Dataset

Ford GoBike Data set is Lyft dataset for thier public bicycle shaing system. The project is a regional public bicycle sharing system in California's San Francisco Bay Area. It is operated by Motivate in a partnership with the Metropolitan Transportation Commission and the Bay Area Air Quality Management District. (Wiki) As it's expands it reaches many places. This dataset was collected from New york city in June, 2020.

This data answers questions like: Where do Citi Bikers ride? When do they ride? How far do they go? Which stations are most popular? What days of the week are most rides taken on? This data exploration help you discover the answers to these questions and more.

This data is provided according to the NYCBS Data Use Policy.
https://www.citibikenyc.com/system-data

image from: https://www.digitaltrends.com/outdoors/ford-gobike-plus-san-francisco/

Main finding

- The Average users age was observed to be between 25 and 35. However, there are a strike of users over the age of 50 resulting the mean users to lean over the right side of the distribution.

- The average trip duration lies between 10 to 50 min, which is a big interval to determine the pattern of users. The distribution is normally distributed giving a mean of around 40 min of each trip.

- The day usage of the customers and subscribers which is inversed in the two groups.

- The age groups who clusters in the casual customers at 50s and in the subscribers at younger ages. The addition of variables life distance and time to these variables showed the habits of the groups, their abilities to travel, and their extremes.

- Most of the casual customers tend to use the bike system in vacations and weekends for fun, and the subscribers uses the system in the normal weekdays which suggests that they use the system for work or school.

- Most of the subscribers are males and least customers are females.

- The subscribers tend to be in younger age from the casual users as the data showed that subscribers are mostly in the range between 20s and middle 30s unlike the casual customers who most of are above 50.

Resources

- https://gist.github.com/rochacbruno/2883505

- https://stackoverflow.com/questions/29096381/num-day-to-name-day-with-pandas

- https://stackoverflow.com/questions/52753613/grouping-categorising-ages-column-in-python-pandas#:~:text=If%20age%20>%3D%200%20%26%20age,so%20on%20