Citi Bike Usage Analysis

by Lisa Greer

Overview:

- Citi Bike is a popular bike sharing company with a thriving location in NYC. Citi Bike is currently receiving customers complaints about bikes not being available when they want them at certain locations. Citi Bike has also received feedback that stations are too crowded for customers to return their rented bikes.

Purpose:

- Create a dashboard that aims to provide helpful insights on the expansion problems Citi Bike currently faces.

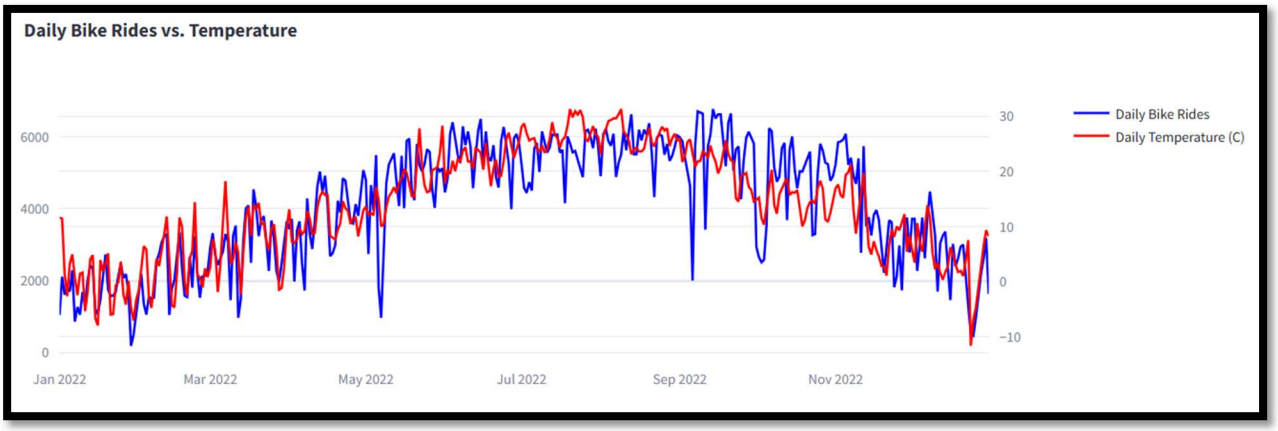
Objective:

- Analyze the potential reasons behind these customer complaints by looking at:
    - Weather and Bike Usage
    - Trips by Season
    - Most Popular Stations
    - Maps with Aggregated Bike Trips

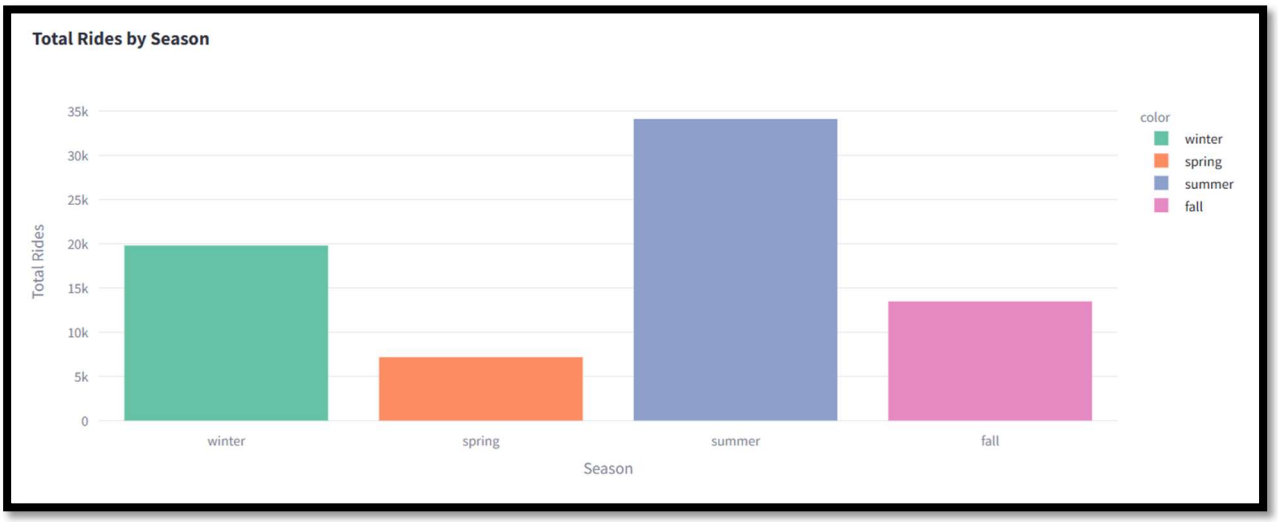
Tools, Skills, Methodologies:

- A number of open-source data sets from Citi Bike and NOAA were used to complete this analysis.
- Python was used to clean, manage, and wrangle the data so that it could be analyzed using statical techniques and visualized. It was then applied to Streamlit to create an interactive dashboard using a Python script.

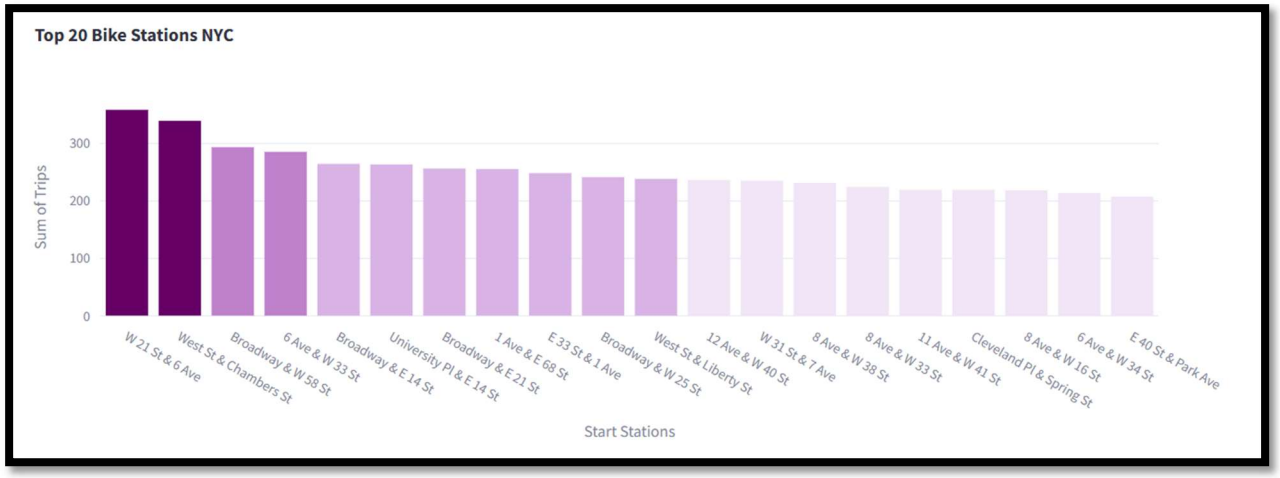
The Findings:



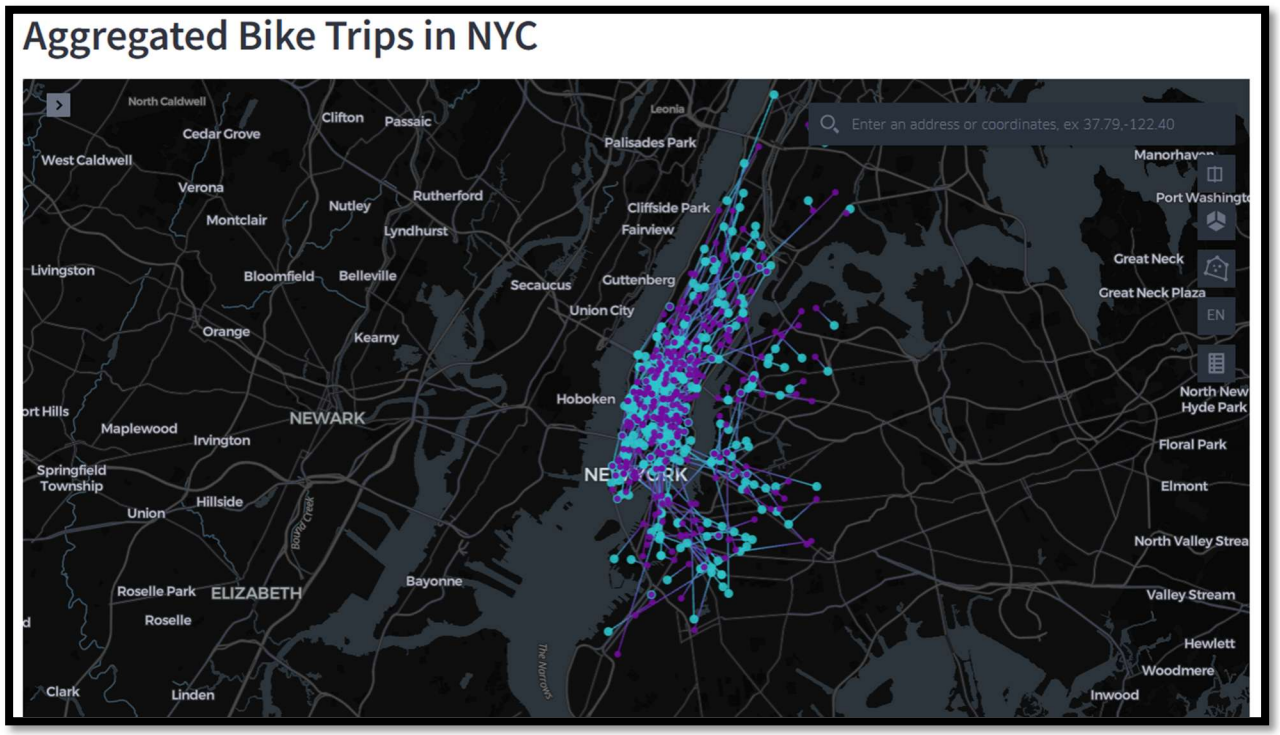
There is a correlation between temperature fluctuations and their relationship to the frequency of bike trips taken daily. As temperatures plunge, so does bike usage. This insight indicates that capacity may be more of an issue in the warmer months, from May to October.



This chart shows a surprising trend that the shoulder seasons, Spring and Fall, have less riders than the Summer and Winter months. This warrants further investigation as perhaps it correlates with school schedules or popular tourism times. After a cursory level of research; the Summer and Winter months correlate with popular tourism times. NYC's population has also been steadily increasing since the decline during the COVID-19 pandemic. Citi Bike may need to look at how their business strategy adjusted/didn't adjust after the pandemic ended.



From the bar chart it is clear that some start stations are more popular than others - in the top three we can see W 21 St & 6 Ave, West St & Chambers St, and Broadway & W 58 St. There is a big difference in usage between the 1st station and the 20th station. This indicates clear preferences for the leading stations.



As a still picture it is not as clear, but by creating an interactive map, the following insights could be found:

The most popular start stations are W 21 St & 6 Ave, West St & Chambers St, and Broadway & W 58 St. While having the aggregated bike trips filter enabled, we can see

that W 21 St & 6 Ave and West St & Chambers St are popular start stations and account for the most commonly taken trips. However, this does not hold true for Broadway & W 58 St.

The most common routes (>6,000) are in the heart of the city surrounding the tourist areas like the Empire State Building, Union Square, Grand Central Station, Broadway, etc. It is worth noting, using the heatmap layer, that start stations are denser in the center of city while end stations have fairly equal density inside and outside of the heart of the city. So, adding more stations in the center might aid overcrowded or empty stations.

Conclusion:

- Add more stations to the locations in the heart of the city. A limitation of this study is how many bikes are able to be parked in each station. Perhaps rather than adding more stations, adding to current stations capacities would provide an adequate response.
- Ensure that bikes are fully stocked in all these stations during the warmer months in order to meet the higher demand. Reduce supply during other seasons to reduce maintenance costs in proportion to demand.
- Reevaluate how Citi Bike adapted to changes after the COVID-19 pandemic. If supply was never restored in full, going back to the pre-pandemic volume metrics may prove helpful.

Limitations:

- This database was large in size and I was limited by my computer's computational power. Stratified samples of the dataset were used to draw these conclusions. It would be helpful to be able to look at a larger portion of the data.

Next Steps:

- Evaluating current stations in NYC by bike capacity would provide an extra level of insight when implementing future changes.
- Looking at competitors and how Citi Bike's locations compare to theirs could provide further insight into parts of NYC with gaps in service.

Follow – Up:

- For more details about the project and data here are links to the GitHub repository, the streamlit app, and my presentation of the information:

- GitHub Link
- Streamlit App
- Loom Presentation