Communicating Data Finding : Ford GoBike System Data

by Dain Russell

Investigation Overview

In this investigation, I wanted to look at different features of the bikeshare dataset that influence trip duration.

Dataset Overview

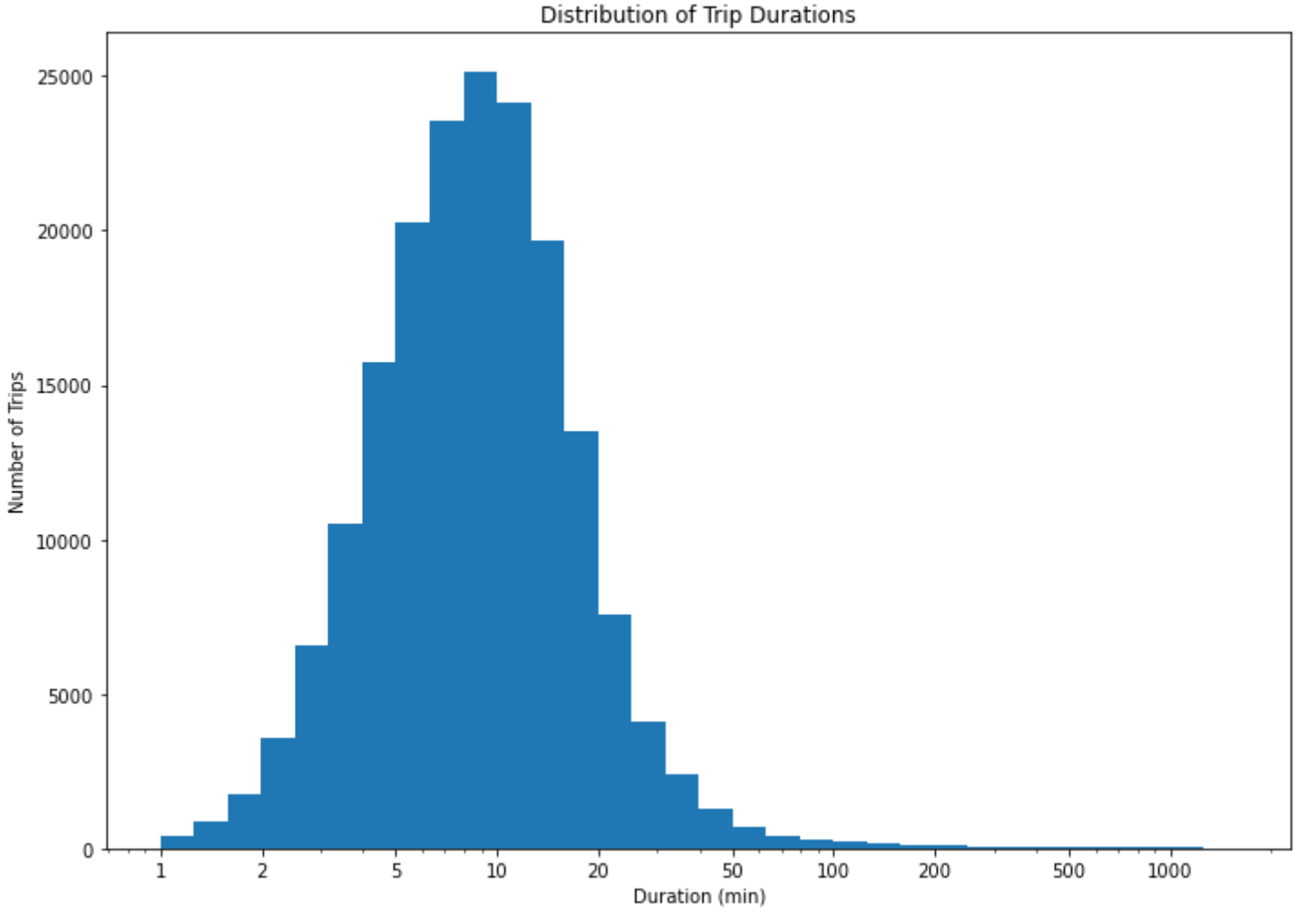
The data consisted of trip start and end times and other variables.

- There were 4646 bikes.
- There are 183412 fordgobike trips in the cleaned dataset with 16 specifications or columns.

Connect to the dataset

	duration_sec	start_time	end_time	start_station_id	start_station_name	start_station_latitude	start_station_longitude	end_s
0	52185	2019-02-28 17:32:10.1450	2019-03-01 08:01:55.9750	21.0	Montgomery St BART Station (Market St at 2nd St)	37.789625	-122.400811	13.0
1	42521	2019-02-28 18:53:21.7890	2019-03-01 06:42:03.0560	23.0	The Embarcadero at Steuart St	37.791464	-122.391034	81.0
2	61854	2019-02-28 12:13:13.2180	2019-03-01 05:24:08.1460	86.0	Market St at Dolores St	37.769305	-122.426826	3.0
3	36490	2019-02-28 17:54:26.0100	2019-03-01 04:02:36.8420	375.0	Grove St at Masonic Ave	37.774836	-122.446546	70.0
4	1585	2019-02-28 23:54:18.5490	2019-03-01 00:20:44.0740	7.0	Frank H Ogawa Plaza	37.804562	-122.271738	222.0

Distribution of Bike Trip Duration



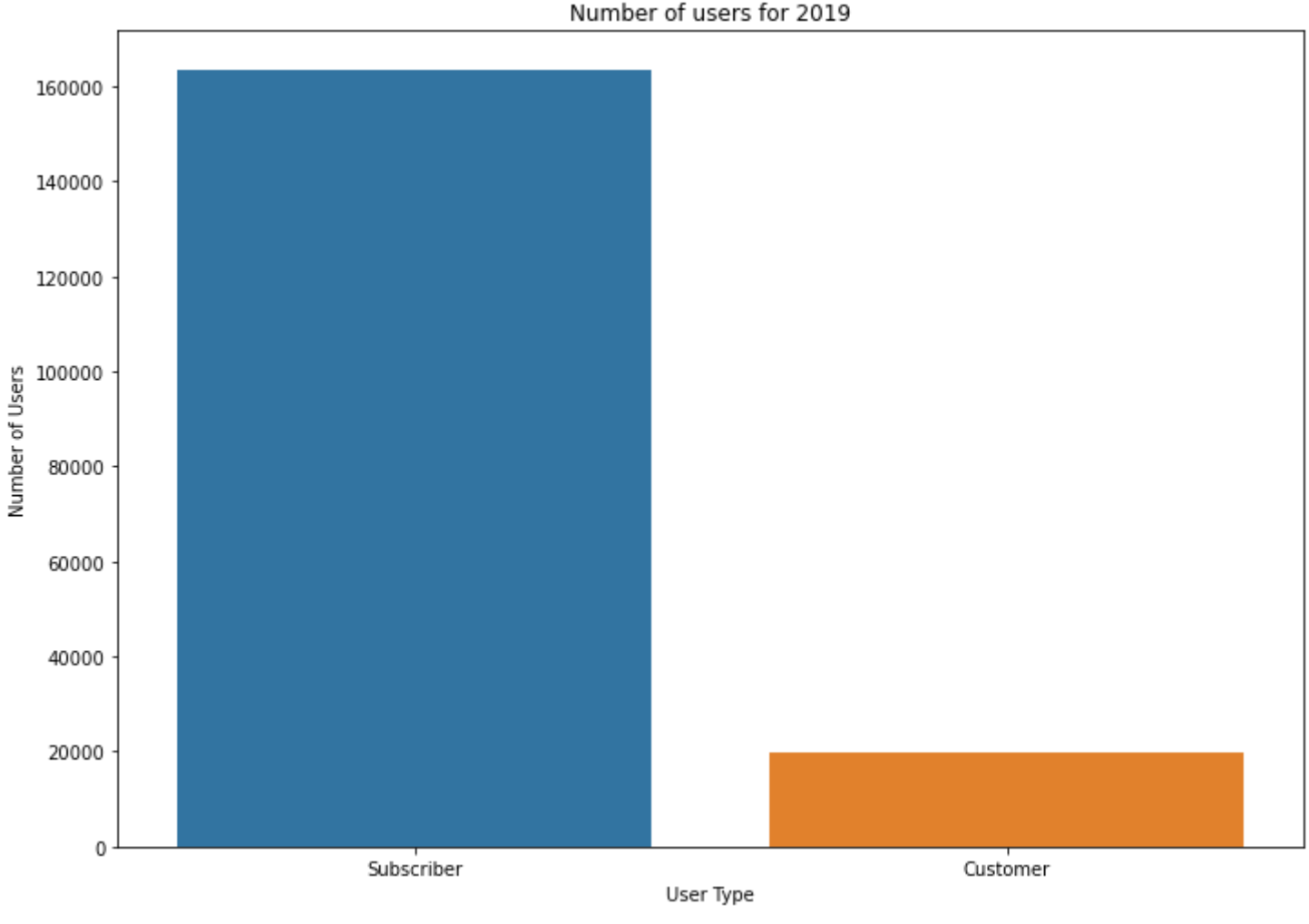
Observations

- Long tailed, bimodal distribution peaking between 8 and 10.
- The average bike trip is 12 minutes.
- The standard devaition is 29.9.
- 25% of the trips are over 5 minutes, 50% over 8 minutes and 75% over 13 minutes.
- The longest trip is 1424 minutes and the shortest being one minute.

With this distribution I really wanted to see how the trip was distributed when looking at just the length of a ride. According to baywheels.com, trips longer than 45 minutes will incur an extra 3 dollars per each additional 45-minutes for those with an annual pass. A single ride cost 3 dollars and only last 30 minutes This fact probably factors into how long a user will ride and why the data shows most trips are below 45 minutes.

Distribution of Users

User Type Key 1.Customer = 24-hour pass or 3-day pass user 2.Subscriber = Annual Member

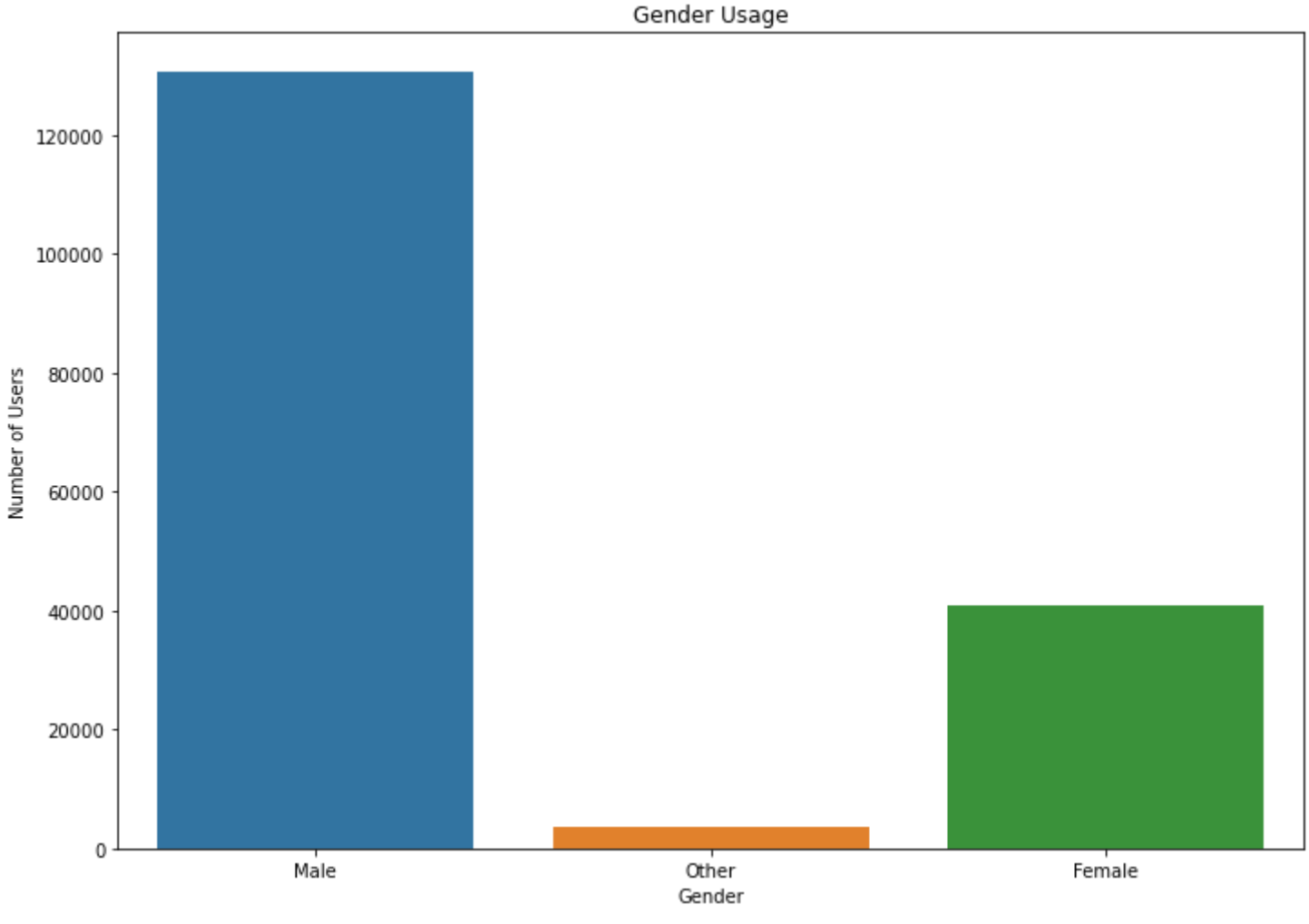


Observations

- Users included Customer that represents those with a 24-hour pass or 3-day pass, and subscribers those with anual membership.
- The bar chart shows over 160,000 subscribing users and 20,000 customers.
- Most users are actually subscribers with annual memberships.

Due to the fact that there are so many more people subscribing that using a single day or 3 day pass I think can be attributed to the cost savings. I think there are so many users that are riding bikes daily, tht it makes more cost sense to sign up for an annual membership.

Distribution of Gender

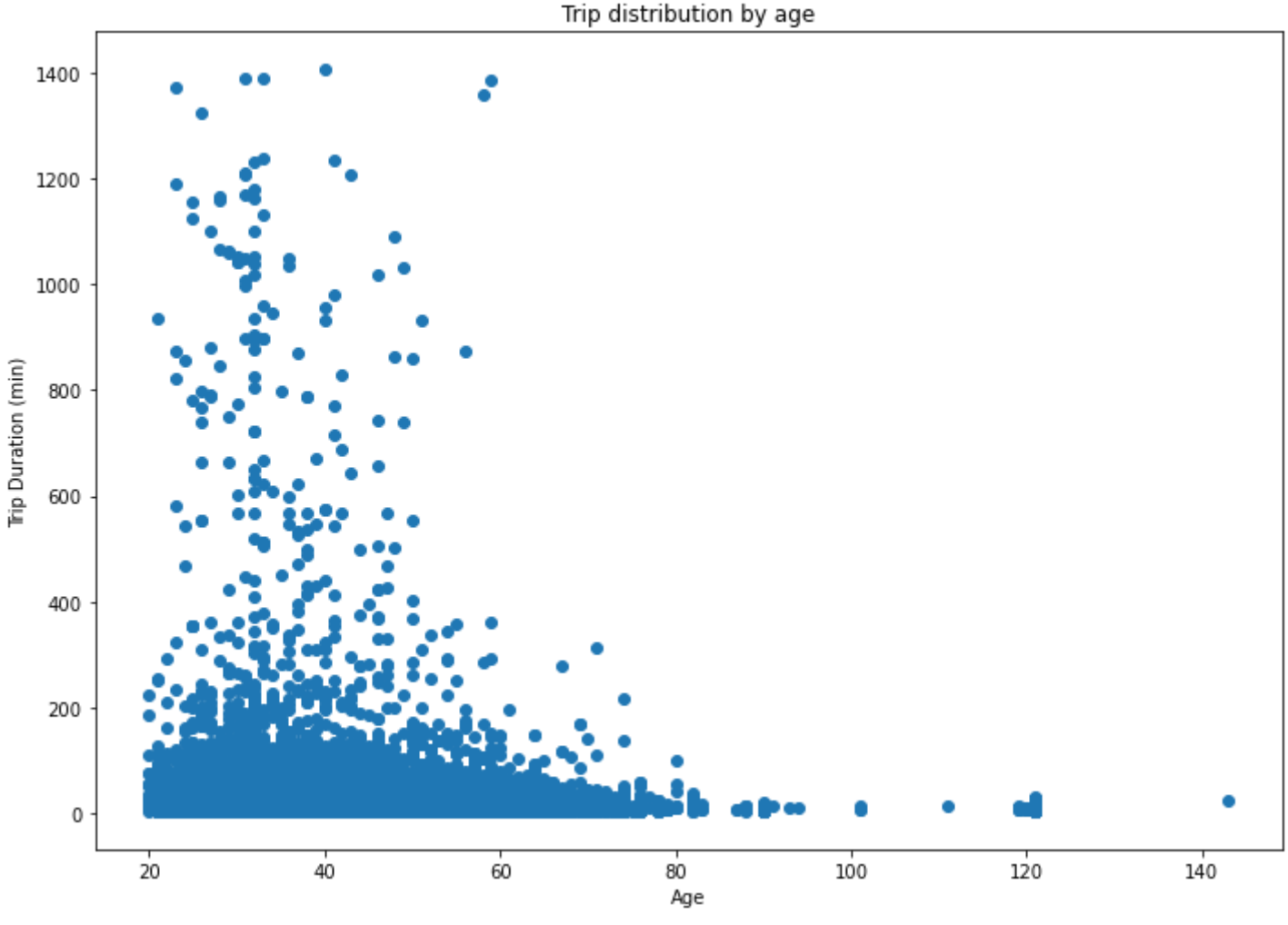


Observations

- Males use the bike service overwhelmingly more than females and other genders.
- Over 120,000 males used the service in 2019.

The intent was to see how each gender type is spread out and how much of each group is using the bike share service.This distribution reveals that males represents most of the riders. Could this be because males are more into this type of fitness activity. We should have a clearr picture if we factor in more variables like age to see what group of males this number represents.

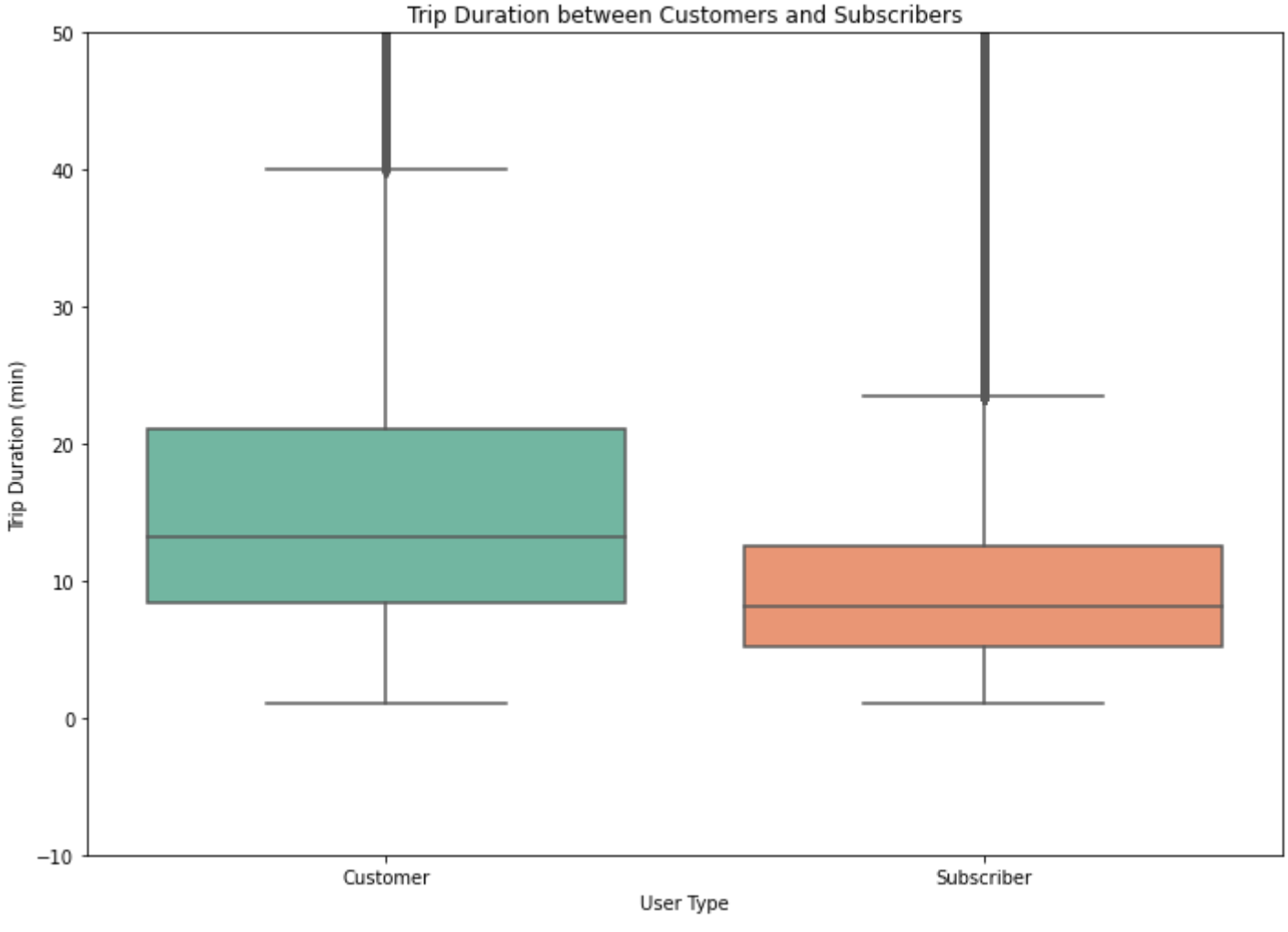
Trip Duration and Age



Observation

So with this plot I wanted to look at how trip duration is affected by the age of the user. You can observe that there is a negative correlation where the age decreases as the trip duration increases. The concentration of rides are for persons between ages 25 and 45 showing the inverse relationship between age and the trip duration. It is not surprising that biking trip duration would decrease as you increase in age. Fitness would be less common or intense mong older groups so the bike trip would decrease among a older group.

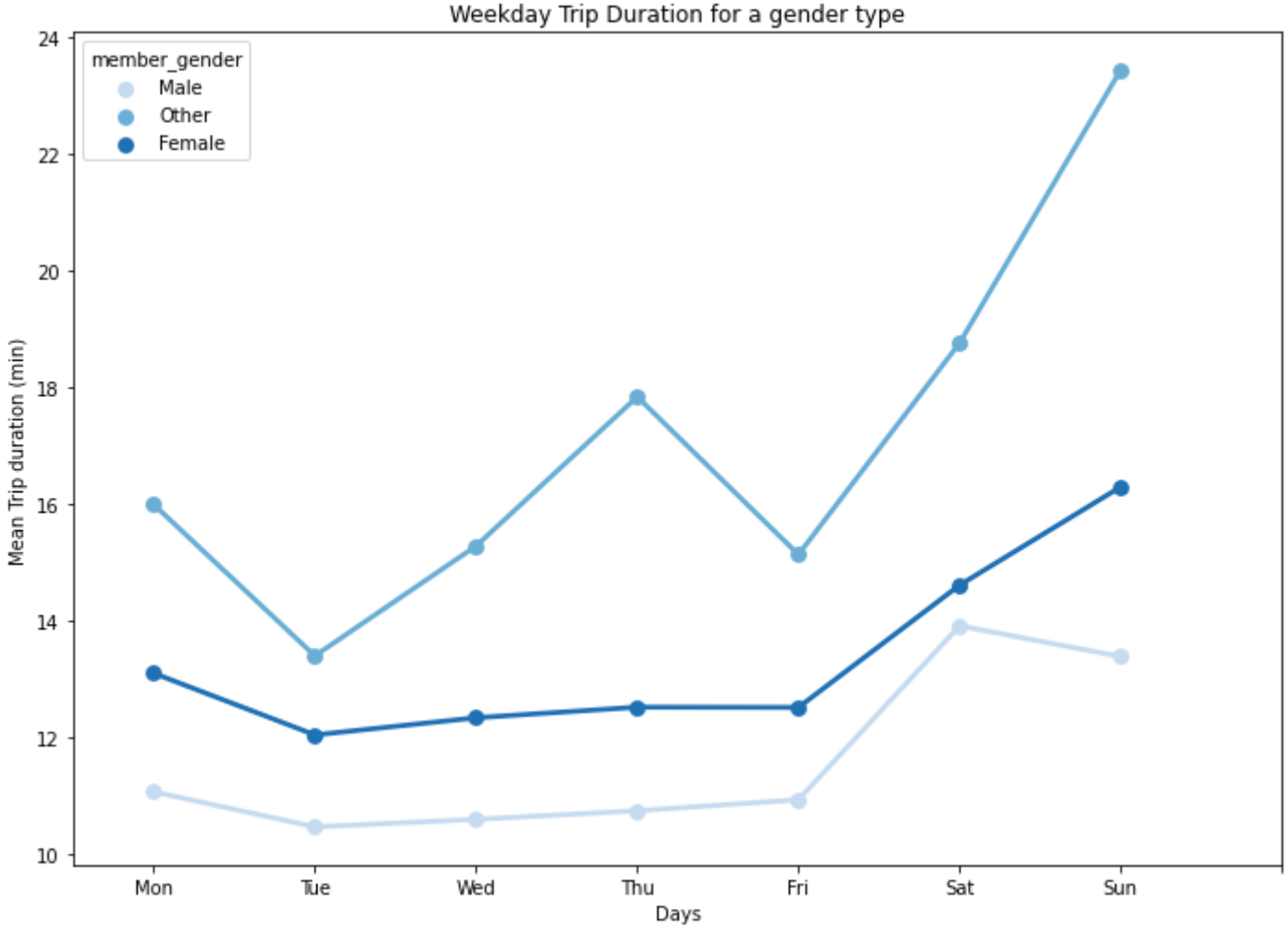
Trip Duration and User Type



Observation

The customer is spending more time on a bike trip than subscribers. Could it be that because customers spend a litle more on a single or 3 day pass that they might be trying to maximize the time that they ride? That may be the reason that there is more than twice as much customers versus subscribers on average riding bikes.

Trip Duration vs Weekday and Gender Type



Observation

- The trip duration start trending up on the weekends from Thursdays to Saturdays.
- Males still have the shortest bike trip.

All the gender groups peaked on Fridays to Sunday and increased the length of their bike trips. That makes sense as those riders probably have more time to dedicate to riding. They are probably riding at a different time of day. I notice that although males increase their trip on Fridays to Saturday they are recording the shortest trips of all gender type.

The End