Citi Bike Analytics with Tableau

By Katelyn Burke

Data Range: January - December 2018

Data Source: https://www.citibikenyc.com/system-data

Tableau Interactive Visualizations:

About Citi Bike:
Renting a Citi Bike is a fun, green, and easy way to travel around New York and New Jersey. Use a bike to commute to work, meet friends for lunch, or sight see around the city. Citi Bike is the nation's largest bike share program. It has 12,000 bikes and 750 stations across Manhattan, Brooklyn, Queens and Jersey City. Using Citi Bike can also save you money. It is much more affordable than taking a Taxi or using the Subway.

Project Goals and Overview:
This project uses the visualization and analytics software Tableau. In this project, I aggregate the data found in the Citi Bike Trip History Logs to build a data story on Tableau to answer many pressing questions about Citi Bike. I also used Python Pandas and Jupyter Notebook to clean and merge the Citi Bike data sets. The Tableau story I have created, helps report interesting facts and data about the Citi Bike program during 2018 and 2019 in an engaging, understandable, and colorful way.
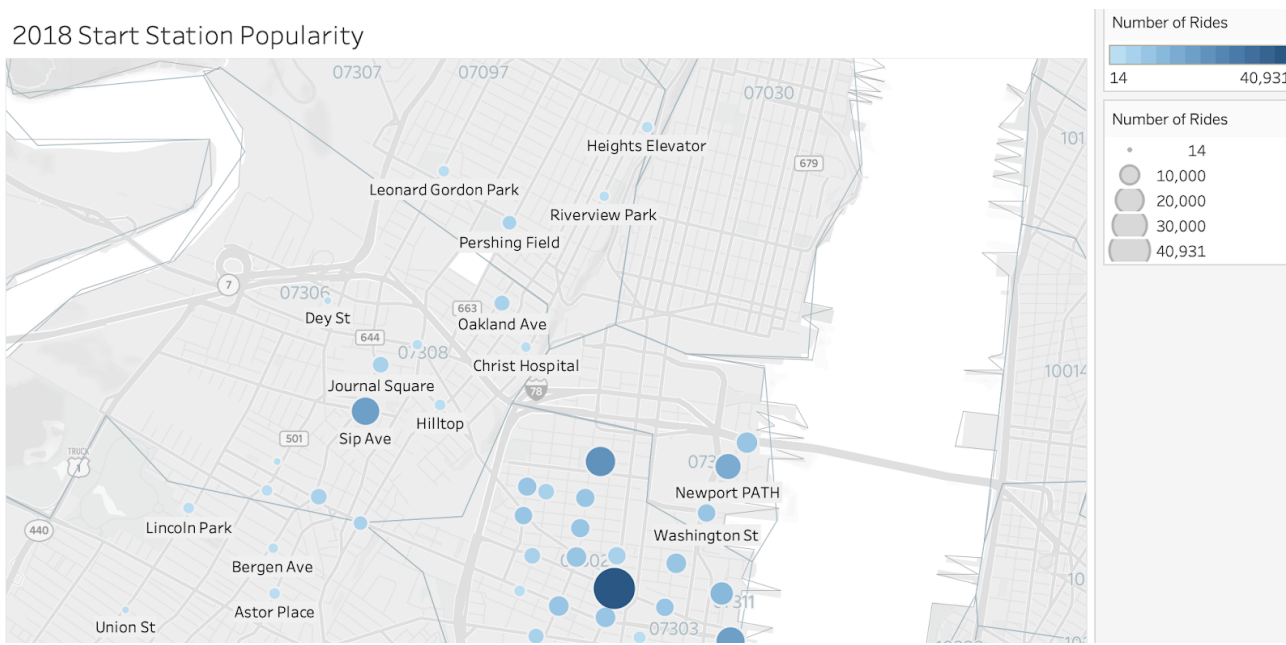
Challenges:
- The Tableau Public program would crash often when trying to save new graphs - this would result in losing work that I had done
- Citi Bike data sets were extremely large - 20,000 or more rows of data for each month in 2018
  - Tableau Public has limitations for the amount of data that can be imported so I was unable to merge multiple months of data in Tableau
- I used Pandas in a Jupyter Notebook to merge all of the files for 2018 so I could create a CSV of yearly data which I was able to import into Tableau
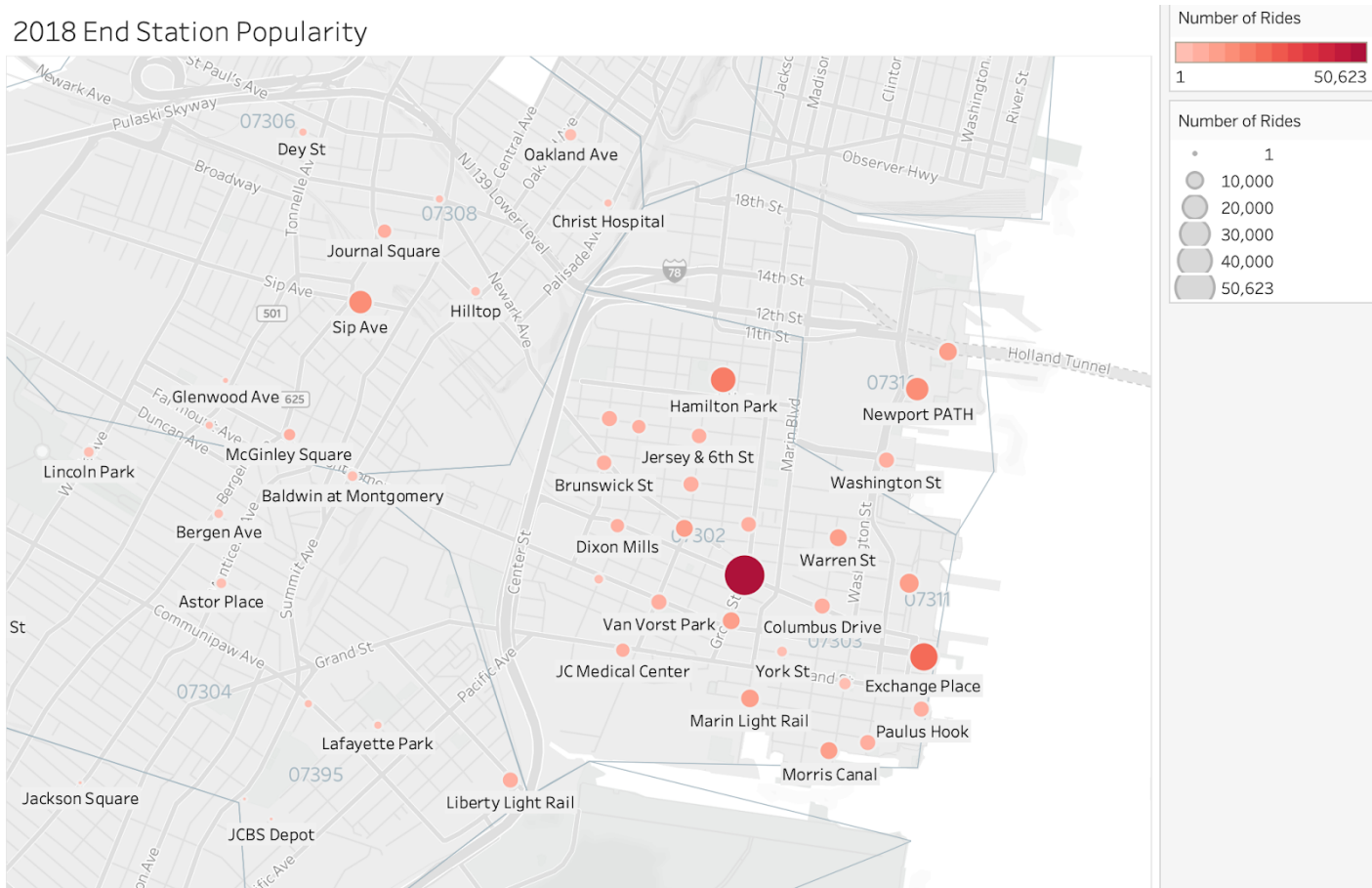
Analysis and Reporting - 2018 Data

Start Station Popularity

- Start and End Station Popularity maps include all 2018 Citi Bike ride data (used Pandas in a Jupyter Notebook to merge the monthly data)
- The most popular start station is the Grove Street Path station with 40,931 rides started from it in 2018
- Clusters of popular starting stations tend to be located within four zip codes: 07302, 07310, 07311, 07303
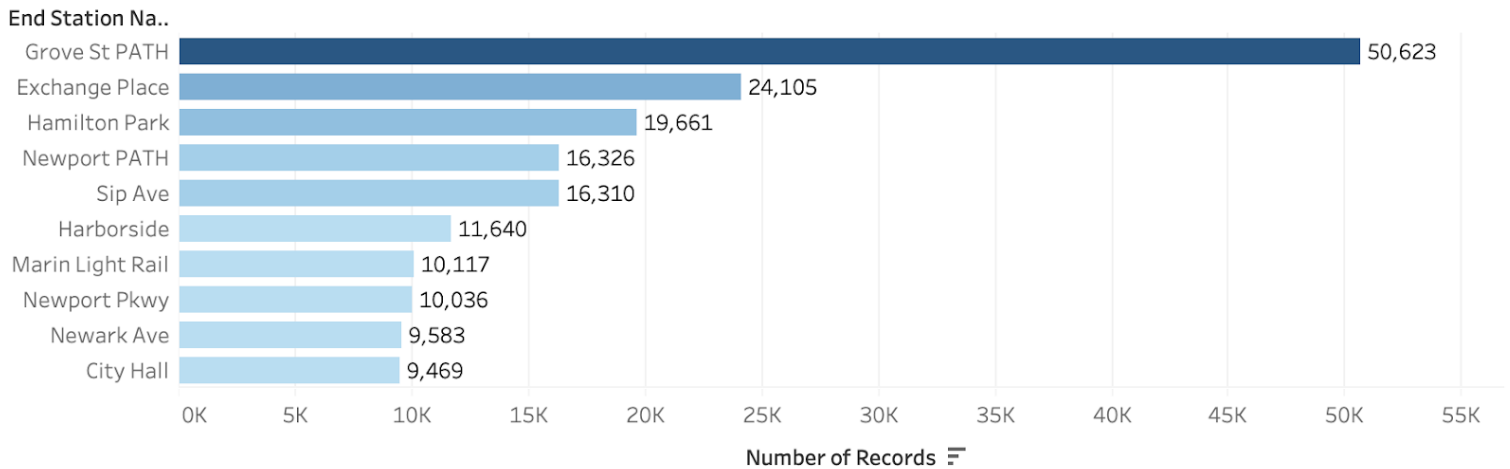- More popular stations fall closer to the water

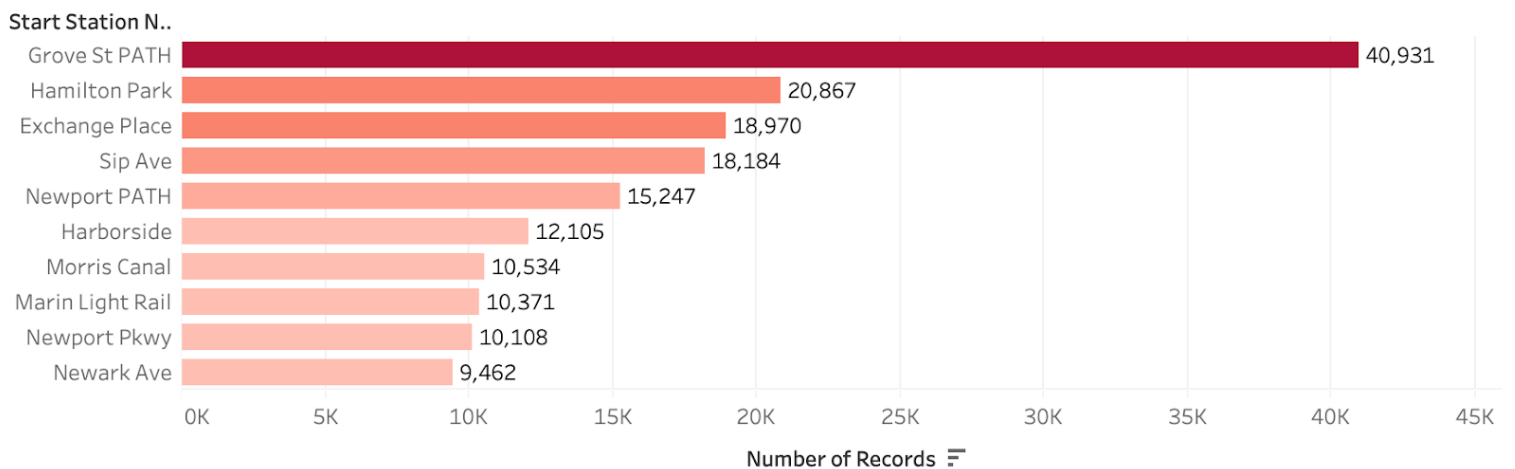**End Station Popularity**

- The most popular end station is the same as the most popular start station - Grove Street Path station
- More rides have ended at this station than started at this station - 50,623 rides
- Approximately 10,000 more rides have ended at this station than started at this station
- I am inferring that many riders start and end their rides at the Grove Street Path Station

Top Stations to Begin and End a Ride
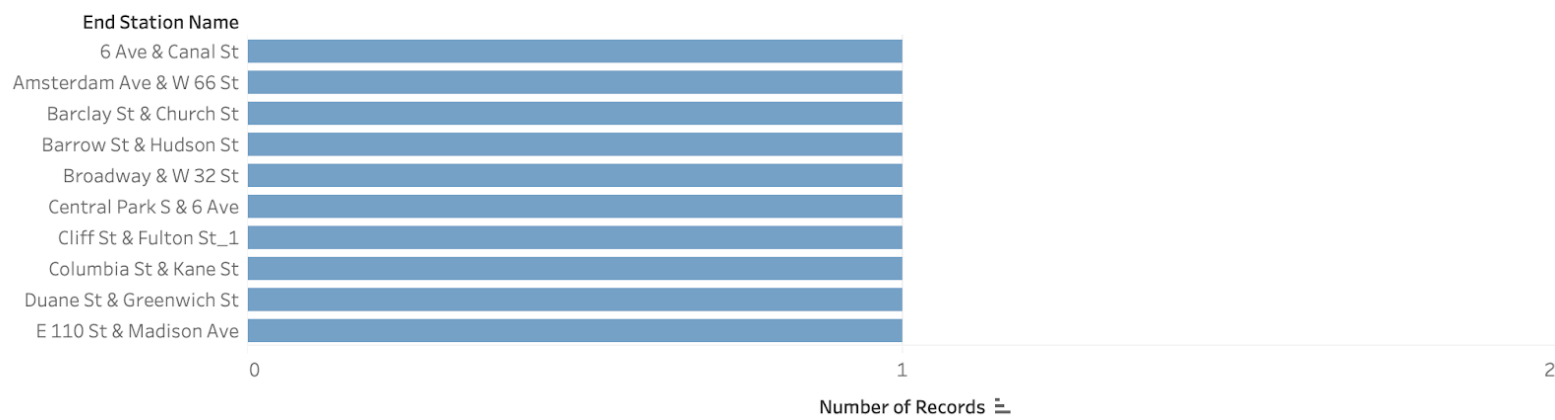
Top Stations to End a Ride 2018

End Station Na..

Station	Records
Grove St PATH	50,623
Exchange Place	24,105
Hamilton Park	19,661
Newport PATH	16,326
Sip Ave	16,310
Harborside	11,640
Marin Light Rail	10,117
Newport Pkwy	10,036
Newark Ave	9,583
City Hall	9,469

Number of Records

Top 10 Stations to Begin a Ride 2018

Start Station N..

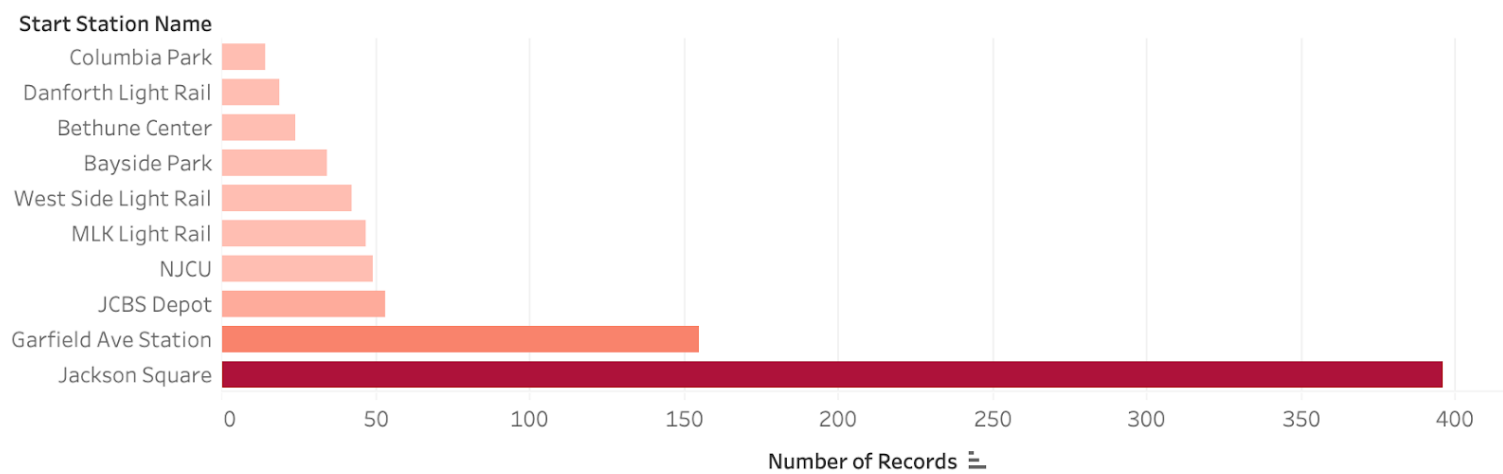| Station	Records
Grove St PATH	40,931
Hamilton Park	20,867
Exchange Place	18,970
Sip Ave	18,184
Newport PATH	15,247
Harborside	12,105
Morris Canal	10,534
Marin Light Rail	10,371
Newport Pkwy	10,108
Newark Ave	9,462

Number of Records

- In 2018, the most popular station to end a ride is the Grove Street Path station, followed by the Exchange Place station second and the Hamilton Park station third
- In 2018, the most popular station to begin a ride is also the Grove Street Path station, followed by the Hamilton Park station second and the Exchange Place station third
- According to Google Maps, the Grove Street Path station is located in an area of Jersey City that has many attractions including restaurants, shops, hotels and other businesses. This could be a reason why this station is so popular to begin and end a ride
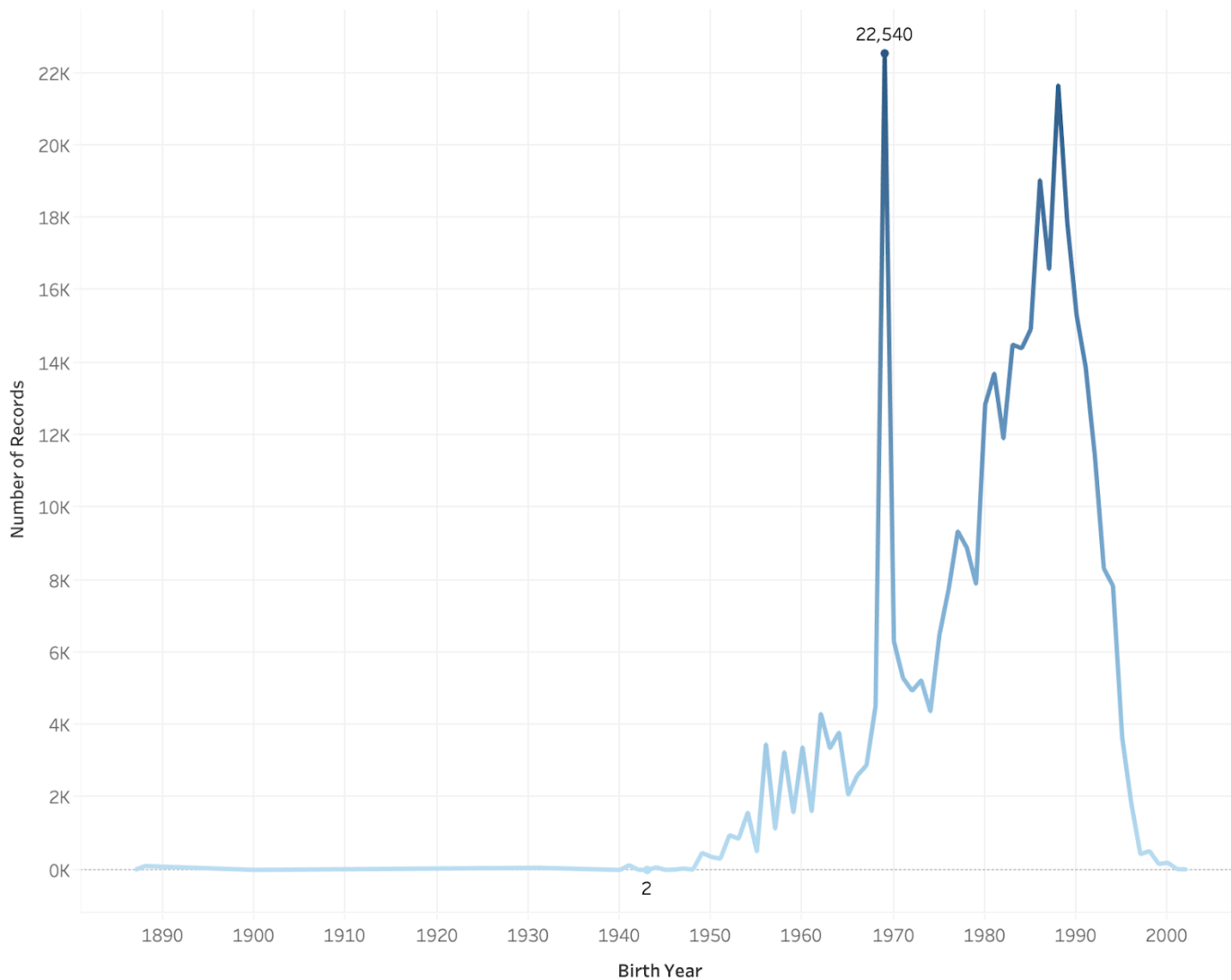
Bottom 10 Stations to End a Ride 2018

End Station Name

Station	
6 Ave & Canal St	
Amsterdam Ave & W 66 St	
Barclay St & Church St	
Barrow St & Hudson St	
Broadway & W 32 St	
Central Park S & 6 Ave	
Cliff St & Fulton St_1	
Columbia St & Kane St	
Duane St & Greenwich St	
E 110 St & Madison Ave	

Number of Records

Bottom 10 Stations to Begin a Ride 2018

Start Station Name

Station	
Columbia Park	
Danforth Light Rail	
Bethune Center	
Bayside Park	
West Side Light Rail	
MLK Light Rail	
NJCU	
JCBS Depot	
Garfield Ave Station	
Jackson Square	

Number of Records

- In 2018, the least popular station to end a ride is a tie with only one recorded ride (see bar chart)
- In 2018, the least popular station to begin a ride is the Columbia Park station, followed by the Danforth Light Rail and Bethune Center stations
- I am thinking that the bottom 10 ending stations may be inactive stations because they only have one ride recorded
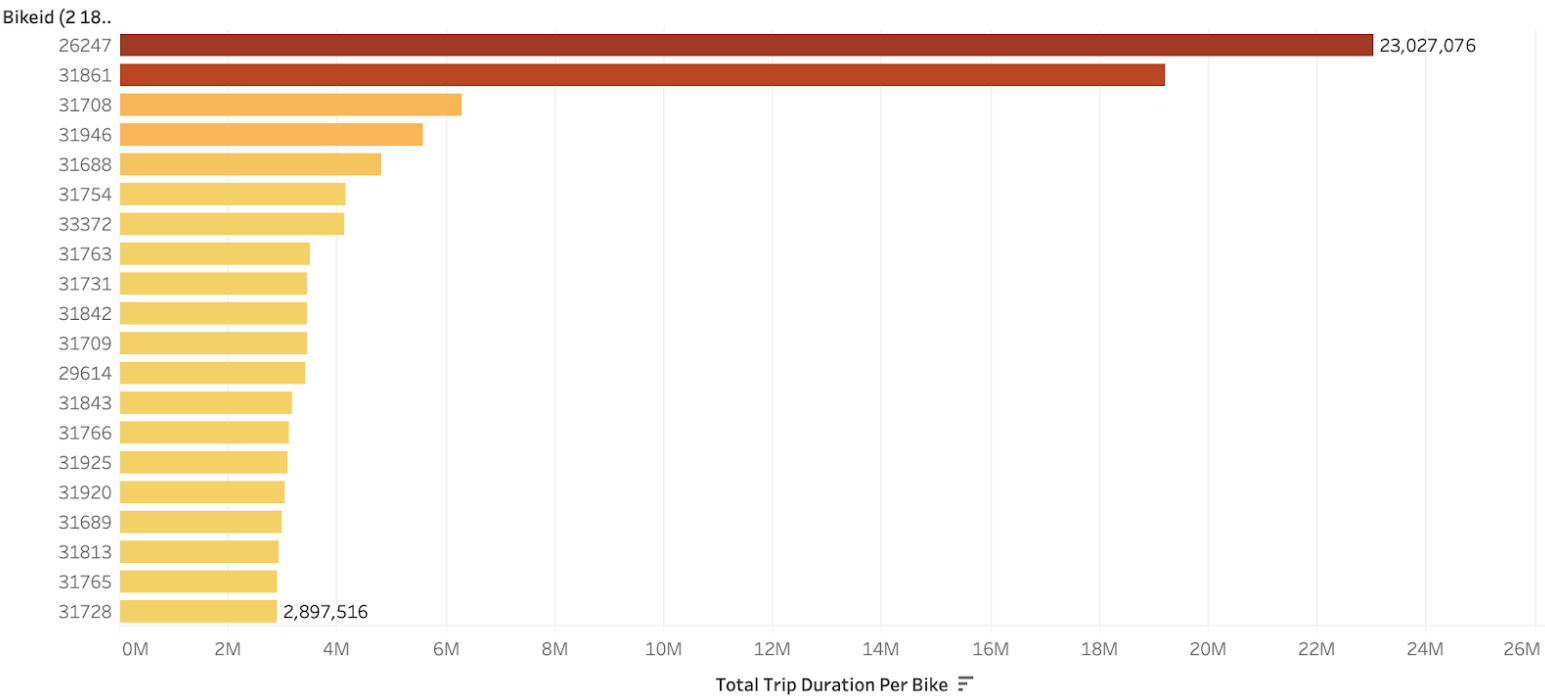
Rides by Birth Year/Age

Rides by Birth Year/Age 2018

- The most rides (22,540) in 2018 were taken by customers born in 1969 - making them approximately 49 years old in 2018
- The fewest number of rides (2) were taken by customers born in 1943 - making them approximately 75 years old in 2018
- I am thinking that there is some error in the birth year data - customers may not have entered their correct birth year when registering for a Citi Bike. An example of error is that there are 111 customers that registered their birth year as 1888 when they rode a bike in 2018. This would make them approximately 130 years old
- According to Citi Bike's policy - you have to be at least 16 years old to rent a bike. The data cuts off at the birth year 2002 (16 years old in 2018) because of this policy
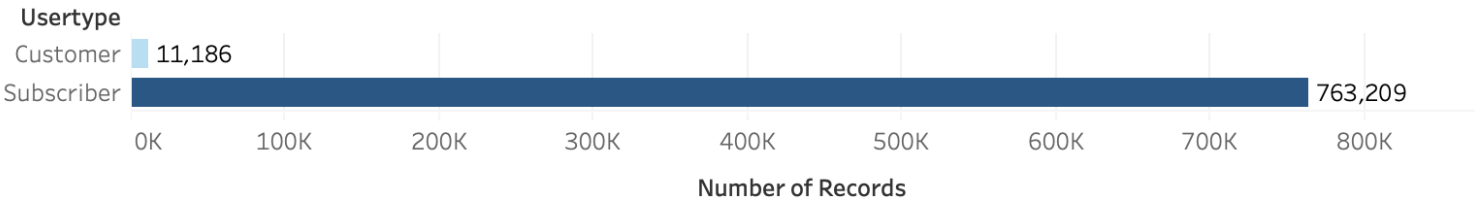
Top Bikes that Most Likely Need Repair

Top 20 Bikes that Most Likely Need Repair

Bikeid (2 18..

- Bikes with higher trip durations (in total) are more likely to require repairs due to how frequently they are used
- The bike with ID number 26247 has had trips totaling in 23,027,076 minutes - the highest trip duration of Citi Bike used
- It would be beneficial for Citi Bikes to track the bikes that are used most often so that they can be inspected for safety and repairs can be made quickly
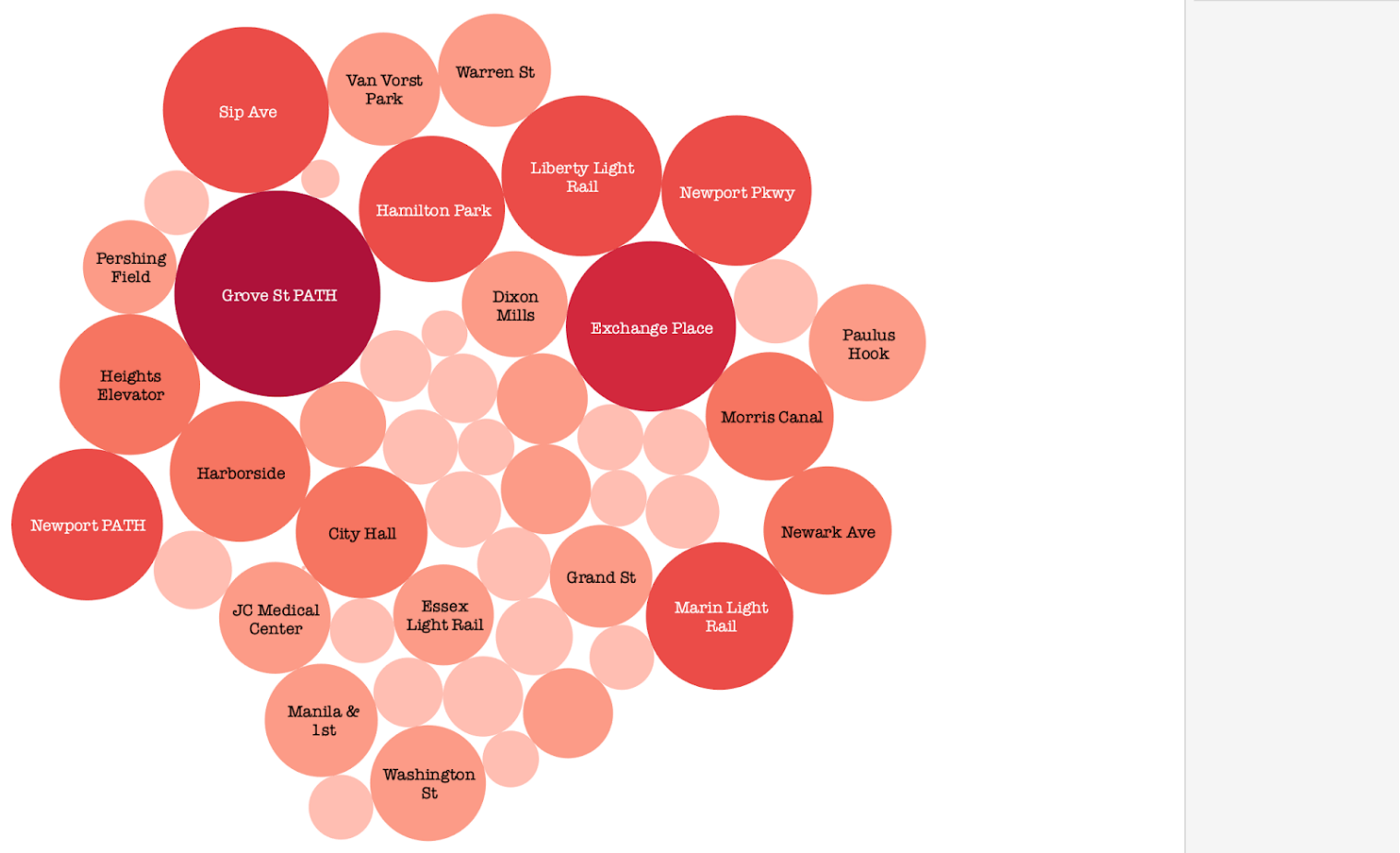
Rider Type: Customers vs. Subscribers

January 2018 Rider Type - Customers vs. Subscribers

- Subscriber option (an annual membership with a one-year commitment) is a much more popular option for Citi Bike riders (763,209 subscribers in January 2018) rather than one-time customers (11,186 customers in January 2018)
- Subscribers end up saving money with the membership, especially if they are using the bikes frequently

Summertime (June 2018) Trip Duration

Summertime (June) Trip Durations from Start Stations

Sum of Trip Durations

4,453,539   13B

- Highest total trip duration from the most popular starting station - Grove Street Path with 2,242,333 minutes of rides
- This is also the number one begin/end Citi Bike station