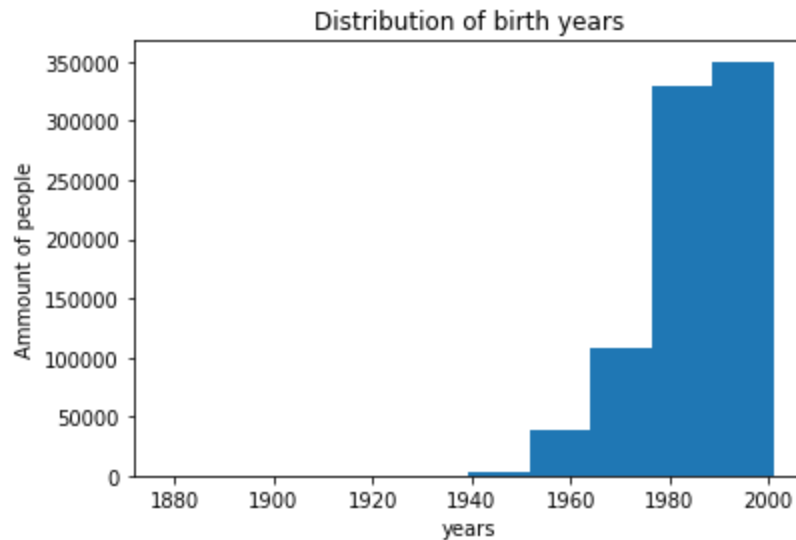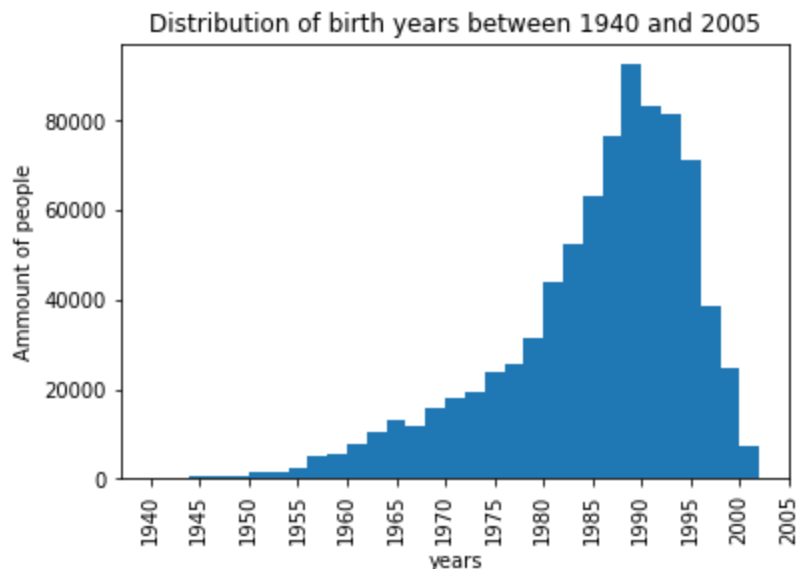# Exploratory Analysis Report

The dataset that I chose was from Ford GoBike.

The first question that I asked during my exploratory analysis was "what is the distribution of the member_birth_year". This question is extremely important since you can target your advertisement to a specific group of people in a certain range of age according to your results. With that in mind, I plotted a histogram of the member_birth_year and this is the result that I got:
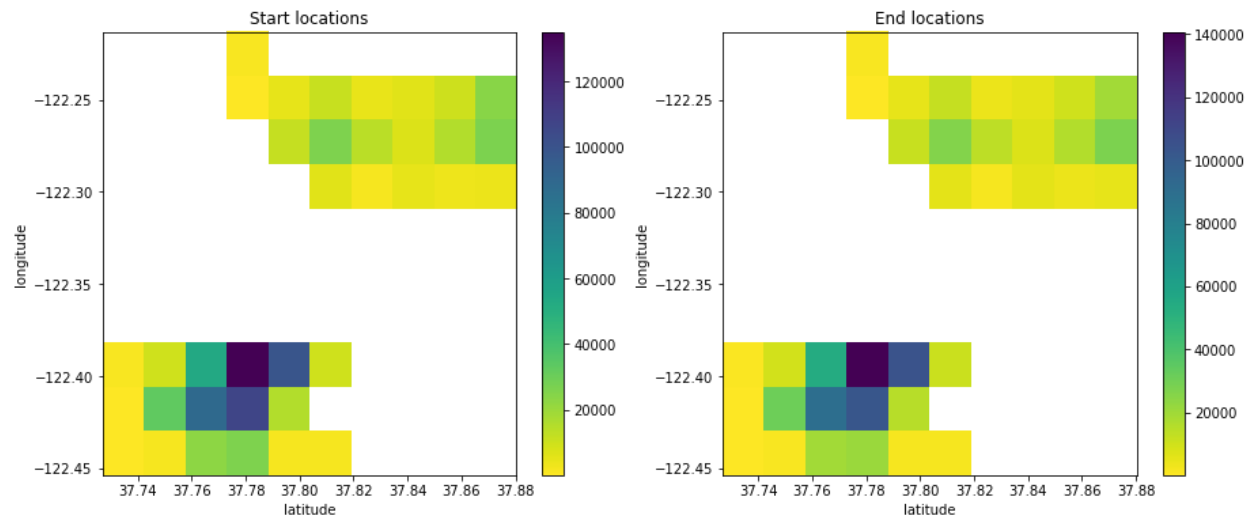

As we can see, there is a long tail up to 1880, but even though that is the case, we should focus in the right part, since more than 99% of the data is concentrated in the right part of the graph (to be more precise, around 99.9% was borned after 1940). With that in mind, I polished the graph to focus on that part.

As we can see, the result is that the graph is left skewed with maximum around 1987. This information can be used, for example, to maximize your profit by targeting certain group of people (around that age).

The next question that I asked is, "what is the distribution of the start and end point according to their latitude and longitude?". This information can be useful in order to understand where most people tend to use a bike and, with that information, start to regulate the amount of bikes that are available in those places.

As we can see, the distribution of the graphs are almost identical, with a big concentration around (37.78,-122.40).