Part II - Ford GoBike System Data

by Darine Battikh

Investigation Overview

This presentation aims at featuring the main conclusion of the exploratory analysis. Although the exploration part has treated thoroughly many variables, this presentation will highlight the interesting insights and conclusions. This investigation will focus on the trip duration and factors that can affect it, mainly, user type, days, members' gender, and members' age.

Dataset Overview

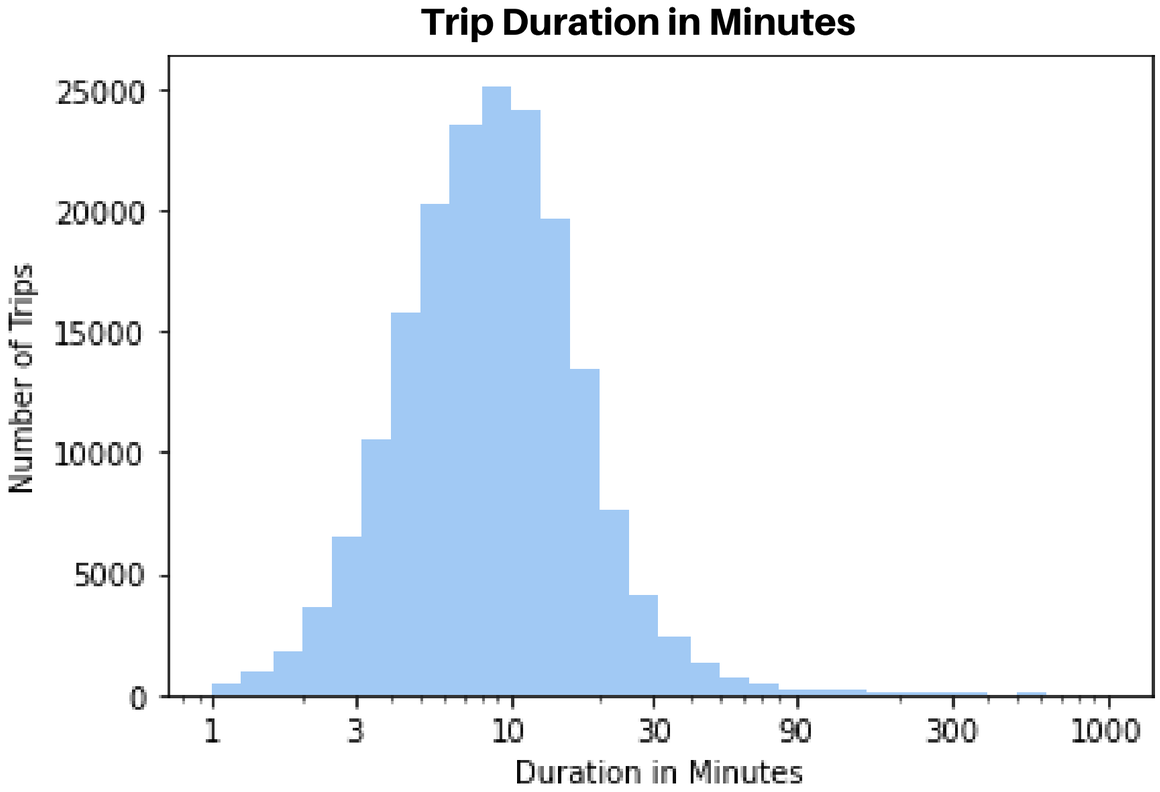
Ford GoBike is a bike sharing system located and covering the San Francisco Bay area. It provides more than 4,000 bikes distributed across different accessible locations allowing customers to get bikes from one station and return to any other station in the system, making them ideal, accessible, and convenient for one-way trips, for special and urgent occasions.

The bikes are available for use 24 hours/day, all days of the week, and all days of the year. Customers have the possibility to become subscribers through a monthly or yearly subscriptions or simply purchase one ride for a determined number of hours.

The Ford GoBike System Data is a data frame was composed of **16 column**s and **183412 entries** and requires some cleaning operations. The data frame has different data types like objects, integers (64), and float64.

Trip Duration and Daily Usage

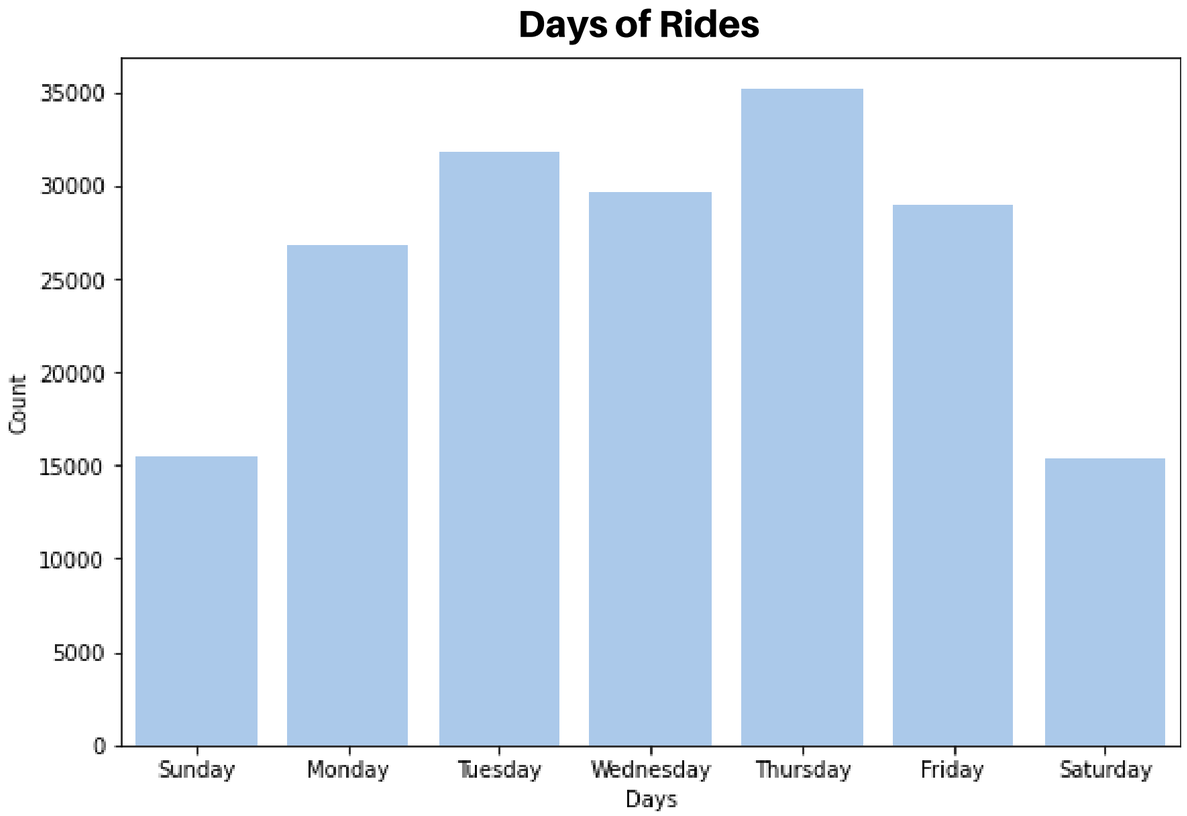
On average, the most of trips last 10 minutes and the majority of rides take between 3 minutes and 30 minutes. Most of rides are taken on Tuesday and Thursday, yet, longer rides are made on the weekends.



Based on the histogram of the Trip Duration in Minutes, we can state the following:

- Nearly 25000 bike rides lasted 10 minutes.
- The majority of bike trips last between 3 minutes and 30 minutes.
- Half of the bike trips last less than 10 minutes.

Trip Duration and Daily Usage



This histogram shows that most members use the bikes on weekdays with two peaks on Tuesday and Thursday which is in accordance to the dates extracted from the previous bar chart.

User Types and Their Service Usage

The Bike users are distributed as follows:

- 89% are subscribers
- 11% are customers

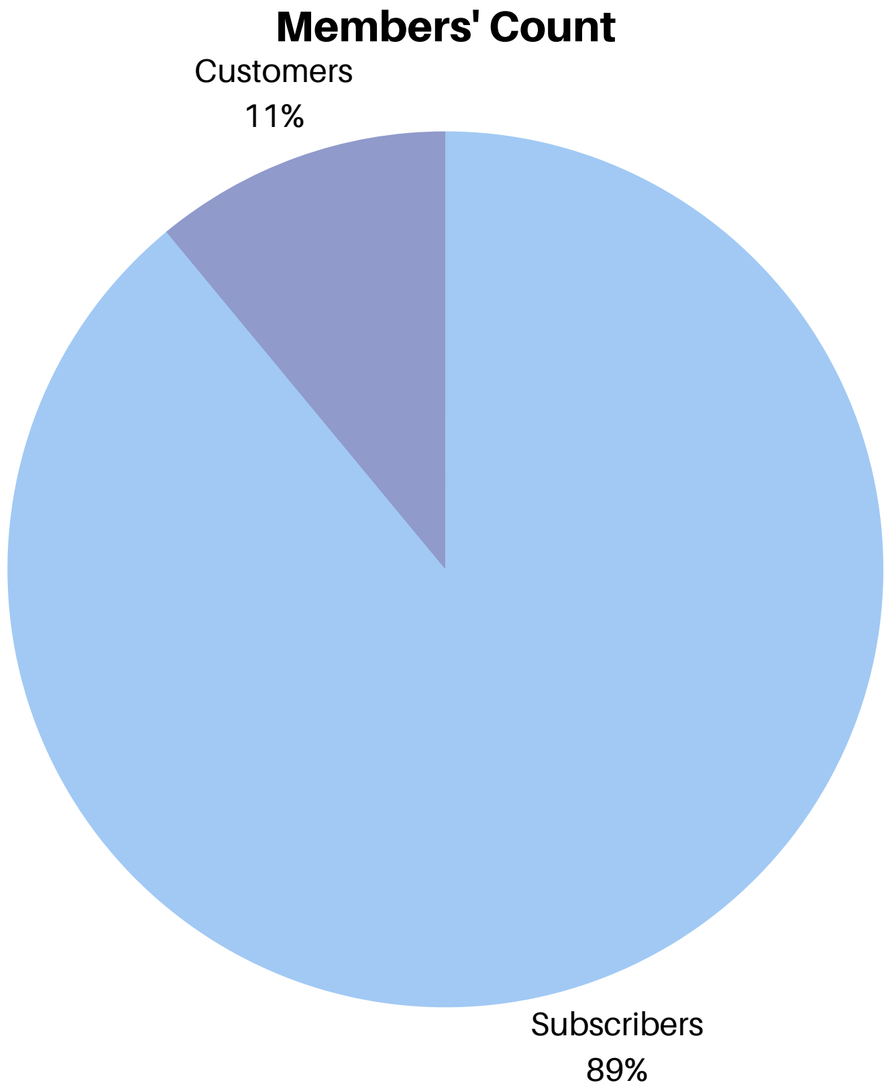
For the service, the usage among user types is different:

- Subscribers' service use is consistent over all days, with a slight increase for the weekdays.
- Customers, in contrast, have a heavy use of bikes on weekends and decrease on the rest of the weekdays.

Also, customers tend to have longer rides than subscribers

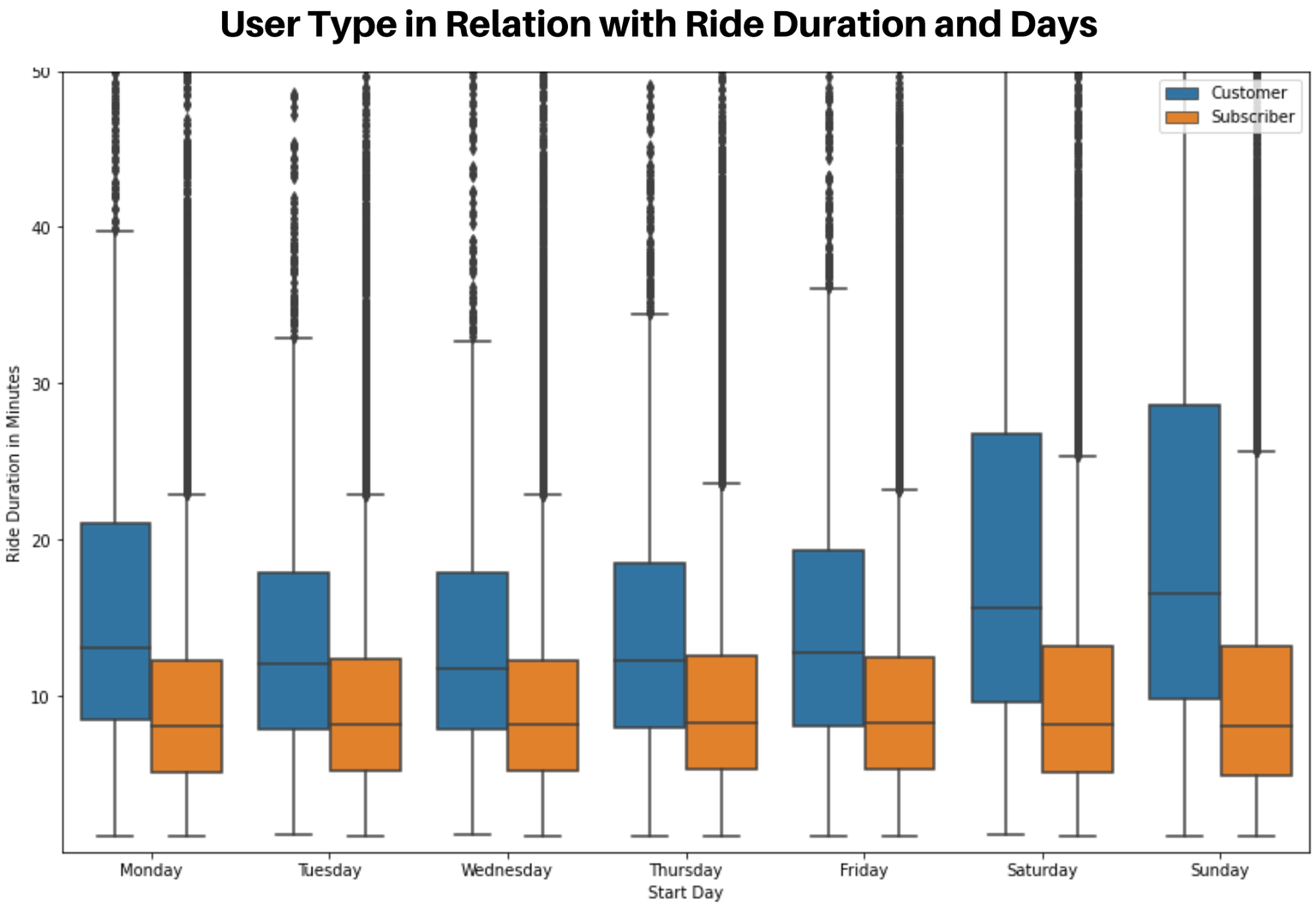
This can be explained b the fact that customers are using the service for leisure and enjoyable trips, whereas, subscribers rely on the service for their daily actitivities (work, school, etc).

User Types and Their Service Usage



The pie chart illustrates that 89% of the members of GoBike are subscribers compared to only 11% of customers.

User Types and Their Service Usage



This boxplot shows that subscribers have a consistent usage in terms of days and duration of bikes in contrast with customers.

Customers have longer durations during the weekends.

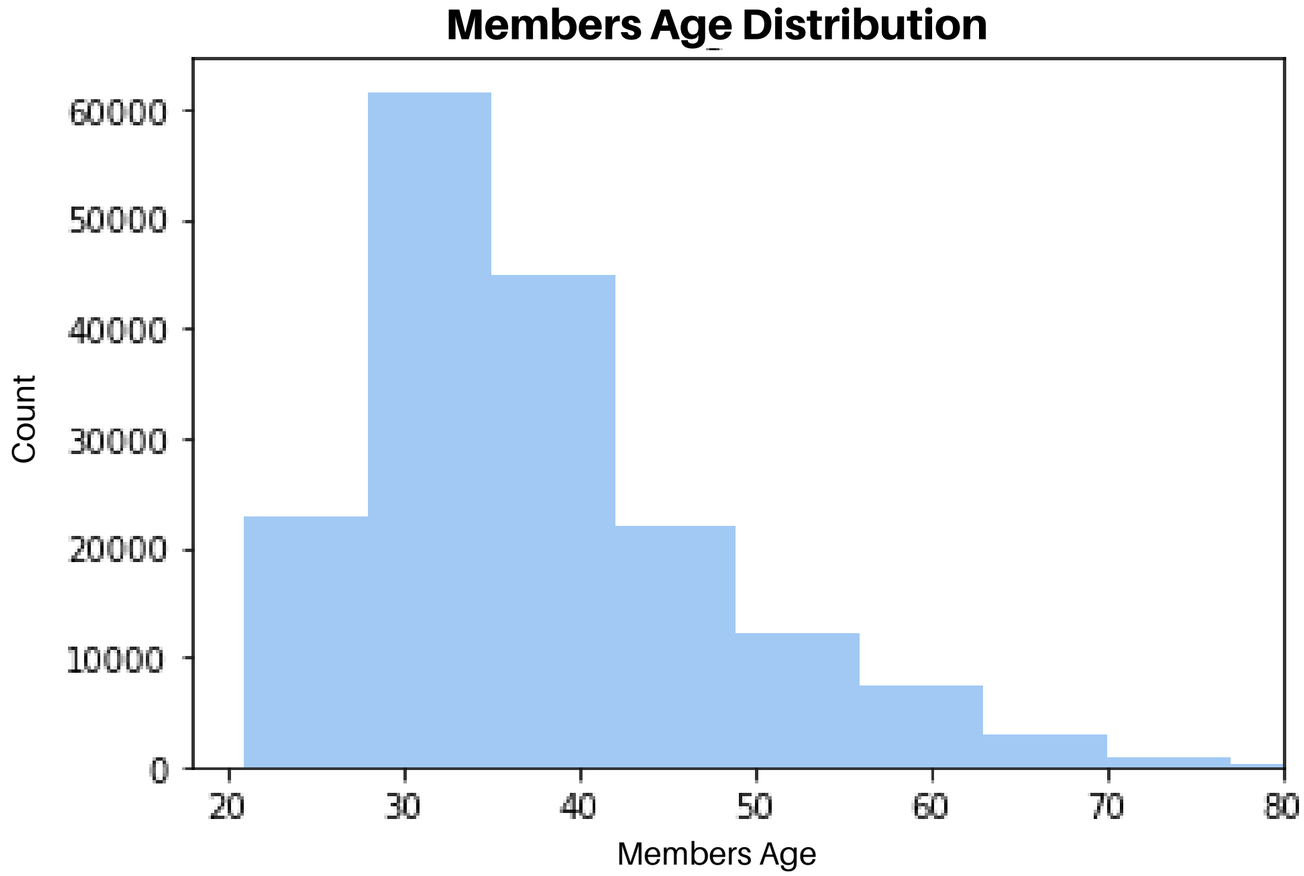
Members Service Usage based on Their Age Groups

Members with most trips are aged between 30 and 35 years old, which can be explained by the fact that this age group is the most present in the dataframe with a peak for the 34 years old.

The trips counts and durations is affected by the members' ages as follows:

- The longest ride belongs to people identified as other aged between 30 and 40 followed by those aged above 50 years old, in the same gender type.
- Members above 40, in almost all gender types, have longer trip duration than younger members.
- Female members in 50 - 60 age groups have longer rides than other younger females and all males.

Members Service Usage based on Their Age Groups

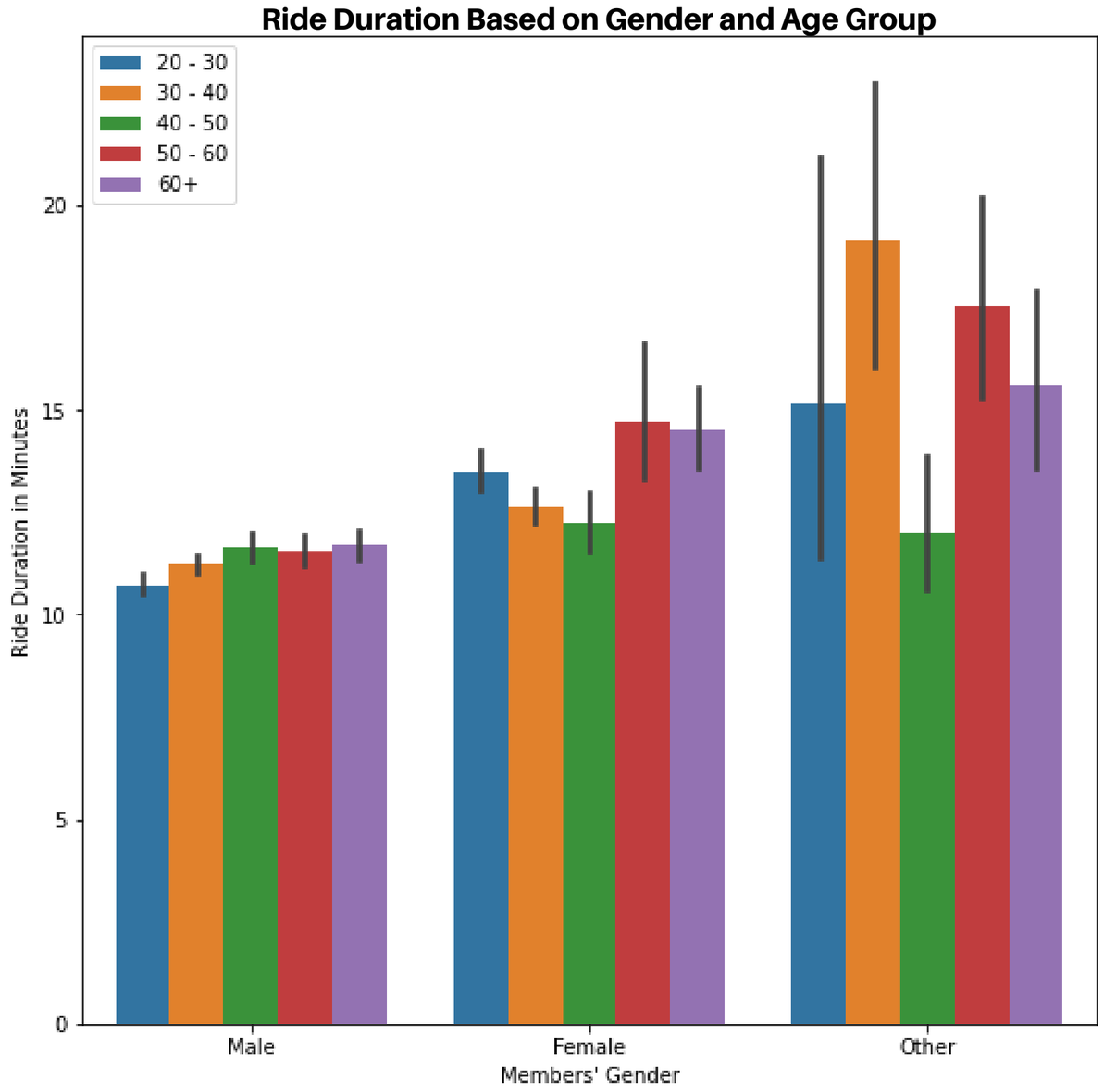


The histogram shows a peak between **30 and 35 years old.**

And according to the previous chart, we can conclude that largest group of members are aged **34 years old.**

Therefore, we can assume that the majority of the members use the bikes as a necessary transportation mean (work, for instance).

Members Service Usage based on Their Age Groups



The bar plot shows that people, identified as **other gender**, aged in the **30-40 age** range have the **longest bike rides** reaching 20 minutes.

Additionally, people aged between **50 and 60** have **long bike rides** more than younger people in all gender.