Project: Communicate Data Findings

By Debarghya Biswas

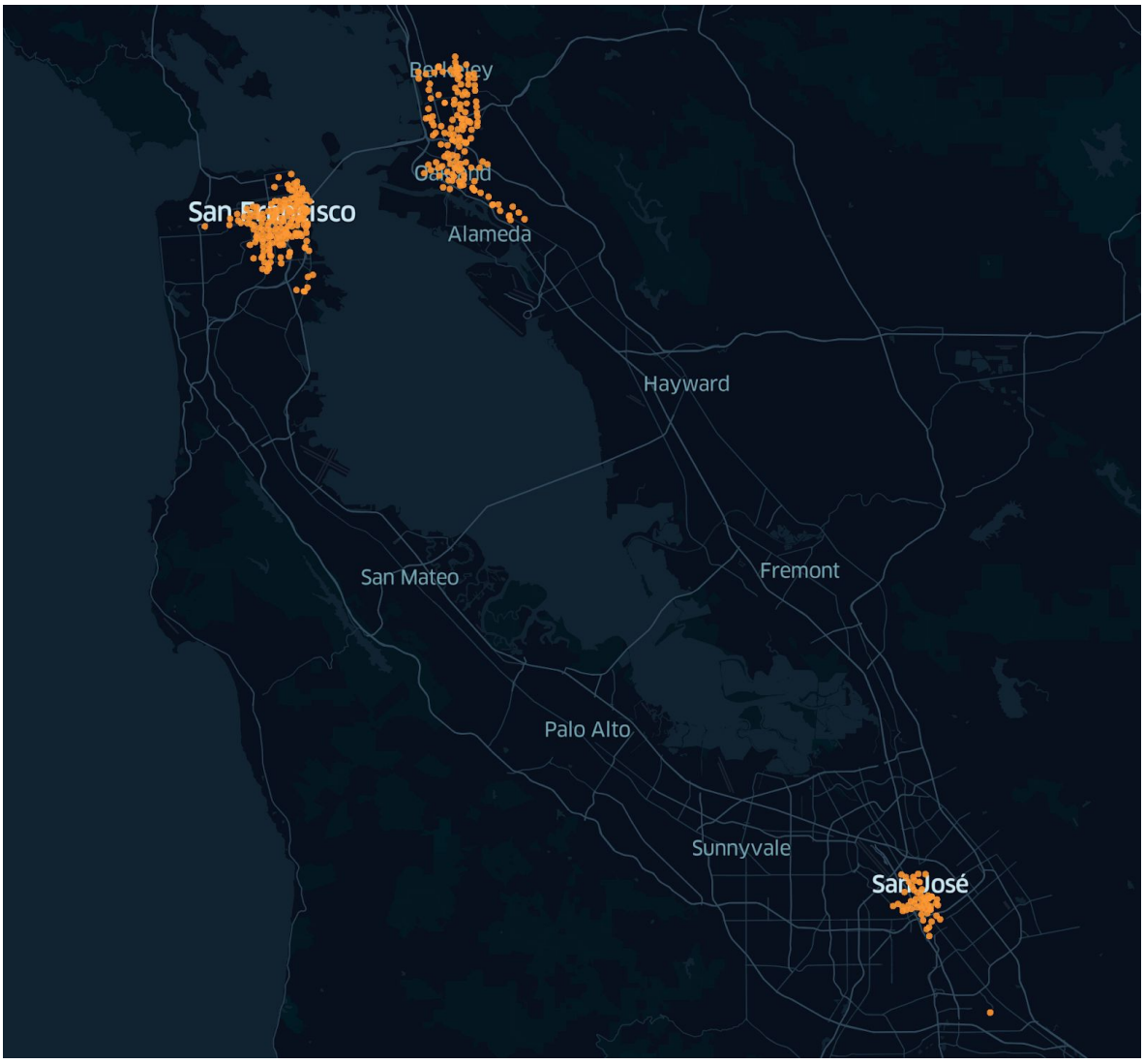
Exploring Ford GoBike Data

Data Sources

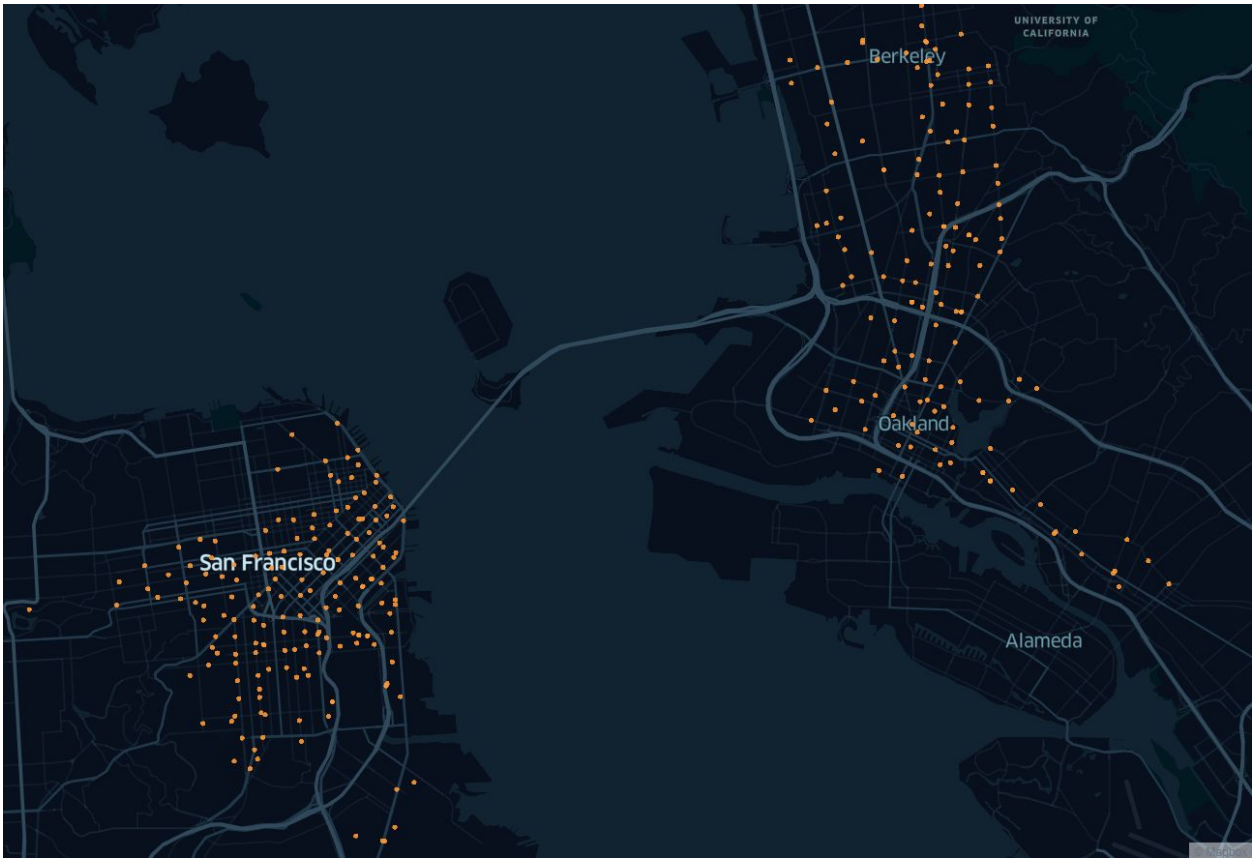
1. **Name:** result.csv
    - **Definition:** Ford GoBike System - Data
    - **Source:** https://www.fordgobike.com/system-data
    - **Version:** Files from 01.2018 - 02.2019

Source: kepler.gl

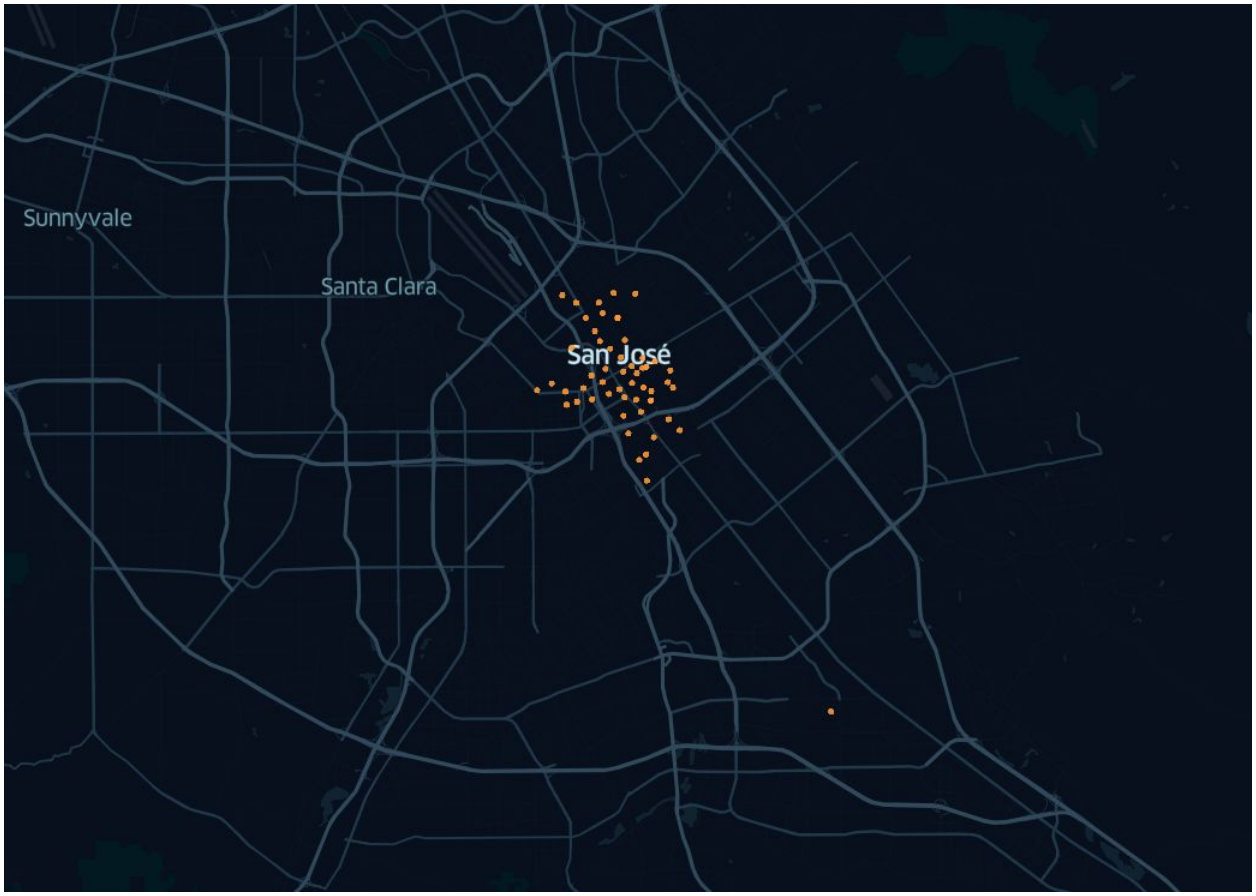
This is on what we will take a closer look on - San Francisco, East Bay, and San José:



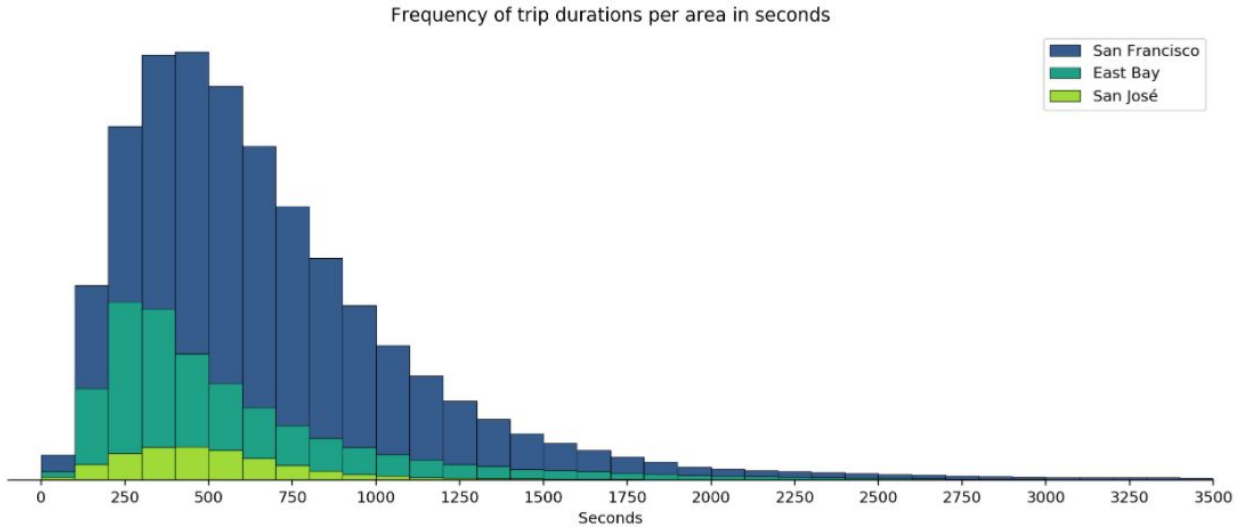
San Francisco and East Bay



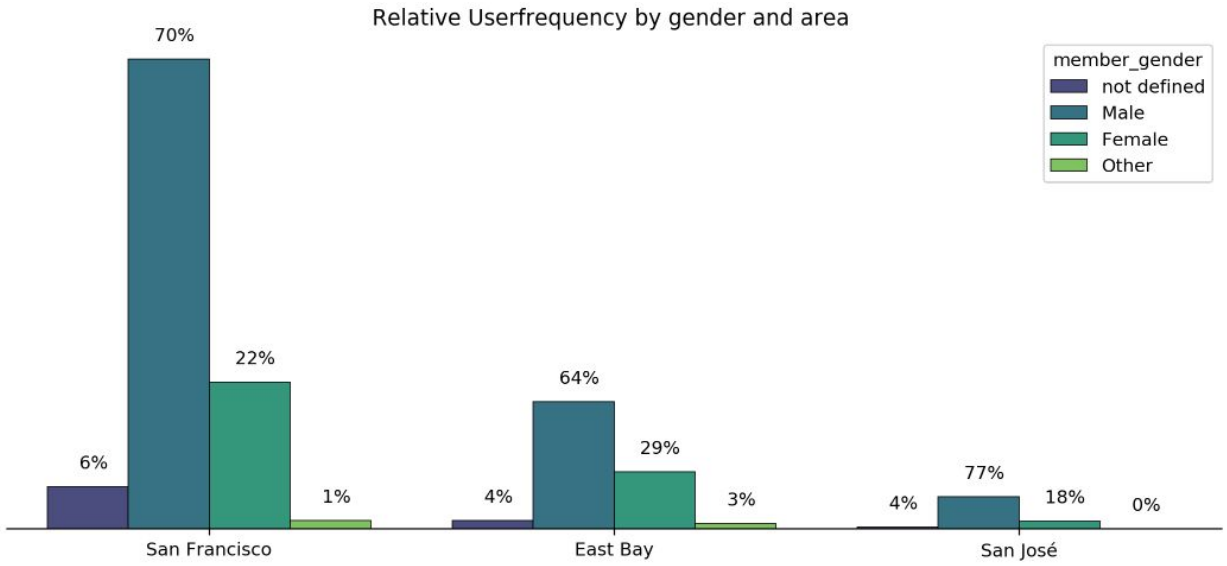
San José



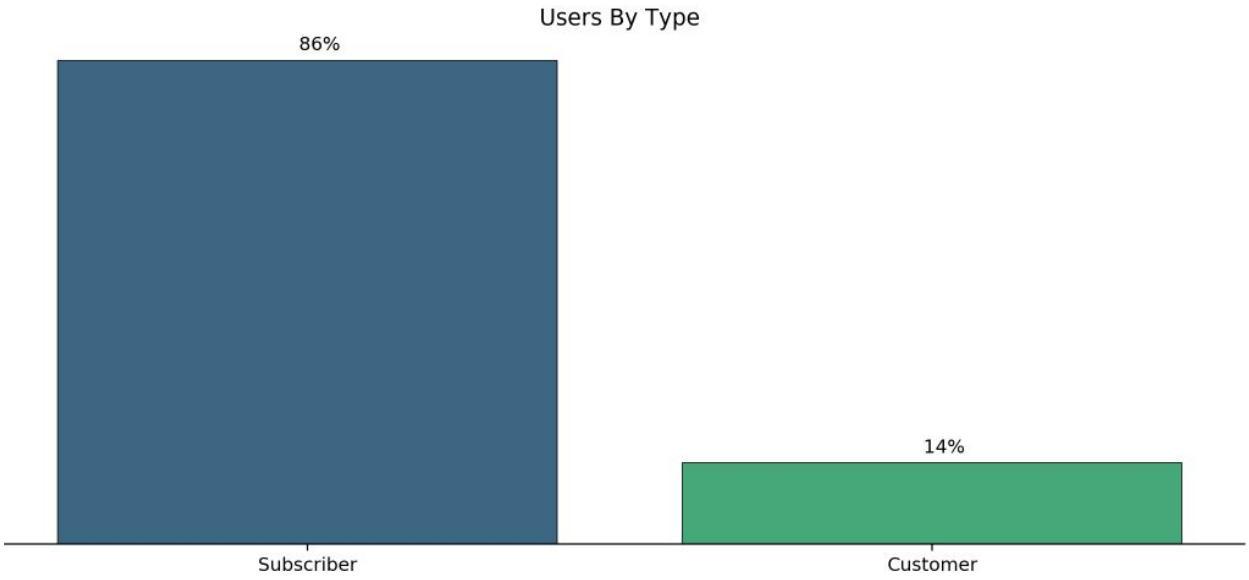
Who are the people that are using this service? Let's find out - at first, we will look on the average trip duration.



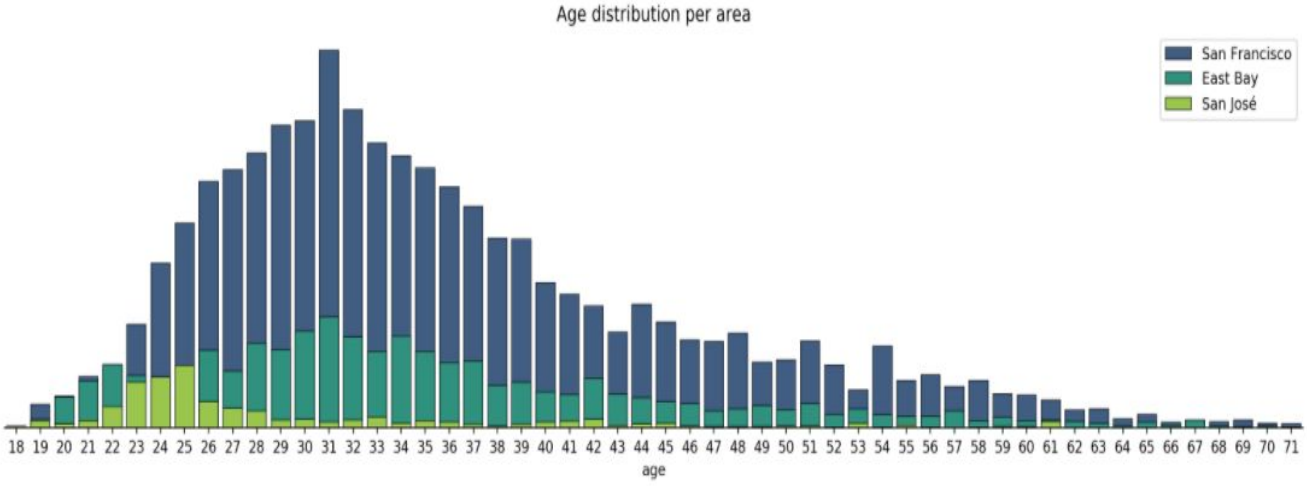
Looking at these, trends are looking similar to each other, although it seems like trips in East Bay are usually a little bit shorter in duration. What about gender?



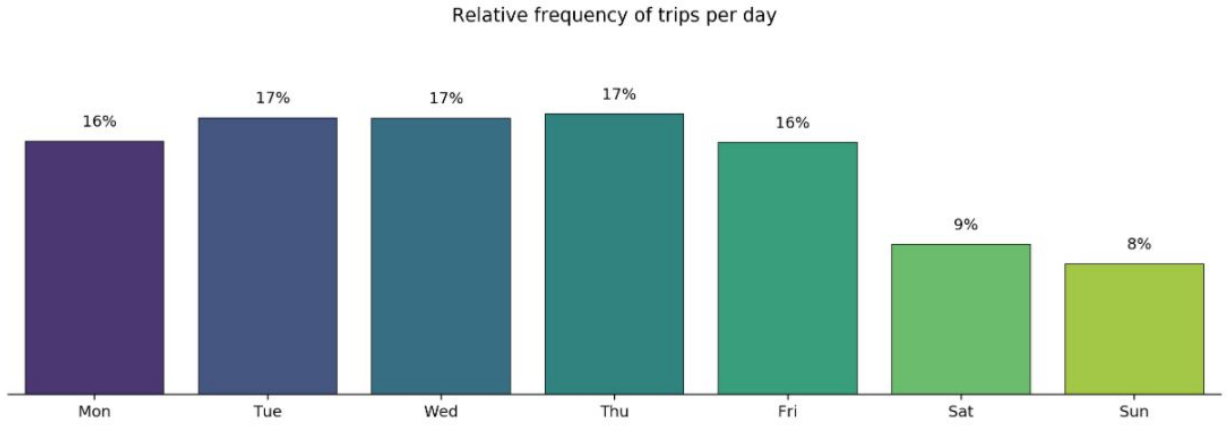
Looking at the plot and the relative frequencies, the male percentage is > 60% for all three areas. But are their users between them who are willing to subscribe to this service or are the most of the 'normal' customers?



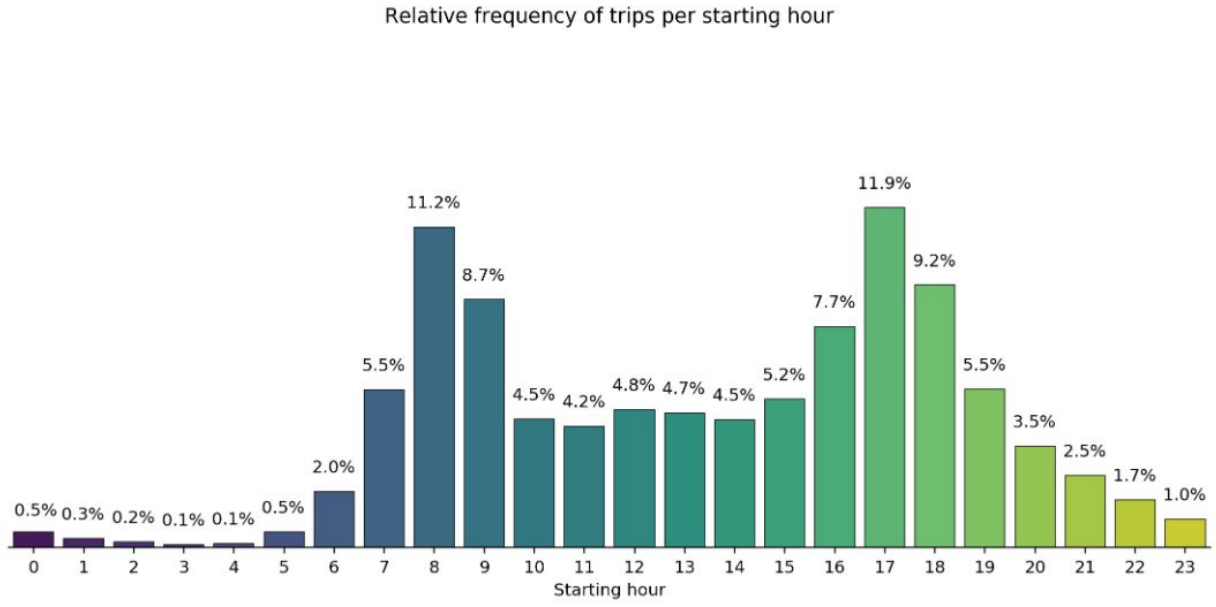
And how is the age of the users distributed?



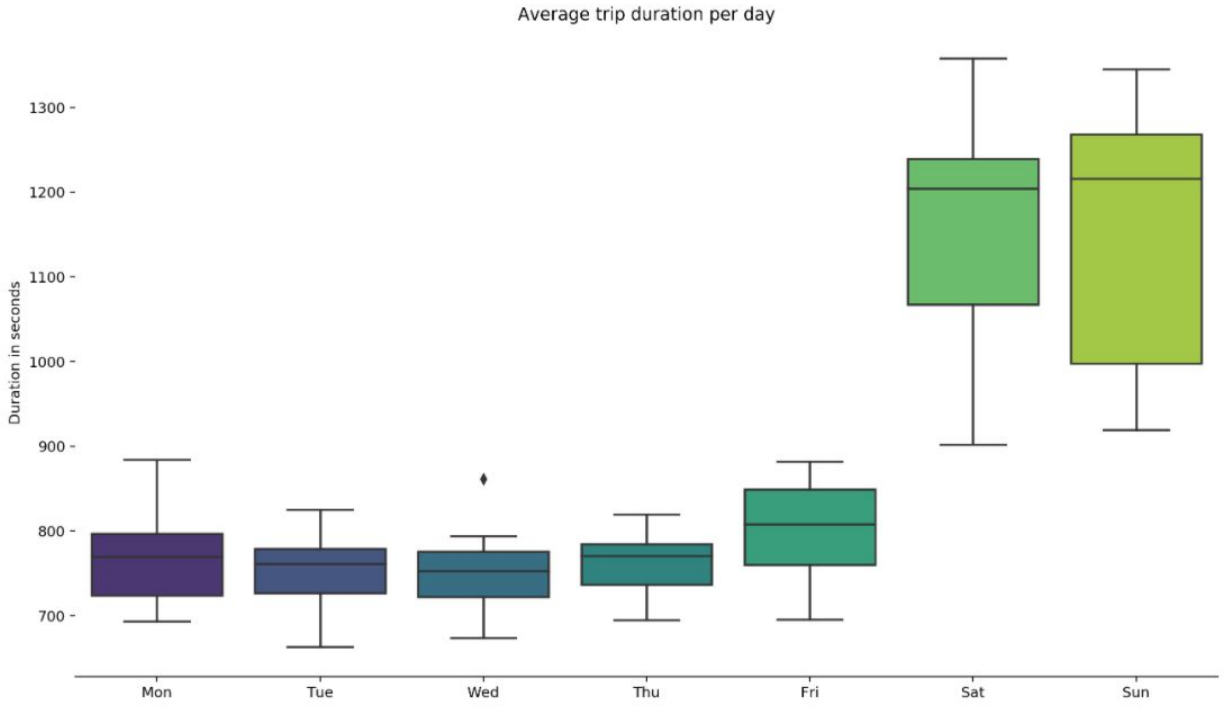
The East Bay age structure is broader than the one of San Francisco and San José has the youngest average group of users. The next plots will focus on the time components of our data. What about trips per day?



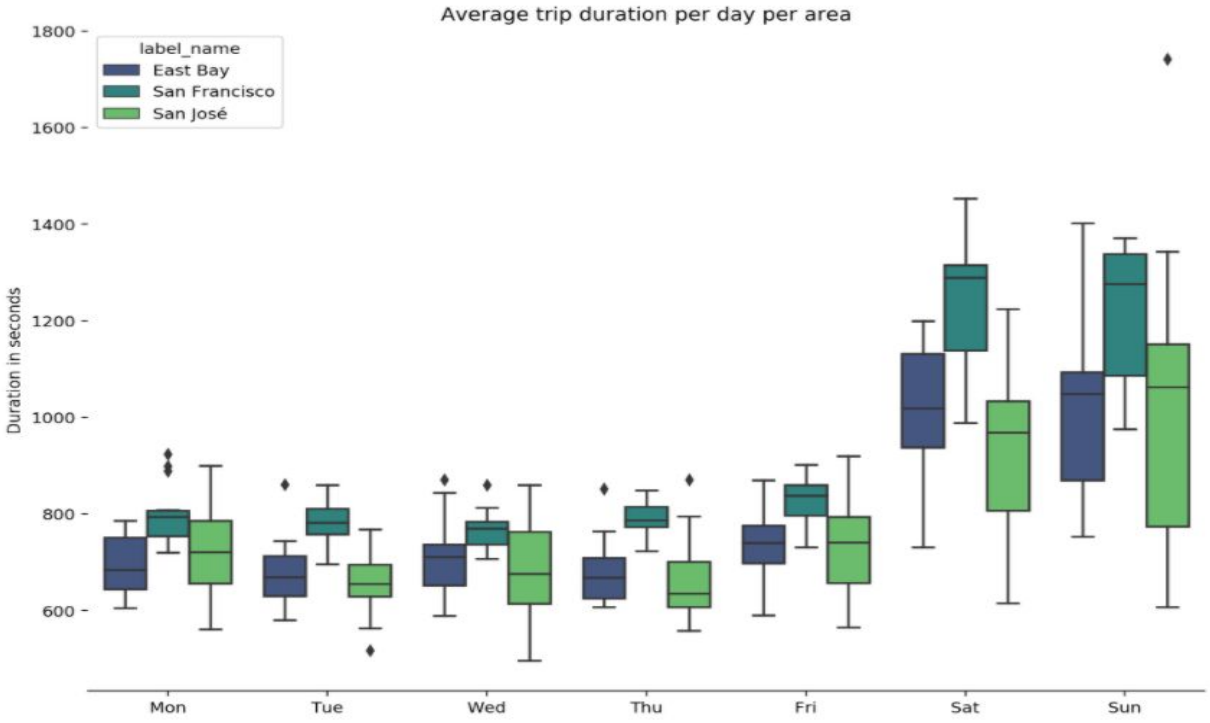
It looks like the users use bikes more frequently during the week than during the weekend. And when do they start their trips?



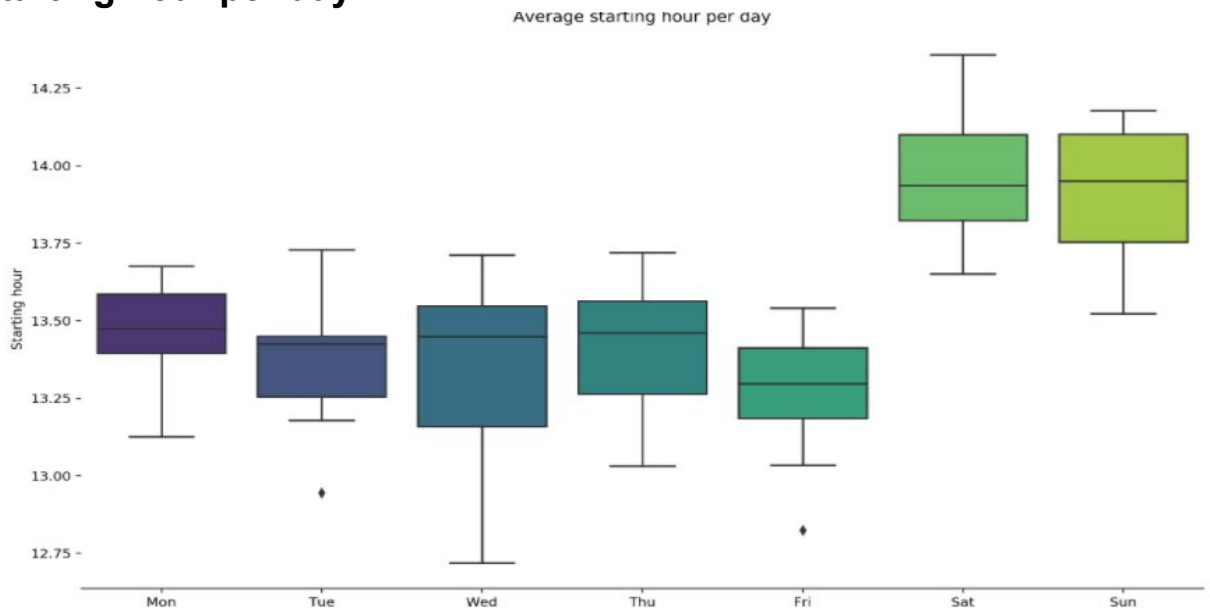
The most frequent starting hours are at 8 and at 17. Maybe people use it before and after work, which would make sense because we have a lot of subscribers in working age in our dataset. You only subscribe to something, if you want to use it regularly. The integration into the working/study life would make sense here! Now we will see if the average duration is dependent on the weekday.



This means, that the frequency of bike usage at the weekend is lower, but the average duration of each trip is greater than during the week!



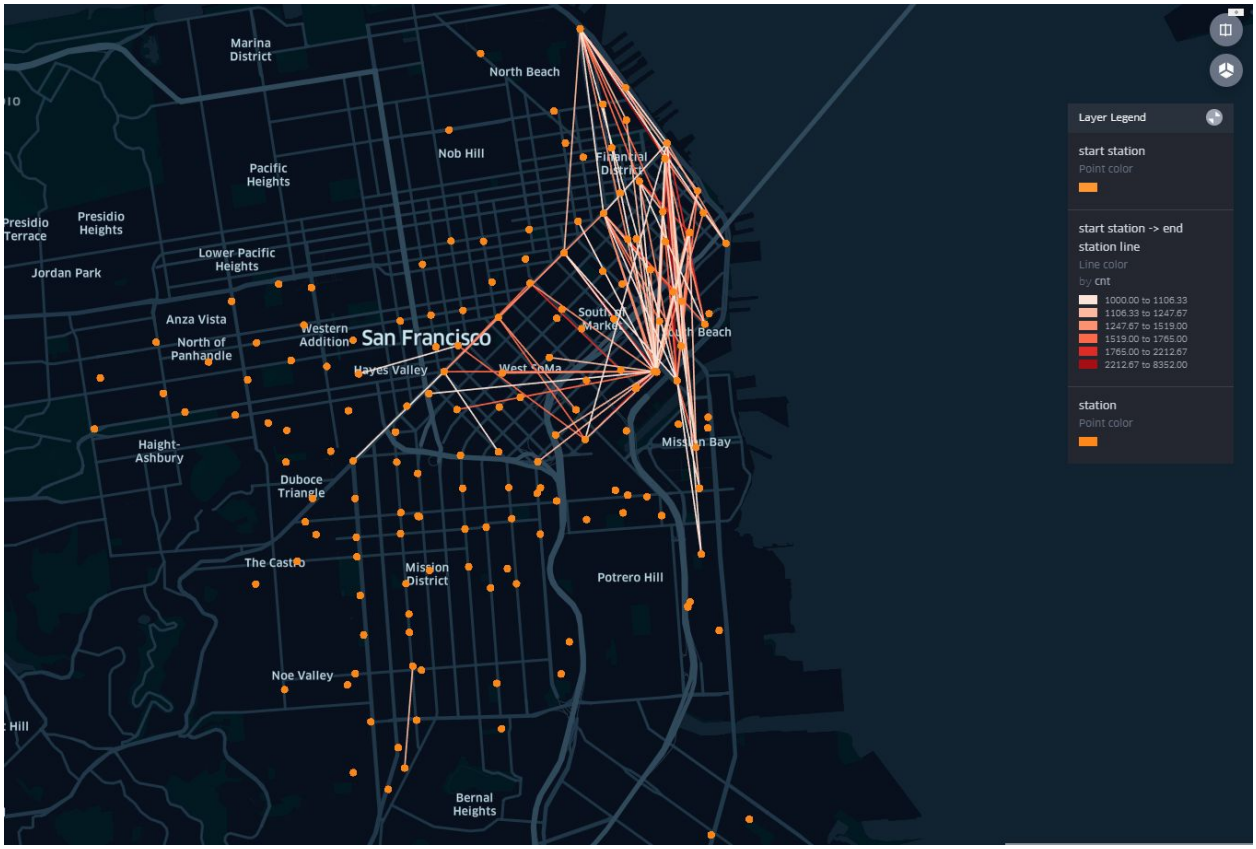
This trend applies for all areas, while we can also see that the users of San Francisco have, on average, the longest duration of trips, followed by East Bay and then San José. And what is the average starting hour per day?



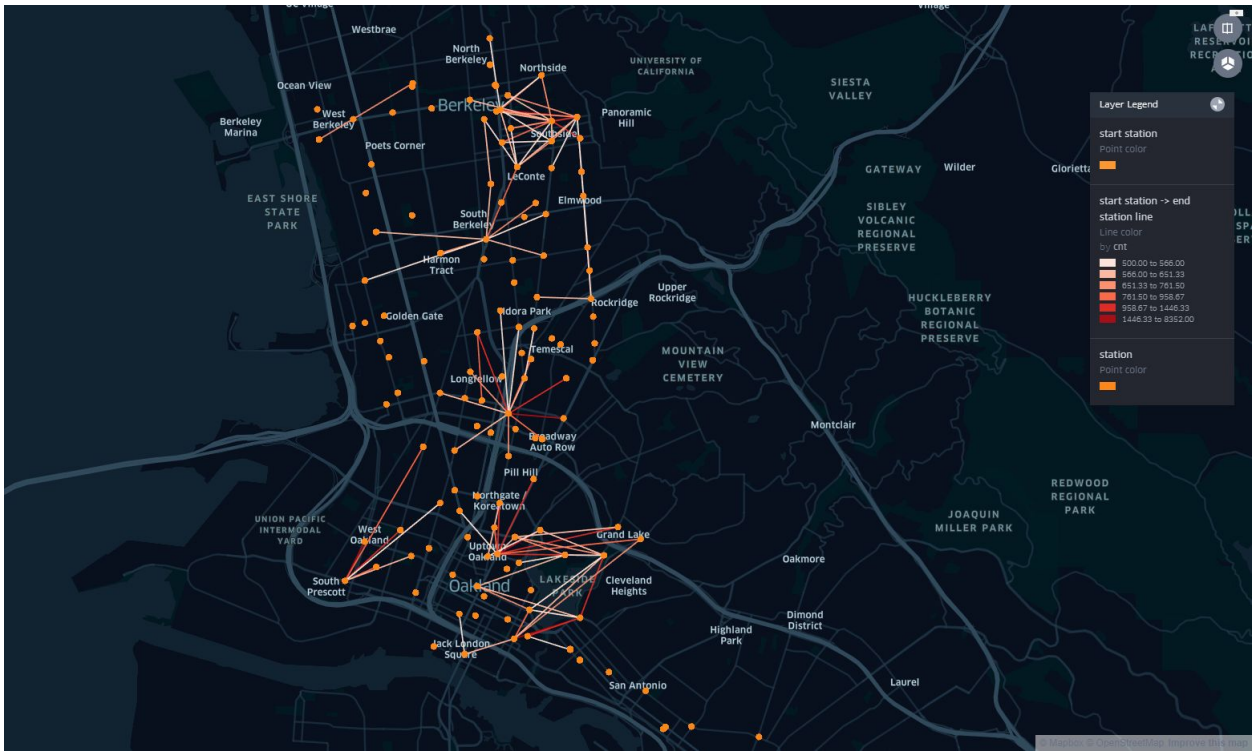
Looking at each area is interesting because Users from East Bay and San José are not only had shorter trip durations on average but also they start their trips later than San Francisco on average.

For the final visualizations, let's visualize the trips.

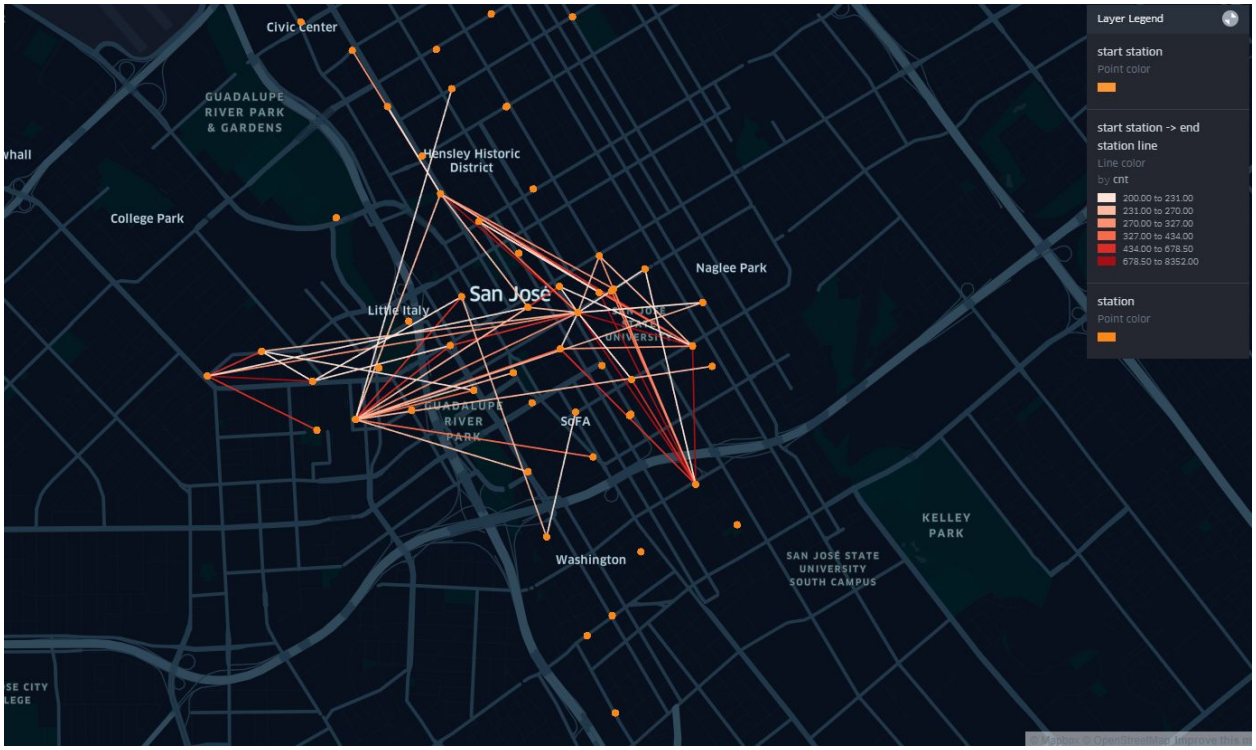
At first, we will look at San Francisco. To get some insight, the visualization will only contain routes with more than 1000 trips:



We can see that most of the trips are close to the beach. Now for East Bay with routes with more than 500 trips:



Here the main routes are much more spread than in San Francisco. Also, it looks like people use this service to quickly overcome smaller distances. For San José, we will take a look at routes that have more than 200 trips.



For San José, it looks spread over most of the stations.