

Abigail Herrup

HW 20: Citi Bike Data Analysis with Tableau

Background & Overview:

In New York, you will commonly see blue bikes being rode by many. These blue bikes are known as a Citi Bike. Citi bike is the countries largest bike share program. These bikes are designed for a quick trip that is convenient (can find the bikes everywhere) and affordable. Currently around NYC there are about 15,000 bikes available for rent. To rent these bikes you use a phone app to either purchase a short-term bike pass or one can become an annual member and use the bikes as often as one wishes.

Data Cleaning:

For the purposes of using the data in Tableau, I downloaded all 12 months of data from the Citi Bike website, which provides data for all of its years of operation for free. From these 12 CSVs I used Python to take a random sample from each of the 12 months to combine them into one CSV that is under 100KB (in order for it to work with Tableau).

Data Analysis:

From the creation of my Tableau story, I was able to analyze a few different aspects of the Citi Bike Data. First, I created a map of the bike stations, with an additional layer in the background showing the population by Census Tract. The dots on the map represent the number of rides originating from said station. From viewing the map, it is clear the most popular place to rent the Citi bike is in lower Manhattan. When you look at the population numbers compared to the numbers of bikes rented, there is a clear correlation. The more densely populated areas of New York City have a higher number of bikes rented.

The next thing I looked at was seasonality. I wanted to know when it is most popular to rent the bikes (season and time of day). When looking at the map, it is clear that while in the colder months there are less bikes being rented, however, lower Manhattan still seems to remain the most popular area to rent these bikes. When looking at the top 30 rental stations, throughout all seasons, Pershing Square North remains the most popular station. When looking at the most popular times to rent a bike, it looks like it corresponds to work commute times as the most popular time is 5pm and second most popular is 8am.

When looking at the demographics who rent these bikes, males rent the bikes about 2x as much as females. Also, the majority of the customers seem to be subscribers, with very few just paying for a short-term pass.

For trip duration, the ages where people ride for the longest seems to be 52 and 86, taking rides on average for just over 30 minutes.

Data Limitations:

The main limitation I found from this data was there was some outliers in the data that looked like they were incorrect data points (i.e., it had riders that were aged over 130 years old).

Future Work Recommendations:

Given more time and data, it would be interesting to compare the trends of bike usage with weather trends. It would also be interesting to see how Citi bike compares with other bike rental services or scooter rental services that have become increasingly popular over the last few years.