Greg Spahlinger

4/21/21

Use Patterns of Citibike before and after the Pandemic in Jersey City.

Introduction

Citibike is a bicycle rental system in the New York City area. The system allows users to pick up and return bikes at various locations in the New York City area, as needed. Renters can pay for the bikes through two separate payment models: “customers” by a pass for one or several days, while “subscribers” purchase a membership by year. The present report focuses on usage patterns in the Jersey City area for the years of 2019, 2020, and 2021. Jersey City was chosen as a focus as it’s files were conveniently separated from the rest of the data, and were a manageable size.

Data Manipulations

The data was obtained from the Citibike System Data website (<https://www.citibikenyc.com/system-data>) in the form of .csv files. One file was obtained for each month of use. Most of the exploration and analysis was conducted using Tableau public, but several manipulations were carried out using Python (including the numpy and pandas libraries). The .csv files were read into a Jupyter Notebook. The files were concatenated into a single dataset, and several columns were added and adjusted. The starttime and stoptime columns were converted into datetime format (these had to be reformatted manually in MS Excel before that could happen, due to an unusual starting format), and a column was added for month, year, and year\_month, for convenience. After this the Haversine formula was used to calculate the displacement of the bike in miles for each trip (i.e. the distance between where the bikes were obtained, and dropped off). Two filtrations were then carried out – trips with individuals claiming to have been born in 1920 or before were removed, as these probably represent unreliable responses, and trips with a displacement of more than 100 miles, as these are impossible given the range of citibike, and represent faulty geocoordinate values. Following this the data was written out into a single .csv, and imported into Tableau for exploration and analysis.

Findings - General use patterns and pandemic impacts.

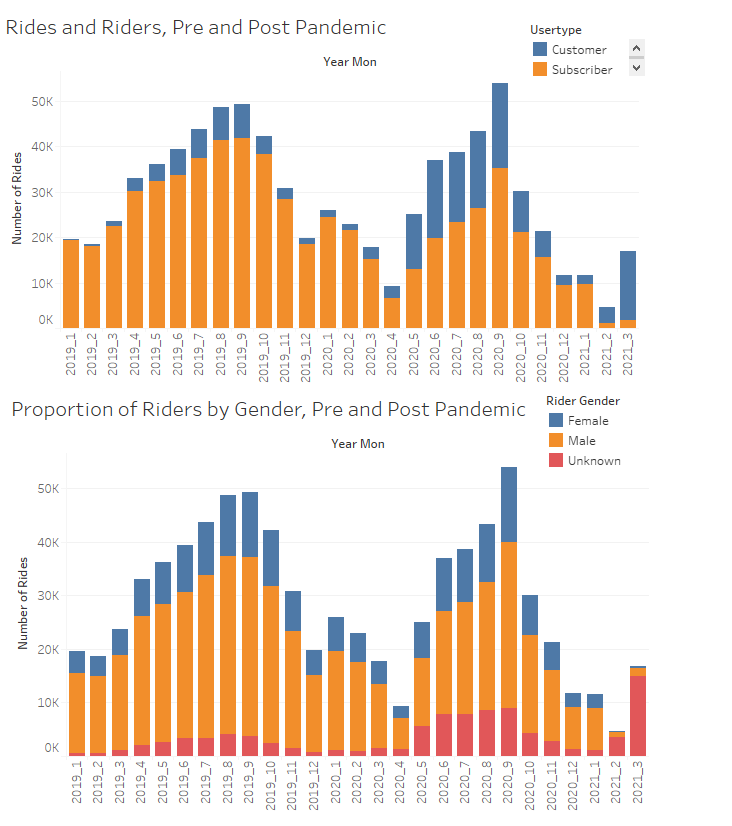


Figure 1 – Riders by Month, Gender Demographic and Payment Model

The first trend we can see in the data above is that the pandemic had some negative impact for Citibike in Jersey City. The months of 4/2020, and 12/2021, 1/2021, 2/2021, 3/2021 all seem anomalously low. The number of subscribers also appears to be going down since 2/2020. The good news is that the number of people paying for the bikes as customers appear to be going up in the same time period. The gender balance of people riding before and during the pandemic appears relatively constant, with an anomalous increase in unknown gender toward the end of the data range. That increase is difficult to understand, but may be a result of the high proportion of customers using the bikes during the same time.

Findings – People in Jersey City Mostly Stay in Jersey City

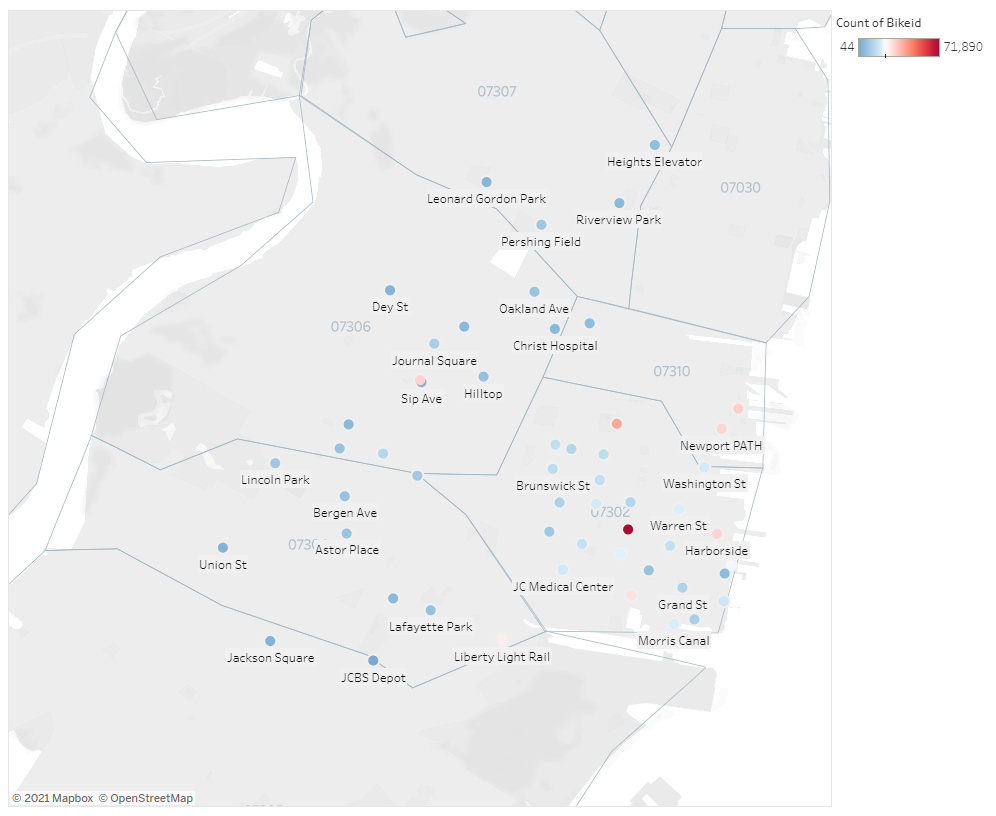


Figure 2 – Popular starting points in Jersey City.

As expected, riders in the Jersey City area pick up bikes in Jersey City. The most popular places to pick up a bike are mostly along the waterfront. People who pick up the bikes return them in a variety of places, some of which are in Jersey City, and some of which are in New York City. In both cases, Grove St Path is the most popular station, accounting for tens of thousands more rentals and returns as Hamilton Park, the number two station.

Findings – The bikes are most popular during commute times.

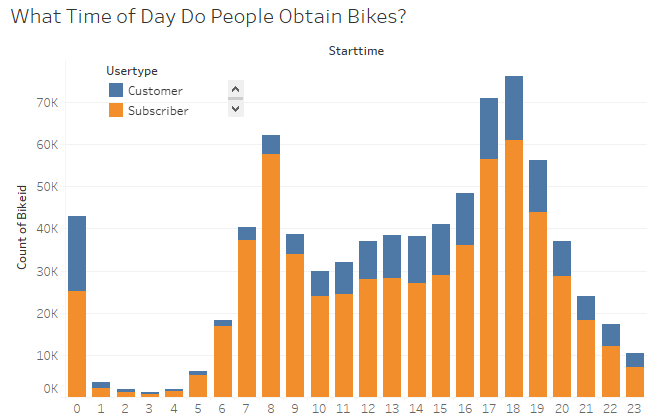


Figure 3 – Bikes are most popular during commute times

The data above show that the most popular time for bike rentals is in the morning around 8 am, and in the afternoon around 5 or 6 pm. A spike also occurs at 12 am, which appears to be an anomaly. This spike is also localized to the spring time (January through April – see full Tableau story for figures) and probably reflects either some kind of error or loss of data, or some kind of midnight biking event. The early mourning spike is mostly composed of subscribers, while customers show up more heavily in the afternoon hours. The logical explanation for the data is that a reasonable number of people who subscribe to Citibike do so in order to use the bike for their commutes.

Conclusion

In summary, several trends have been observed from the data obtained from Citibike. Jersey City riders are mostly male and mostly subscribers. While the number of subscribers is dropping, and the number of customers increasing, the gender balance is relatively constant. People use citibikes to get around Jersey City, and a lot of them appear to be using the service to get to work.