Citi Bike Data Analysis

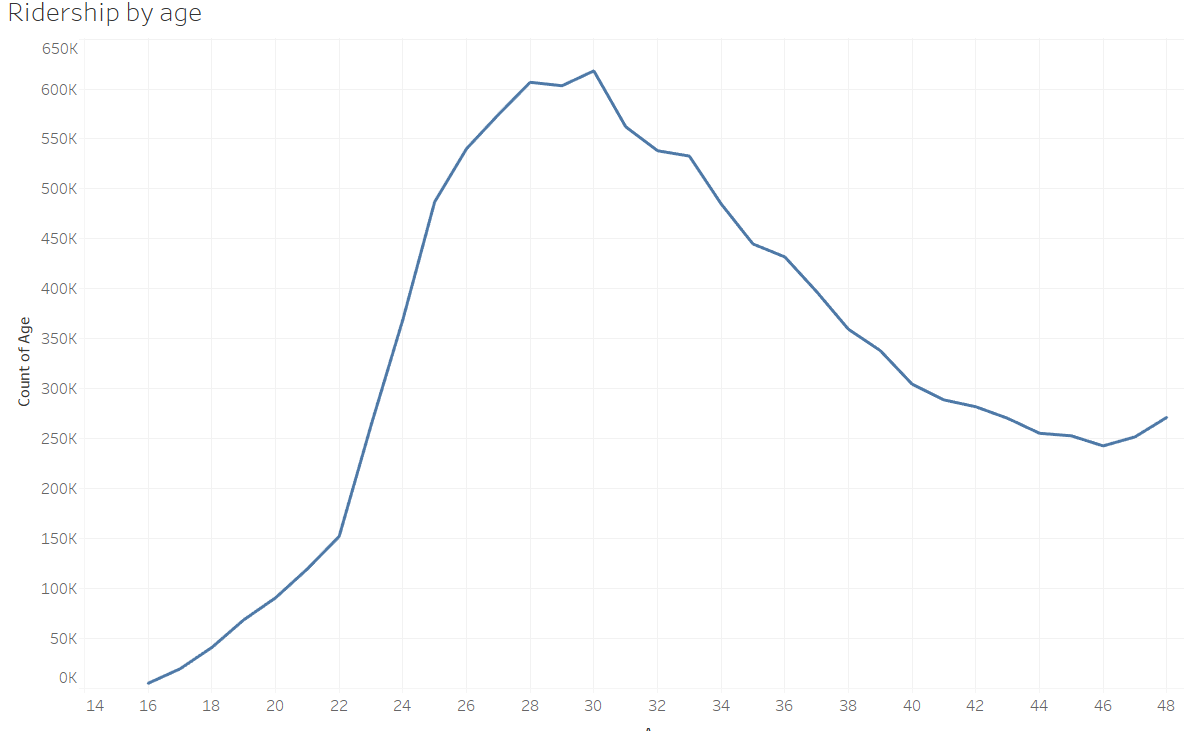
An overview of usage and demographic statistics

Ben Farniok

Demographic Data

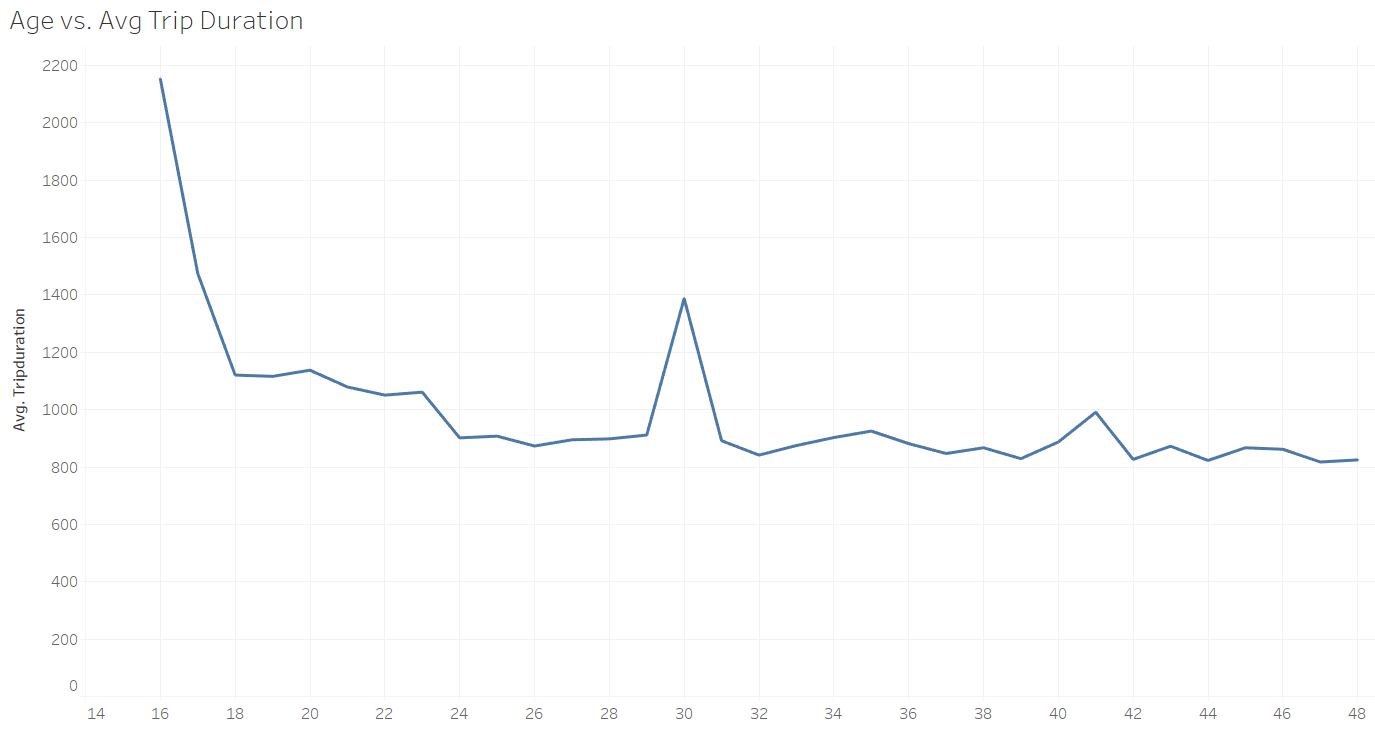
Over the process of analyzing the data for the New York Citi Bike Program, a number of data points showed up and showed clear trends within how the people in the city and visiting the city are making use of the program.

First off, ridership is significantly higher among people within the age ranges of 25-30, but stays higher than the average for riders in their twenties and through their thirties.



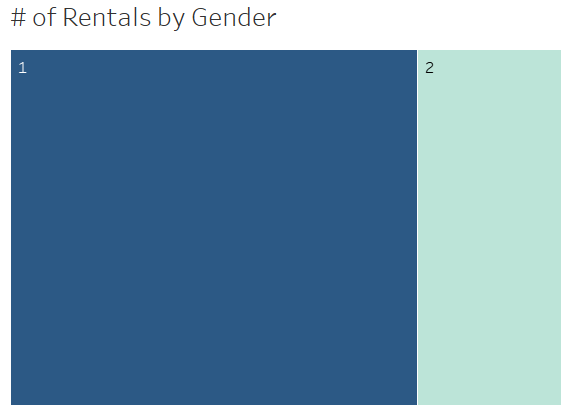
Further, ridership does not fall off as steeply as it picks up with age, this may be due to adults in their mid and late twenties finding more value in the program, and then only slowly falling off as their living situations change and move on in their careers, purchasing cars and other potential shifts. Ages 50 and above are excluded on account of unreliable reporting.

Among that ridership, the analysis also revealed useful trends in the average trip times undertaken by age. Young adults and teenagers take by far the longest trips, peaking with the youngest age group of 16 taking much longer trips than their older counterparts, with increasing age acting as a strong predictor of lower and lower trip durations. Again, ages 49 and over were excluded due to significant outlier data.

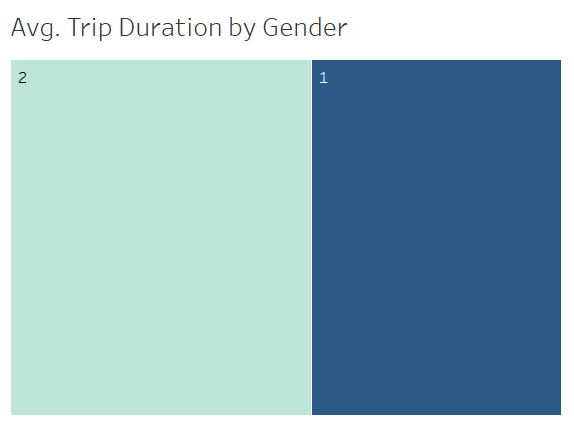


Analysis of gender among riders showed that men are much more likely to take advantage of the Citi Bike program.

The study showed that men make use of the system more than three times as often as their female counterparts.



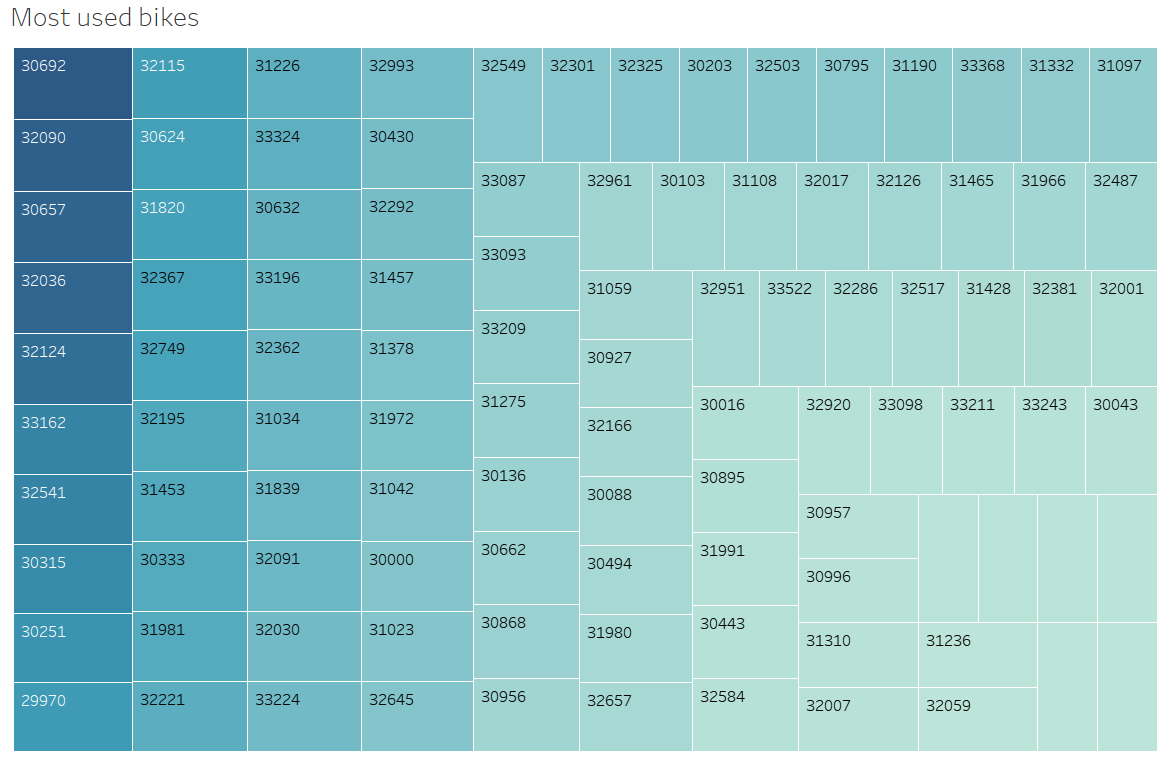
Here, we can see that male riders (represented by the label “1” in the database) make use of the system much more often than women (represented as “2”). This has been a consistent issue within the program, women consistently make less use of the system in favor of other means of transportation around the city.



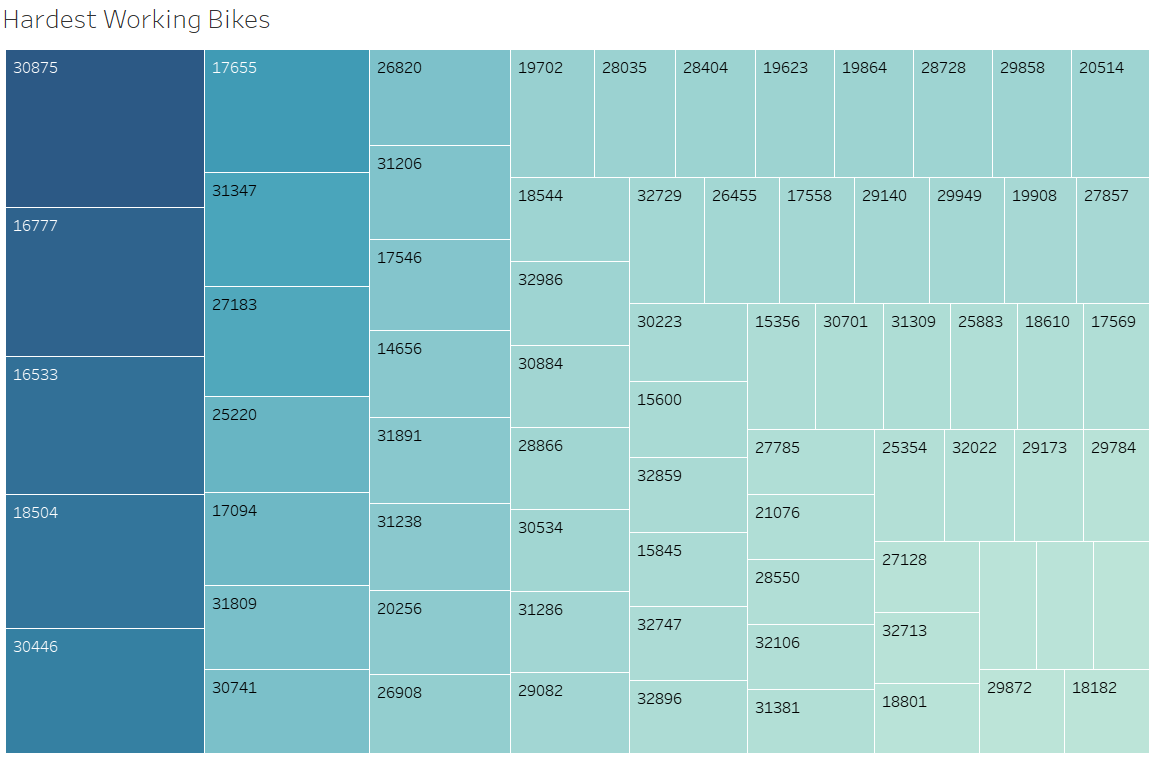
While they represent less of the total ridership, we can see that when women are using Citi Bike, their average trips last a significantly longer time. This could be put down to multiple reasons, such as making longer journeys on the bikes. Unfortunately, we do not have data to accurately say what is the case, only that the trips are generally longer.

This demographic information could be used to push marketing efforts as it can show us which groups are under-engaged and which ones have responded well to the program already.

On the topic of usage of the program we can see a variety of factors that could help program management by pushing maintenance efforts in the most effective, data driven ways.



First off, we have representations of the most-heavily used bicycles. These bikes, represented in this chart by ever darker shapes, gives us a picture of which bikes are rented most often by subscribers to the program as well as those who are only making short-term use of the bicycles by their bike ID. The most heavily used in 2018, with the Bike ID 30692, has been rented 1,799 times over the course of the year. This chart, filtered down for effect in this report, shows only the most-often used bikes across the city. The least often-used bike, with the ID 35443 has been rented only once, making it less likely to go through the wear-and-tear that the more-often rented bikes have experienced.



A similar chart here could also serve to direct maintenance efforts, the above graphs shows representations of the bikes with the highest total trip duration for the year. The longest-running bike in the bunch, numbered 30875 has spent more than 4,400 hours out on the streets and sidewalks of the city.

Whether through the inspection of the most often-used bikes or simple maintenance work, this visualization provides insight into how the bikes are being ridden by Citi Bike users and how the city can best care for the bikes, ensuring they will continue to be in good repair for any current and future bicyclists