Assignment - Citi Bike Analytics

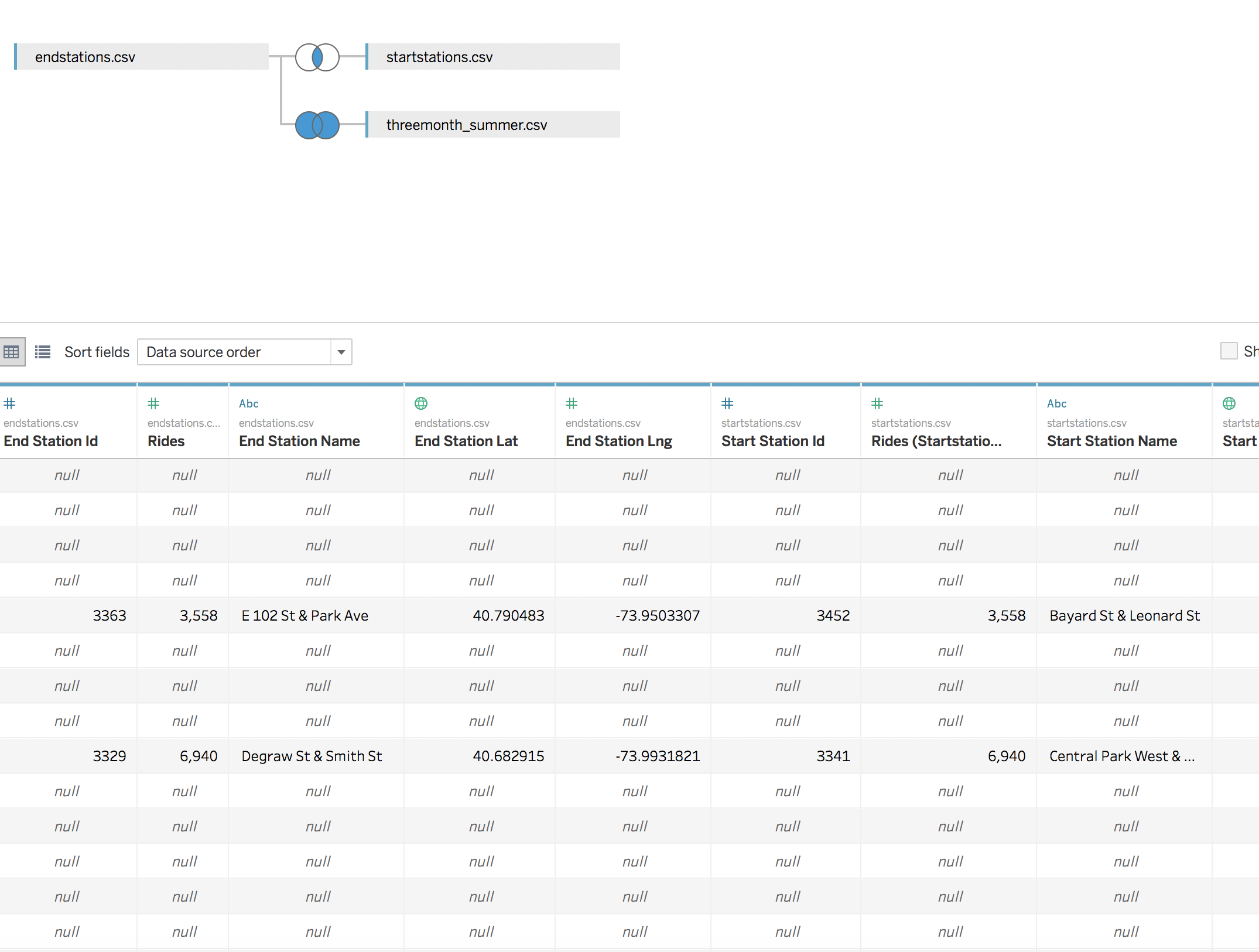
Ghassan Aleqabi

Data Exploration and Joining-

Including the data downloaded from https://www.citibikenyc.com/system-data) webpage.I also downloadeed the following file Bicycle-Routes.geojson from https://data.cityofnewyork.us/Transportation/Bicycle-Routes/7vsa-caz7

and used csv2geojson to convert lat and lon of bike stations to geojson. An outer join allowed a connection between the two files that facilialted creating maps of bike routes and stations locations.





\*\* Please note for explanations consult the story file

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

I chose June, July and August of 2017; there are more than 5500K trip in those three months.

\* By what percentage has total ridership grown?

Ridership is increasing within less than 4% and mostly during July (weather related)

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

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

It seems a **bimodal** distribution with two peaks describe peak hours; 7-9 in the a.m. and

4 to 6 in the p.m. Also the weekend show a small increase in the bikes usage (Figure attached)

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

Winter months were not considered in this analysis.

\* 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?) I believe those stations are within the heart of the

New York City boroughs.

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

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

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

Those station are visited for reason other than tourist or recreational and probably more industrial locations than parks.

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

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

\* How does the average trip duration change by age?

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

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

Bikes with ID between 21692 and 24849 potentially out of service during this period.

\* How variable is the utilization by bike ID?

Bike ID as a variable is limited in utilization as it describe the bike only and technically it is a

Static variable does not change during bike rental operations.

\*\*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 locations to start and end a journey with zip code data overlaid on top.

<https://public.tableau.com/profile/ghassan2357#!/vizhome/Book1_39787/Dashboard4?publish=yes>

\* 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.

<https://public.tableau.com/profile/ghassan2357#!/vizhome/Book1_39787/Dashboard5?publish=yes>

\*\*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.

Potentially, both females and senior citizens will keep stacking the number of bike ridership.