2018 Citi Bike Analysis – Harry Chapman

**Data Overview**

Monthly data from 2018 was retrieved from citibikenyc.com and then a 2% random subset was created from that for more easy manipulation.

**Density Map**

The density map shows the frequency of trips that end at a certain location given our data subset. Apparently, stops in Midtown Manhattan are the most popular. This is most likely due to this area’s position as the economic hub of NYC.

**Usage and Duration by Birth Year**

Bike usage gradually increases from those born 1975 and peaks in 1988 before dropping sharply for younger users. In 2018 when this data was aggregated, those born in 1988 were 30 years old. It makes sense that there is a larger population of people in the prime of their career that use the Citi Bikes in Manhattan and explains the steep drop off for younger age users as there is less career opportunity for them in those locations initially.

Interestingly, average trip duration is longer for those born in the second half of the 1990s-2002. This may be due to this age group using the bikes for leisure instead of commuting and thus their trip times are more protracted.

**Customer and Subscriber Count per Year**

Customers are defined as Citi Bike users who purchase daily or 3-day passes and subscribers are those who are yearly members. As those born in 1988 were found to be peak users, it makes sense that those in that age range +/- 2 years make up the majority of customers and subscribers. Correspondingly, the younger age groups (and older age groups before 1988) make up smaller proportions similar to their usage metrics in the preceding visualization.

**Trip Count by Gender**

Men use Citi Bikes far more often than women according to the 2018 data set. This may be due to women’s propensity to use other means of public transportation such as the subway or bus lines or because there are less women working in Manhattan than men.

**Trip Duration by Gender**

The median female trip duration exceeds that of male trip duration. This could be due to women living farther away from their destinations or perhaps their propensity to take longer Citi Bike rides rather than use other means of public transportation. In light of less women than men using Citi Bikes on an absolute basis, the latter point is less likely. Additionally, women could be using Citi Bikes for leisure more than men, which would extend trip duration as they are not using it as a means to get from one point to another in the most efficient way possible.

**User Type by Gender and Usage by Gender**

There are almost 4x as many male subscribers as there are female subscribers. Additionally, there are 23x more male subscribers than customers. The disparity in subscriber to customer split is most likely due to the financial incentives people receive for purchasing longer-term memberships with Citi Bike.

Similar to what has been seen in the above visualizations, males dominate Citi Bike usage. There is also a relatively large group of “unknown” gender users that is a little over 1/3 the size of female users. Citi Bike most likely provides an option upon sign up for people to choose not to disclose their gender. This raises the question of how accurate gender is as a classifier as a whole since the veracity of the user input cannot be known. This may also hold true for date of birth but may be corroborated with credit card information.