

Citi Bike Analysis

For this specific example, I decided to evaluate one month of Jersey City Data (May 2020). The reason I decided to use this data sample is it was easy to use and I wanted to set a benchmark after the COVID-19 breakout.

I am in advertising, particularly, media strategy. In my field, it's important to set a baseline and then analyze. With the data provided, it gave me enough information to tell a story about most common Station IDs, rider demographics and if I were to work for Citi Bike, how to market accordingly to increase growth amongst riders in the Jersey City area.

The first analysis I did was looking at stations, and which ones were the most common to both start and stop. There wasn't much difference between the two but this information could also help me decide which stations I would recommend to potentially cut and which ones have growth opportunities.

The second analysis revolved around the demographics of riders, both customers and subscribers. I noticed that these audiences use the Citi Bikes differently. Customers on average 3x longer rides than subscribers and their peak times of when they start their trips are a little bit earlier than subscribers. I also noticed that customers are a little bit younger according to their birth years than subscribers. If we removed the statistical anomaly of the majority of customers being born in 1969 (which more than likely isn't true), they are about 3-5 years younger than subscribers. As somebody in advertising, I would take that data and come with a media plan of how to target customers and then do a separate media plan on subscribers. They are different audiences and I would form an analysis of how and where to place advertising to reach these people.