A group of blue letters

Description automatically generatedNew York Citi Bike Analysis

Seasonal variation in Citi Bike Usage in New York City

This analysis aimed to discern whether there is a profound impact on Citi Bike usage in New York City, across different seasons. Bike usage across 2 months of the year (January and July) were used as markers for their seasons. Various aspects such as usage throughout the day, the length of trips, and the usage and locations of various stations were explored.

Firstly, data was obtained from Citi Bike via their public website <https://citibikenyc.com/system-data>, in the form of CSVs containing trip figures for January and July trips. 2023 was the chosen year, as it is more recent, the data it contains would be more pertinent.

The CSVs were processed and cleaned in Jupyter Notebook, before being exported to Tableau for visualization. The cleaning process consisted of concatenating the tables, removing and fields with null data, converting data types, and creating a new field for trip duration. This was all done in Python as it provided a greater degree of functionality when it came to data manipulation.

As for the visualization aspect, Tableau provided a much greater degree of freedom when it comes to displaying facts and figures.

As displayed by the “Seasonal Bike Usage” dashboard, it is evident that summer is clearly the more popular month for Citi Bike rides – almost twice as many bike rides took place in July, when compared to January. The most likely explanation for this could be better weather in summer. As for daily variances, there were some peaks in usage, for both seasons: 07:00 – 09:00 am, and 16:00 – 17:00 pm. Both coincide with office hour traffic. The “Trip Duration” dashboard indicates that trips are generally short (between 1 – 10 minutes), and there is no variation between seasons.

Citi Bike stations appear to be clustered methodically throughout the city, with some areas having almost no stations at all. The overall density increases greatly the closer we move to lower Manhattan, as evidenced by both station location graphs. A possible explanation for this could be income, as there seems to be a slight correlation between median household income and station density throughout the city.

As the 2 maps show, there are stations clustered throughout the city, but a few are significantly more popular than others, as signified by the large blue circles on the maps. The top 10 start stations also appear to be highly popular end stations, signifying that most Citi Bike trips occur between a select few locations. Midtown appears to be a popular starting point, with lower Manhattan and Central Park appearing to be popular end points, possibly due to their functions for work and leisure.

There does not appear to be a significant impact from seasonality, as usage is consistent across both seasons, save for some exceptions like Central Park.