**Report of the Analysis of Citi Bike Trip Data for Jersey City (25th March 2020) and New York (5th January 2021)**

**by**

**Nalishebo Meebelo**

**Introduction**

This report presents an analysis undertaken as part of my Bootcampspot homework on the topic: Tableau. For this homework assignment, I evaluated two datasets from New York Citi Bike data – one representing bike trips for 25th March 2020 in Jersey City and the other, bike trips for 5th January 2021 in New York. The two cities face each other over the Hudson River. I have taken cognizance of the fact that the analysis is for different dates in the two cities and therefore the comparison may be limited in its deduction of facts.

Citi Bike is a privately owned public bike sharing system[[1]](#footnote-1) and, the following features can be deduced from the analysis of the two Citi Bike datasets. Please see the visualisation of this analysis can be found at:

<https://public.tableau.com/profile/nalishebo.meebelo#!/vizhome/CitiBikeTripsComparisonJerseyCityandNewYork/CitiBikeTripsComparisonJCNY?publish=yes>

To assist in my analysis, I used Tableau Public, and began by inner joining the two datasets. I also used some calculations for change ‘gender’ to ‘male’ and ‘female’, and the ‘birth year’ to ‘age’. I also changed the start time from ‘year’ to ‘hour’ to assist with generating better visualisations for the story.

**Station Locations**

It is clear from the analysis that there are more station locations in New York, than there are in Jersey City. Both cities overlook a water feature (the Hudson River), and the famous Statue of Liberty, which makes the two cities potential tourist locations. As noted in my research about the cities, a number of other tourist locations can be found in these cities. This makes cycling convenient for those who want to enjoy the scenery or travel short distances possibly a cheaper form of transportation, and in a less congested manner. It is also noted that while cycling is healthy, it can also be risky compared to other forms of transport[[2]](#footnote-2).

**Station Popularity**

With regard to station popularity, Grove St PATH is the most popular starting station in Jersey City, followed by Sipp Ave and Hamilton Park, respectively (highlighted in shades of green on the map). However, you will note **an unexpected phenomenon** which is that the popularity of the ending stations is not necessarily the same as the start stations. In order of popularity Grove St PATH and Sipp Ave remain at the top, while Hamilton Park is replaced by Harborside Station. In New York City, the popular starting station is W 21 St and 6th Av, followed by E 17th St and Broadway and Grand Elizabeth. These three stations are equally, the popular ending stations.

One **other peculiar phenomenon** is that while we see a greater number of Citi Bike stations in New York, and fewer stations in Jersey City as mentioned in the above section, the lowest count of starting station trips in New York is 6 at White Street and John son Ave, while the lower count of trips in Jersey City is 1,517 at Communipaw and Berry Lane. The highest count of trips for start stations in New York is 4,813, while for Jersey City it is 68,264. Individual stations in Jersey City seem to have more clients than New York stations.

Note in the maps that the dark green and large size of the marker (i.e., a circle or dot in this case) reflects the most popular station (in count of trips), while dark red and a small size dot reflects less popular stations comparatively. Lighter shades of red and green fall in between in terms of trip counts.

**Trip Duration by Age**

In terms of trip duration against the ages of customers and subscribers, it is interesting to note that the ages 16-19 seem to have cycled longer in the data set, doing better that other ages including the 22 - 23-year-olds and the 36-year-olds. In addition, I noted that the trip duration and the count of trips data seem to be the same when I later generated the *No. of Trips by Age* visualizations for both cities. With regard to New York, I found that the highest age group in terms of trip duration were the 36-year-olds. This group was in 5th place in the data for Jersey City.

**User Type**

In two datasets, the variable *“Usertype”* appears highlights more short-term customers than annual subscribers. In terms of **gender,** we note that there a more male short-term customers than females, and we also note the same trend for annual subscribers. It is also noted that the trip duration is higher for annual subscribers in comparison to short-term customers, also signaling that annual subscribers use the Citi Bike service more that the short-term customers. This brings me to question of whether this service is used more by tourists or regular residents of the two cities.

**Start Time and Stop Time**

The analysis shows that that the trends for the start time and stop time against count of trips are not too farfetched in both cities. They patterns are similar as can be seen when they are placed side by side in the .tbwb. However, the curves for the two cities are not similar as is witnessed in the visualization titled “Starttime by Trips” for each city in the first Dashboard.

Peak starttime for Jersey City is 8hrs followed by 17hrs to 18hrs, while for New York peak time was 15hrs. This may signify a different type of demand for clients in each of the cities at given times.

1. https://en.wikipedia.org/wiki/Citi\_Bike [↑](#footnote-ref-1)
2. https://theconversation.com/cycle-walk-drive-or-train-weighing-up-the-healthiest-and-safest-ways-to-get-around-the-city-100238 [↑](#footnote-ref-2)