

Ford GoBike Data Exploration

Tanmay Gangurde

Outline

1. Introduction

2. Questions for Analysis

- i. Most common start and end stations.
- ii. Gender trends observed in the dataset.
- iii. Range of age group.
 - i. Age vs User Type
- iv. Riding Durations
 - i. Riding Durations for Subscriber and Customer

Introduction

Ford GoBike is a **public bicycle sharing platform** in San Francisco. For ease of transportation, **bicycles are available for the people at any point of time**. Either you can become a subscriber or avail this service as a customer.

This dataset has **details of users** like what was the **starting point, end point, age of user, duration** and many more.

This dataset has **183412 rows and 16 columns**.

Questions for Analysis

1. Most common start and end stations.
2. Gender trends observed in the dataset.
3. Range of age group.
4. Riding Durations and distance travelled.

i. Most common start and end stations.

The most common **starting point** is **Market St at 10th St** and most common **ending point** is **San Francisco Caltrain Station 2**.

Top 5 Starting and Ending point are Listed below.

Top 5 most common starting points

Market St at 10th St
San Francisco Caltrain Station 2 (Townsend St at 4th St)
Berry St at 4th St
Montgomery St BART Station (Market St at 2nd St)
Powell St BART Station (Market St at 4th St)

Top 5 most common Ending points

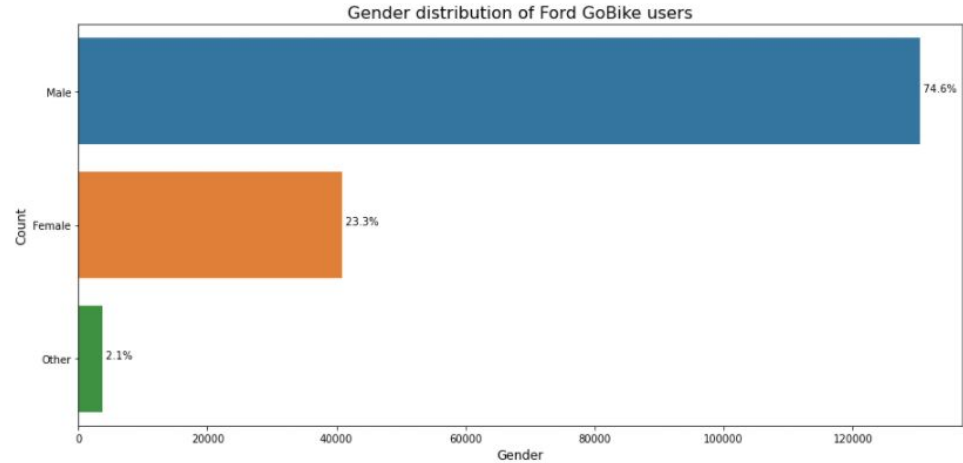
San Francisco Caltrain Station 2 (Townsend St at 4th St)
Market St at 10th St
Montgomery St BART Station (Market St at 2nd St)
San Francisco Ferry Building (Harry Bridges Plaza)
Powell St BART Station (Market St at 4th St)

ii. Gender trends observed in the dataset.

According to data set there are 3 types of people who uses bicycle.

1. **Male**
2. **Female**
3. **Other**

According to the graph, **Male** users comprises of **74.3%** followed by **Female** with **23.3%** and then **other** with **2.1%**



Clearly male users are more as compared to other two category.

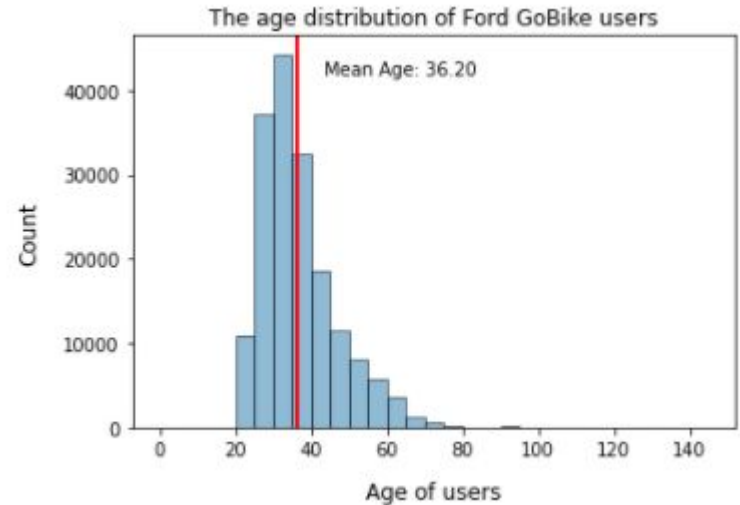
iii. Range of age group.

According to the graph, we have variety of people belonging to different age groups.

Minimum age: 20.

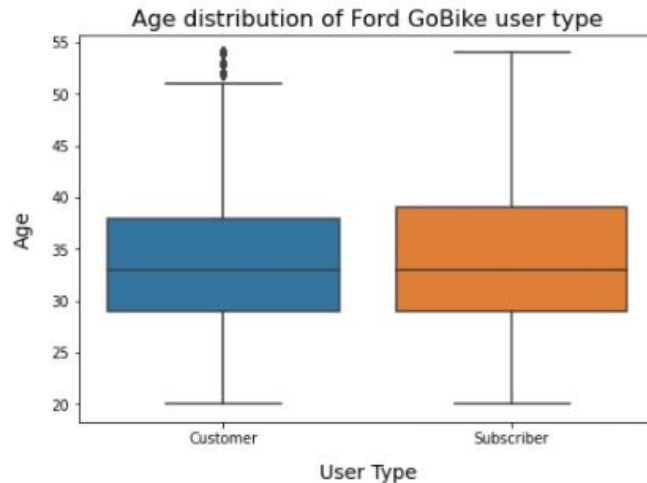
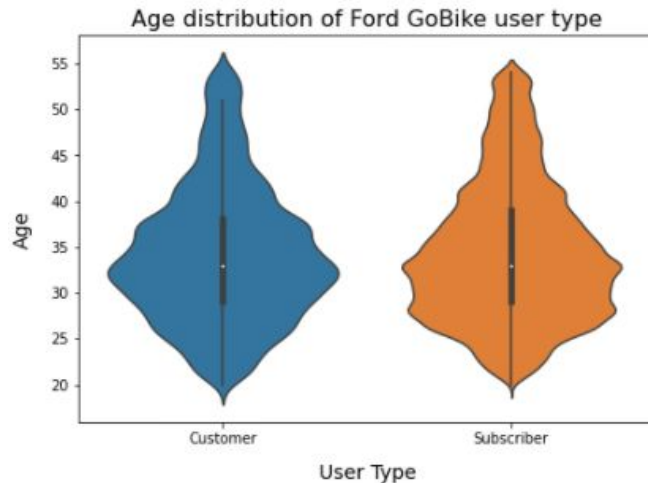
Maximum age: 143.

NOTE: But due to the presence of Outliers, dataframe was modified and constricted till the age of 54.



iii.i Age vs User Type

With the help of Violin Graph it is observed that, maximum people who come under **Customer type**, their **age is in between 28-37**, where as, maximum people who fall under **Subscriber**, their **age is in between 28-38**. Similar trend is observed in box plot as well.

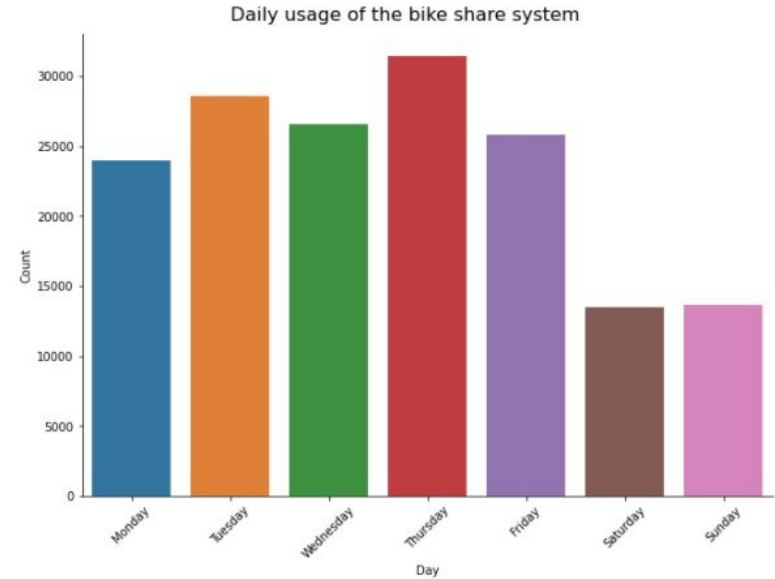


iv. Riding Durations

Looking at bar graph, majority of users prefer using bicycle on weekdays.

They might use bicycle to go to office or Schools, and this notion fits perfectly.

On thursday, Ford GoBike experience rush as compared to other days.



iv.i Riding Durations for Subscriber and Customer

With this we can observe that, Subscriber's workplace is near to the start point as they take less time to reach to their destination. On the other hand, Majority of Customers use bicycle occasionally.

