Tableau Challenge

Christopher Denq, 1/5/22

**Summary**

**Dataset**: Citi Bike Data, 12 datasets totaling 650k data points over 12 months in 2021.

**Project Conclusion**: We can assess and iterate upon three underlying patterns / phenomena found with our Citi Bike data:

1. Subscription members account for ~80% of all rides, with each ride averaging between 0-2 hours.
2. The most popular end stations are densely concentrated in the 07302 zip code.
3. There are certain “gateway” stations into all the other stations in the Jersey region.

**Detailed Report / Story Walkthrough**

Chart

Description automatically generated

There seems to be a major density of the number of rides concentrated to only a few stations. This exploratory graphs prompts the search for why certain stations are overwhelmingly popular compared to others.

Chart

Description automatically generated

When we graph the top five co-occurrences of start-to-end stations, we find that a vast majority of rides at these top 5 stations are rides that start *and* end at themselves. This means that a significant portion of rides are local and generally much shorter. This is potentially telling of the behavior / purpose for why users use our bikes.

Map

Description automatically generated

Graphing the stations on the map, we find the most popular stations are densely located within the 07392 zip code. This means that a vast majority of our users start and end their bike trips within the same *area*, not just within certain bike stations.

Overlaying the streets and points of interest, we find some possible explanation: this region is an entertainment center and/or tourist location!

Chart, bar chart

Description automatically generated

By graphing the total number of rides against time, we see that the bulk of our rides occur within the summer and fall seasons, which correlate heavily with vacation time and pleasant weather.

With all this in mind, we can make an educated guess about the behavior of our bike consumers: they are likely renting bikes to tour and visit the local entertainment centers and businesses!

Let’s dive a little deeper into the profile of our users.

Chart, box and whisker chart

Description automatically generated

Graphing the trip duration via a boxplot, we can highlight a divergence in the user types: on average, customers spend longer on their trips while subscribers (repeat customers) spend shorter times in their trips. Additionally, we see very large range in the trip durations, so large that the Trip Duration axis had to be graphed in log scale. While the bulk of the trips are under 2 hours, there are several trips that range from 3 – 211 hours!

This affirms that the bulk of our riders tend to bike locally, but also highlights the fact that customers do view CitiBike as a potential competitor to longer-distance rides as well.

Chart, bar chart

Description automatically generated

When we breakdown the details of the boxplot, we reaffirm that subscribers make up the bulk of our rides but also discover that there are certain end stations that receive the bulk of the “long-distance” rides.

Map

Description automatically generated

Graphing the stations by average trip duration on the map this time, we are able to highlight several key stations that serve as “gateways” into the area we predominantly service. These stations help outline the direction in which future stations could be added, as to better draw in new riders.

**Actionable insights**:   
1. Given the dominant use / behavior of our biggest customer segment (subscribers), we can better integrate our bikes with the local businesses and entertainment options:

* Create and advertise a local list of must-see’s that would encourage users to get and stay on bikes
* Partnerships with eco-friendly businesses: discounts for CitiBike users while CitiBike includes these businesses on their list of tourist areas
* Integrate with Postmates/UberEats for easier delivery options
* Install / relocate end stations closer to the main attractions, such as having one on the street next to a museum

2. Run a marketing campaign that promotes manual locomotion:

* Given the closeness of all the entertainment sectors, emphasis the convenience of bike riding (easier movement and parking) and environmentally sustainable impact over cars/buses.
* Can also emphasize the low-level and cheaper commitment to a bike, as people flying into the region or sailing over the river from New York do not need to rent a car to get the full experience.

3. Branch out in the direction of the “gateways”:

* As mentioned before, there are certain end stations that receive a higher proportion of longer trips, suggesting that they are receiving longer-distance rides from outside the area and/or there are fewer bike stops around to deposit the bike
* As a way to bring more users to the local entertainment in Jersey, we can grow the number of stations in the direction of the most populated long-distance stations, as these stations will be less congested and more accessible to those who want to enter.