Ford GoBike Data

Exploratory analysis

Contents

- Introduction

- Univariate exploration

- Bivariate exploration

- Multivariate exploration

Introduction

- Dataset containing information on over 3023992 bike trips between 2017-2019.

- Columns include: trip duration, start and end coordinates, start time, user age and user gender.

- Can be found on Kaggle: https://www.kaggle.com/victorinoeng/fordgobike20172019

Introduction

Exploratory data analysis

- This slide deck presents the results of an exploratory data analysis conducted on this dataset.

- The data is explored from a univariate, bivariate and multivariate perspective.

Univariate exploration
What are the months during which bikes are rented the most?

March and April are the most popular months for bike trips

May and June are the months during which the least bike trips were made.

number of Ford Gobike trips per month (2017-19)
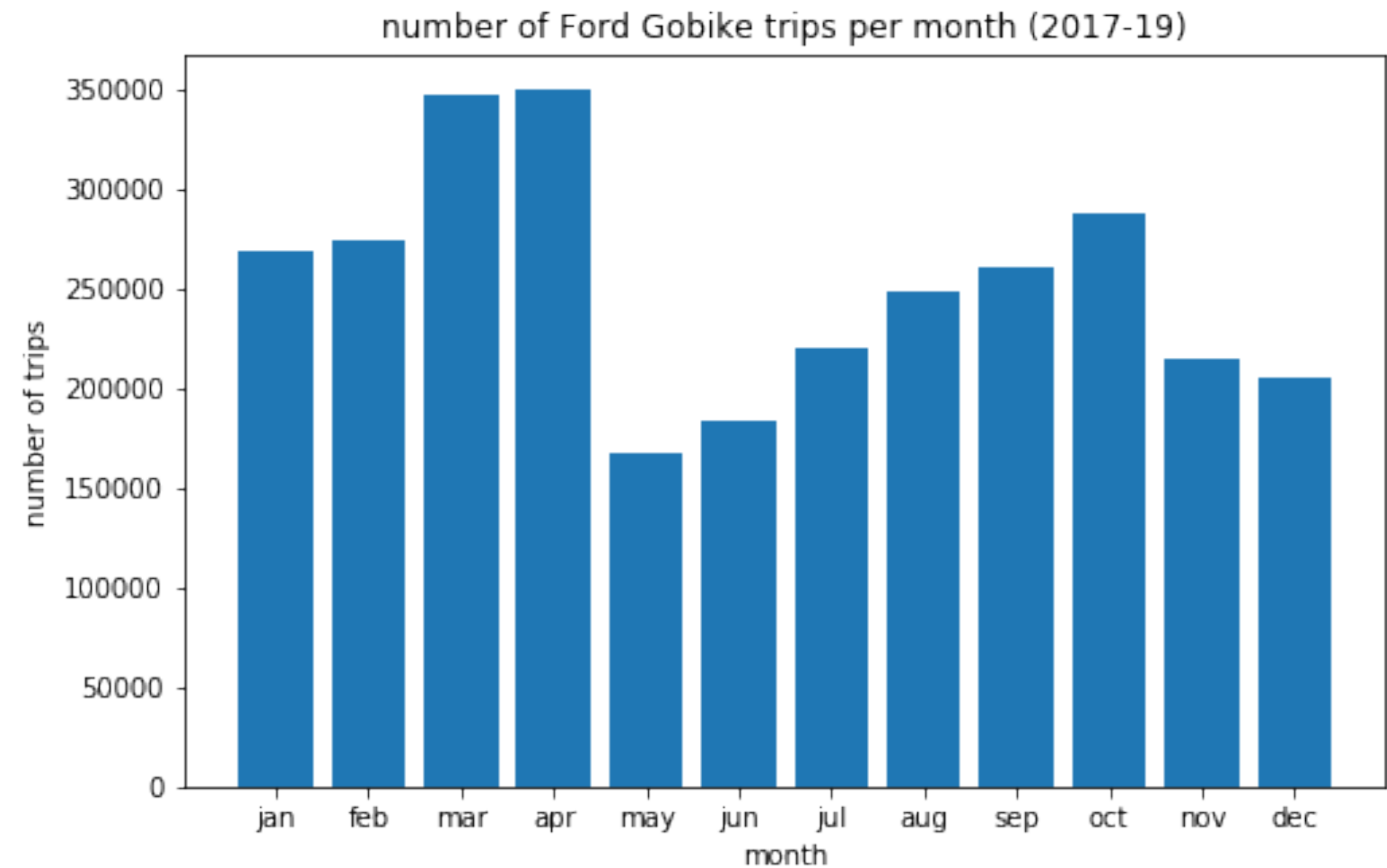
Univariate exploration
What are the days of the week during which bikes are rented the most?

There are significantly more bike trips made during the week than on the week-end.

Could this indicate the Ford GoBikes are mostly used for travelling to work?

Number of bike trips per weekday (2017-19)
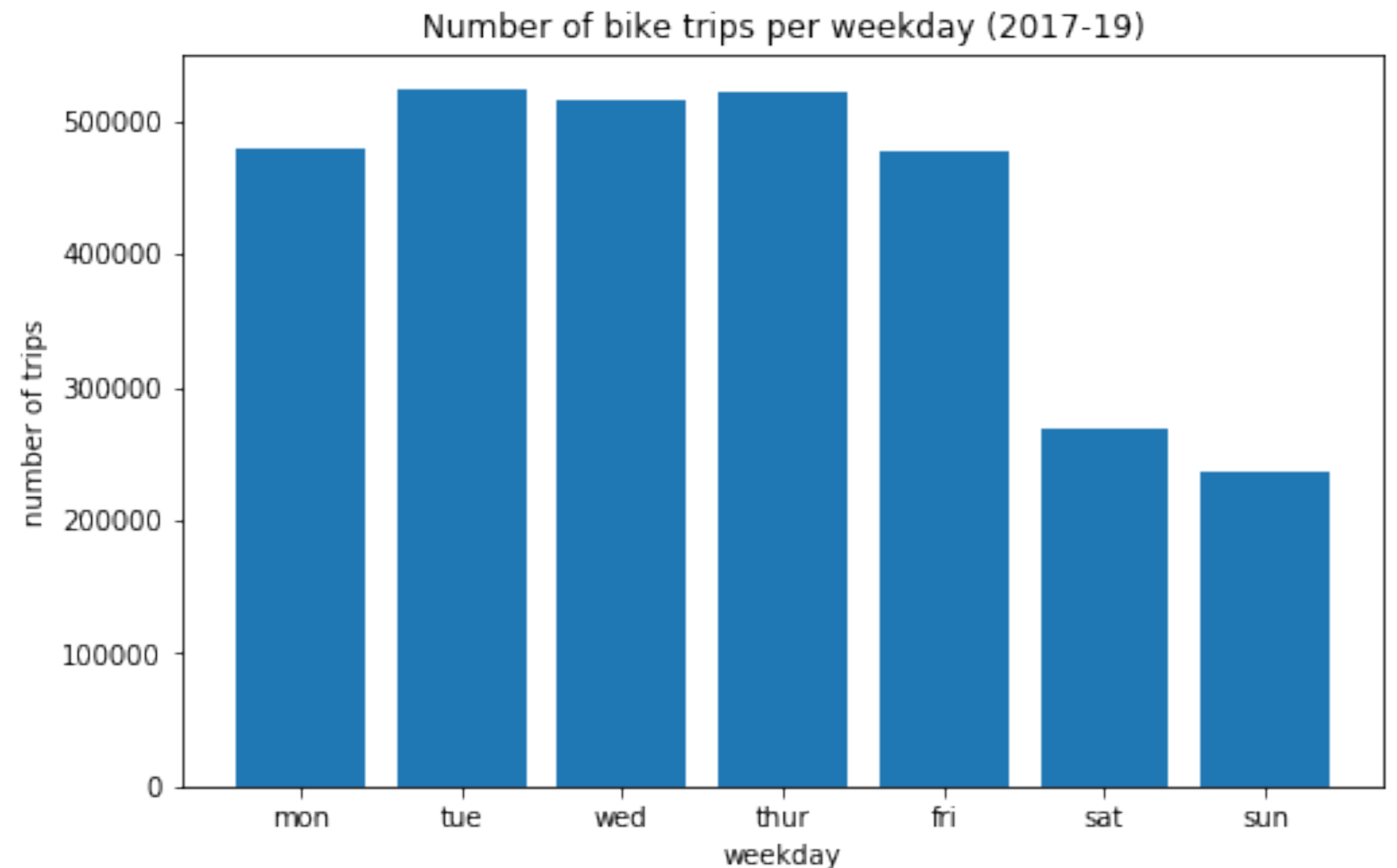
Univariate exploration
What is the trip duration distribution like?

On a log scale, the trip duration distribution looks normal with a mean of about 10min.

distriubtion of trip durations (2017-19) (log)
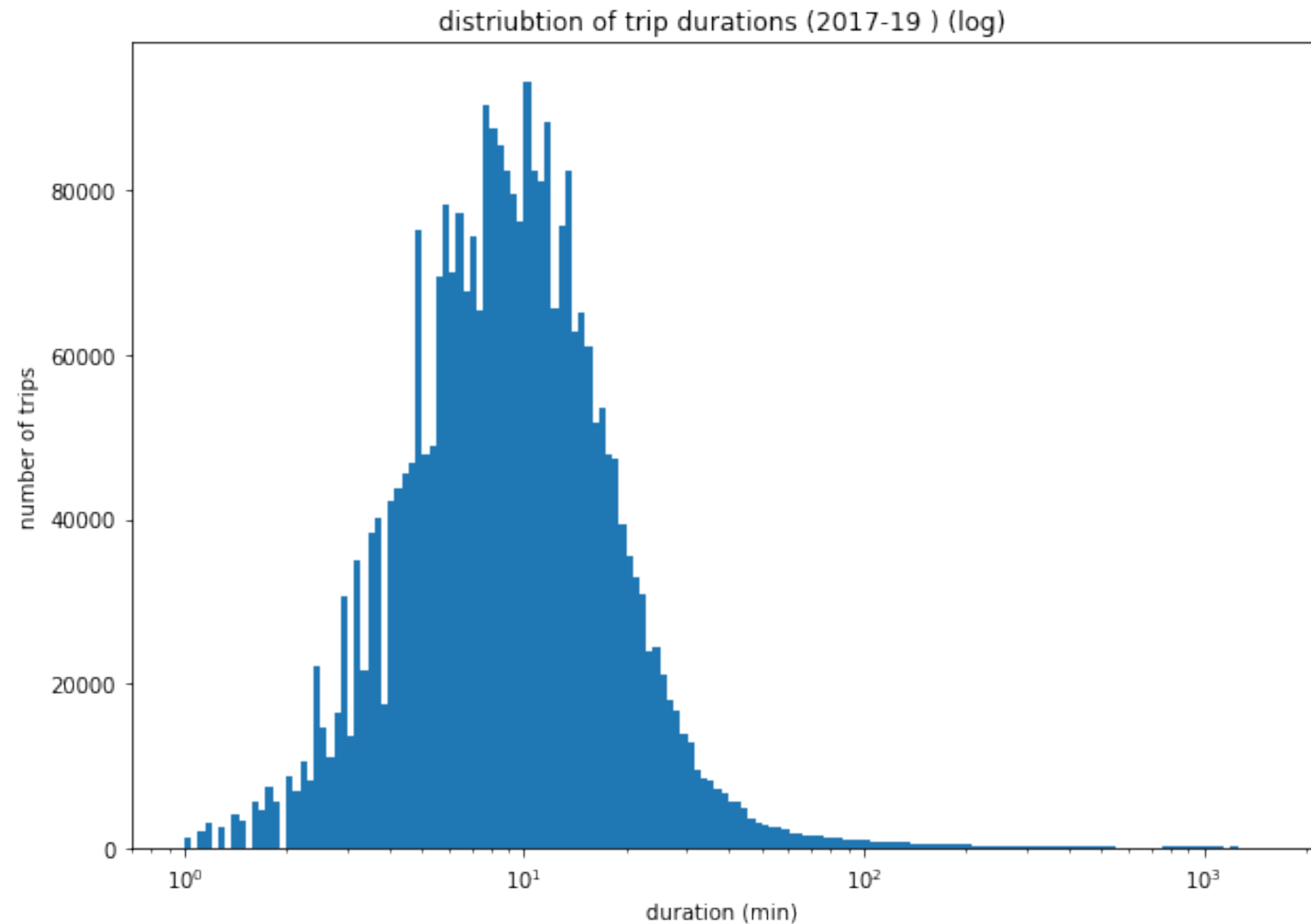
Bivariate exploration

Do subscribers take longer trips than non subscribers on average?

We see that on average, non subscribers take longer trips than subscribers.
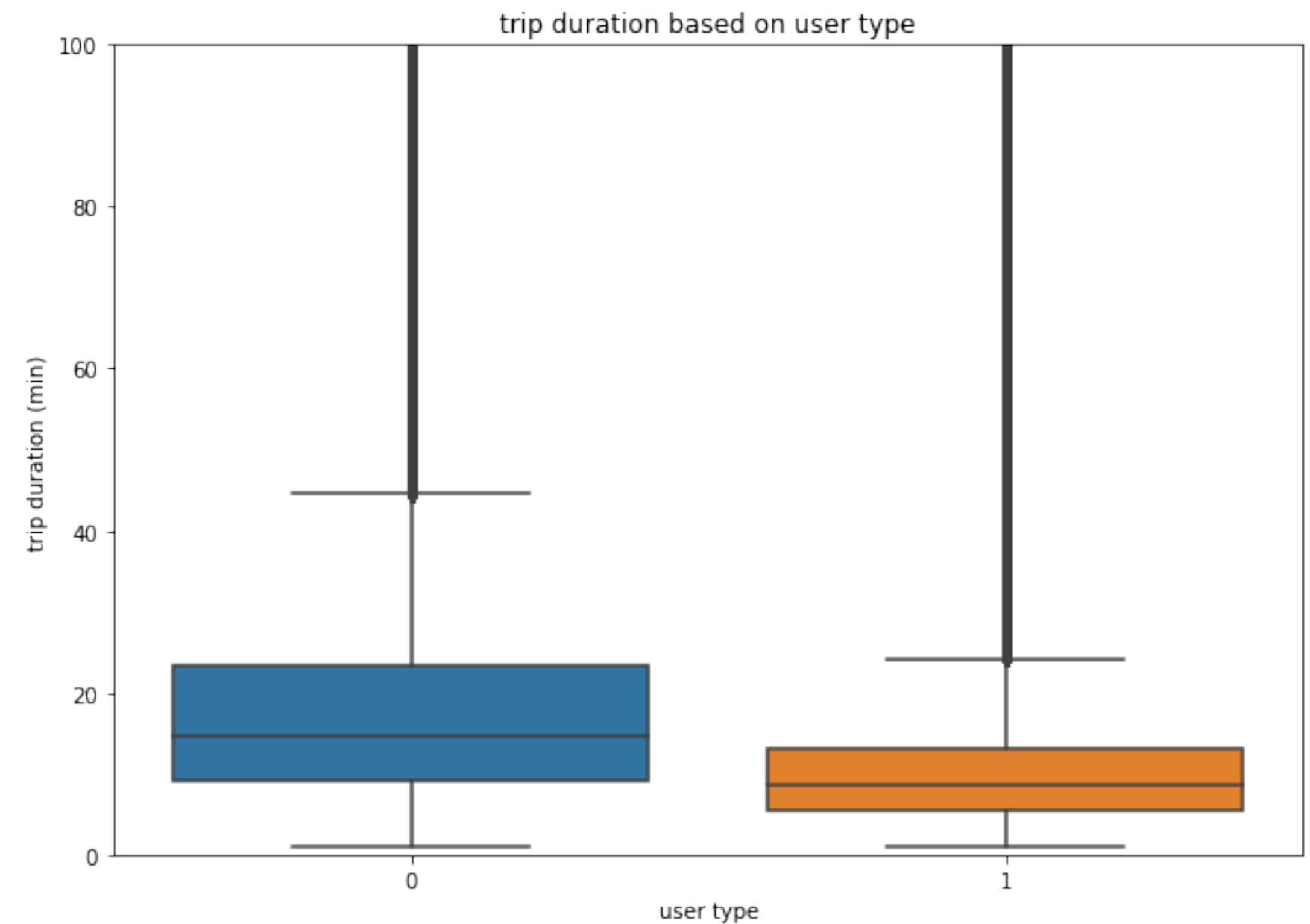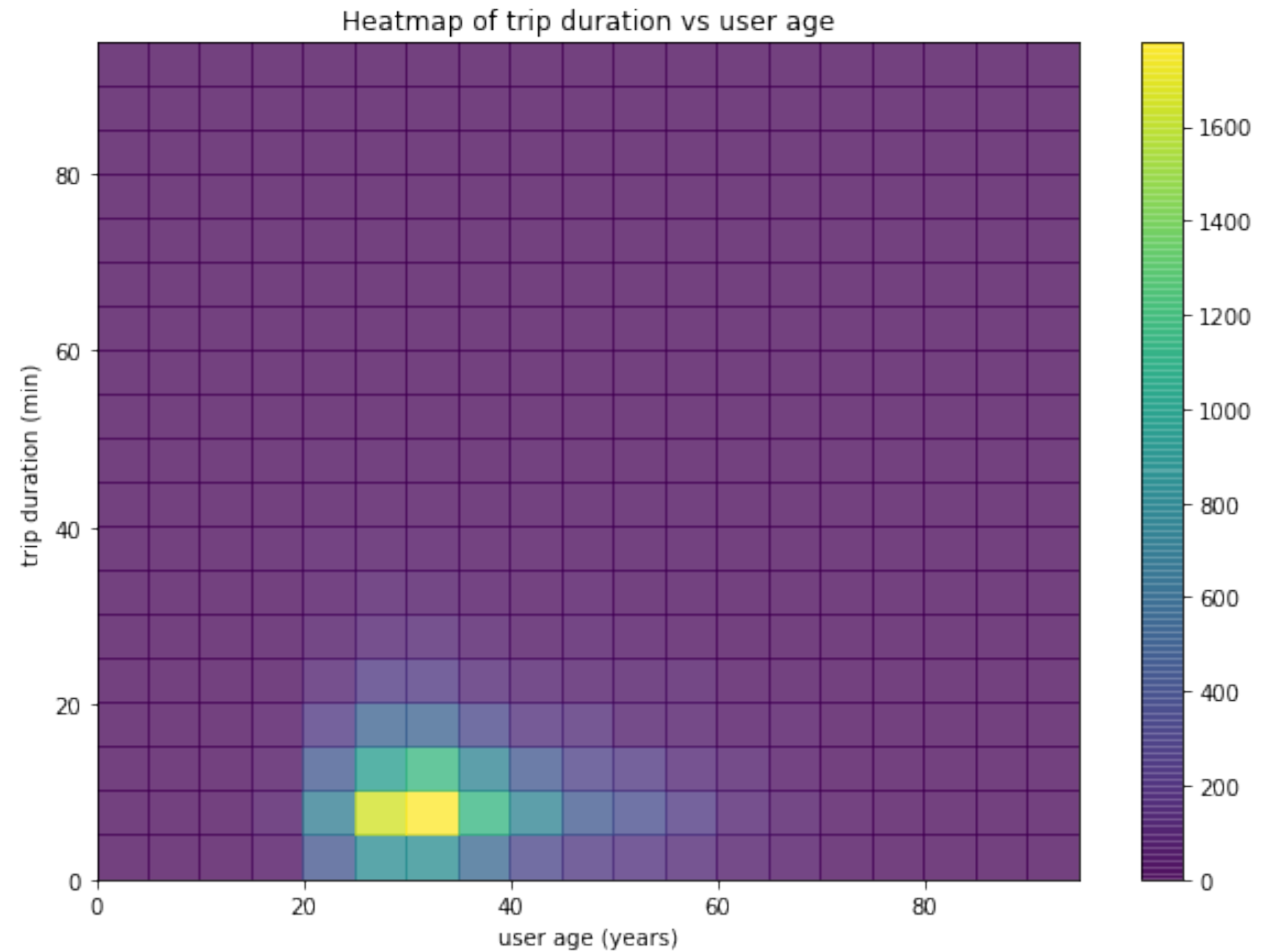
Bivariate exploration

Do younger users make longer trips than older users on average?

The heat map indicates that most trips are made by users aged 20-40 and last between 5-20min on average.

Heatmap of trip duration vs user age

Bivariate exploration

Do we have evidence to support the claim that Ford GoBikes are used mostly for work commute?

These results are interesting: we see a high amount of bikes are rented on weekdays between 8-10AM and 4-6PM. These are usual commute times.

We can say that overall, Ford GoBikes are mostly used for work travel.

Heatmap of day of the week and trip start time
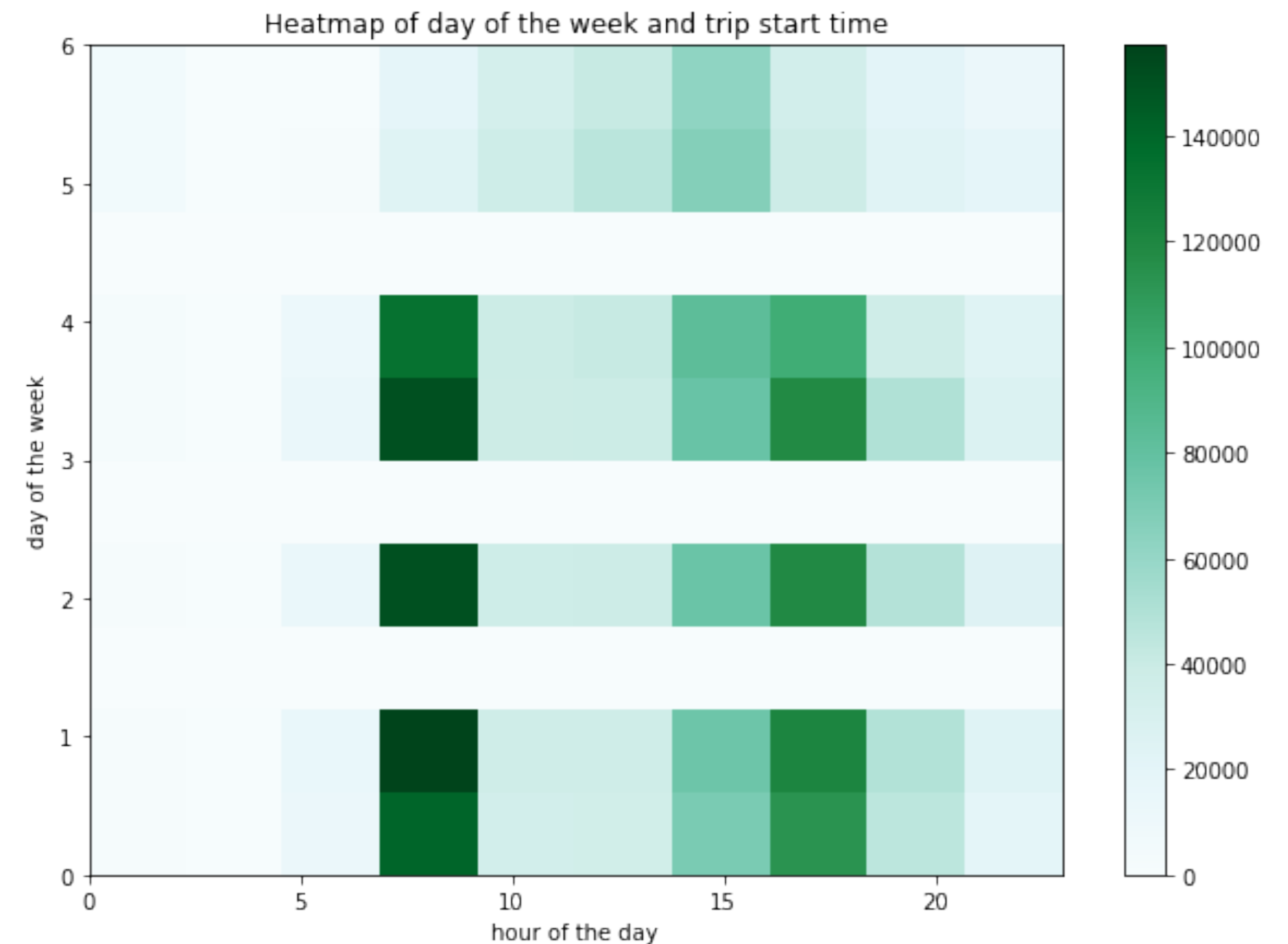
Multivariate exploration

Is there a correlation between average trip speed, user age and gender?

For gender: we notice that on all age brackets, men are slightly faster than women.

For age: we notice a peak in trip speed for both men and women between 20-50years and then an overall decline in velocity.

Averge trip speed versus age and gender