Analyzing Changes in Citi Bike Trips after the Introduction of E-Bikes

For my project, conducted data analysis on how Citi Bike trips have changed after Citi Bike introduced about 200 e-bikes on August 20th, 2018. To do so, first I grouped the docks into the neighborhoods in which the docks are located for the period before the introduction of e-bikes and the period after their introduction. I ran this analysis twice to control both for Citi Bike's year-over-year growth and seasonality patterns. I broke down the two control groups the following way:

- Period over Period (70 days before and after the launch of e-bikes): 06/11/2018-08/19/2018 and 08/20/2018-10/28/2018
- Year over Year (50 days after the launch of e-bikes in 2018 and that same period in 2017): 09/12/2017-10/31/2017 and 09/11/2018-10/30/2018

The below visualizations explore the change in movement after Citi Bike introduced e-bikes.

Reading the Visualizations

The scatter plot compares growth rate between the two periods and number of trips before e-bikes were introduced (time = 0). The goal of the scatterplot is to help users put the growth rates into perspective, as not all are equal. For example, the first scatter plot shows that bike trips between Park Slope, Brooklyn and Steinway, Queens grew 11x, but when look at the y-axis and see that the original amount before e-bikes were introduced as only 1 trip, we can assess that the change wasn't as significant, as perhaps the growth of trips between Williamsburg, Brooklyn and Bushwick South stations. Although trips between Williamsburg and Bushwick grew about 2.2x, the first period had 35 trips, which would make the overall increase more than triple the number of trips between Park Slope and Steinway.

Moreover, the network visualizations represent only the growth rate between the periods before and after Citi Bike introduced e-bikes. The nodes represent neighborhoods and are mapped according to the centroid of each neighborhood in New York City. The sizes of the nodes represent absolute connectivity or the number of edges (lines) that connect back to a neighborhood. The network's edges represent only positive growth rates between neighborhoods before and after the introduction of e-bikes. For the edges, the larger and brighter the edge, the larger the growth rate.

Period over Period--Comparing the Weeks Before and After the Launch of E-Bikes

(Period over period scatter plot and network go here)

Year over Year--Comparing the Weeks After the Launch of E-Bikes in 2018 with the Previous Year's Period

(Year over year scatter plot and network go here)

From examining both visualizations, we can begin to hypothesize that after the introduction of e-bikes, riders began traveling longer distances and moving between boroughs more often. The possible increase in distance traveled by Citi Bike users after the introduction of e-bikes could be a sign of mobility patterns that can change as the bike share program rolls in a growing fleet of e-bikes across the City in the next few years.

Evaluating E-Bike and Regular Bike Travel Distance and Duration

After seeing the potential visual impact of e-bikes in the neighborhood network, I decided to take a look at how travel distance and duration was different for e-bikes than regular bikes. To do so, I tried guessing which bike ids belonged to e-bikes, by looking at bikes which were only present after August 20th, 2018 and cross-checking these with Citi Bike's station status API, which shows how many e-bikes a particular station has.

After narrowing down the list of e-bikes to about 230, I created a station network and ran these through MapBox's directions API. I then unnested the JSON responses and joined the directions dataframe with the data on individual trips taken by Citi Bike rides.

The analysis results seem to confirm the network visualization conclusions. It looks like e-bike trips, on average, were longer in distance and duration than the average non-e-bike trip. I also found it interesting that distance over duration was smaller for e-bike trips, which makes sense when you consider that e-bikes give users a push and move faster than regular bikes.

is_ebike	distance	tripduration	dist_durat
0.0	2096.548328	762.728991	3.568102
1.0	2104.313039	868.176800	3.564633