**CitiBike Statistics**

The data utilized was Jersey City (JC) from October 2018 to September 2019 to evaluate the trends happening for the past year.

A Jupyter notebook was utilized to clean the data prior to use it in Tableau. From the data selected, there was no missing values and the tables headers were the same. The following steps were taken to clean the data as much as possible:

- A new column was created to add the age of users and use that instead of year of birth.

- There were some years of birth dated to 1800s or early 1900s which made it difficult to believe there were real data. To solve this, the bottom Percentile 2 on ages was calculated and removed. So, only users born after 1961 were taken.

- The gender “unknown” was also removed.

- After these changes, the tables were concatenated to obtain a single table which included all the date from the period selected.

STORY

**Start Stations/End Stations by Popularity Maps**

These maps show the most common stations utilized to start or end a journey. When a map layer showing income per capita by Zip Code, it’s noted that the most popular stations are localized and more utilized in areas with the highest income.

DASHBOARD #1

**CitiBike Statistics by User Type**

First phenomena observed was that there were substantially more subscribers (Annual members) than Customers (24-hour pass or 3-day pass user). This is important to take into account as the number of residents benefiting from the service is much higher than casual customers or tourists.

From this, a graph was obtained to compare these records per month. As expected, trips were more commonly done during the months of warm weather like May through September.

Another 2 graphs created included the top 15 stations to start or end a journey. These had some variations depending on filter customer/subscriber. Perhaps these phenomena are related to areas with more touristic attractions. One thing to notice is that popular stations to start are same to popular stations to end. This could be related to the fact that, since we have more subscribers data, the bikes are being used as a transportation to work and then left back at start station.

DASHBOARD #1

**CitiBike Statistics by Gender**

Second phenomena observed was that there were substantially more Male than Female bikers.

The bubbles graph demonstrates that the relationship by gender remains stable throughout the different starting stations.

The linear graph shows most common user’s by age and gender. As expected, bikers were mostly on their late 20s and early 30s.

Last linear graph shows distance trip (miles) performed by gender. Here we can conclude that most of bikes are used for half mile trips in both genders. Also can see the longest trips performed, divided by gender.