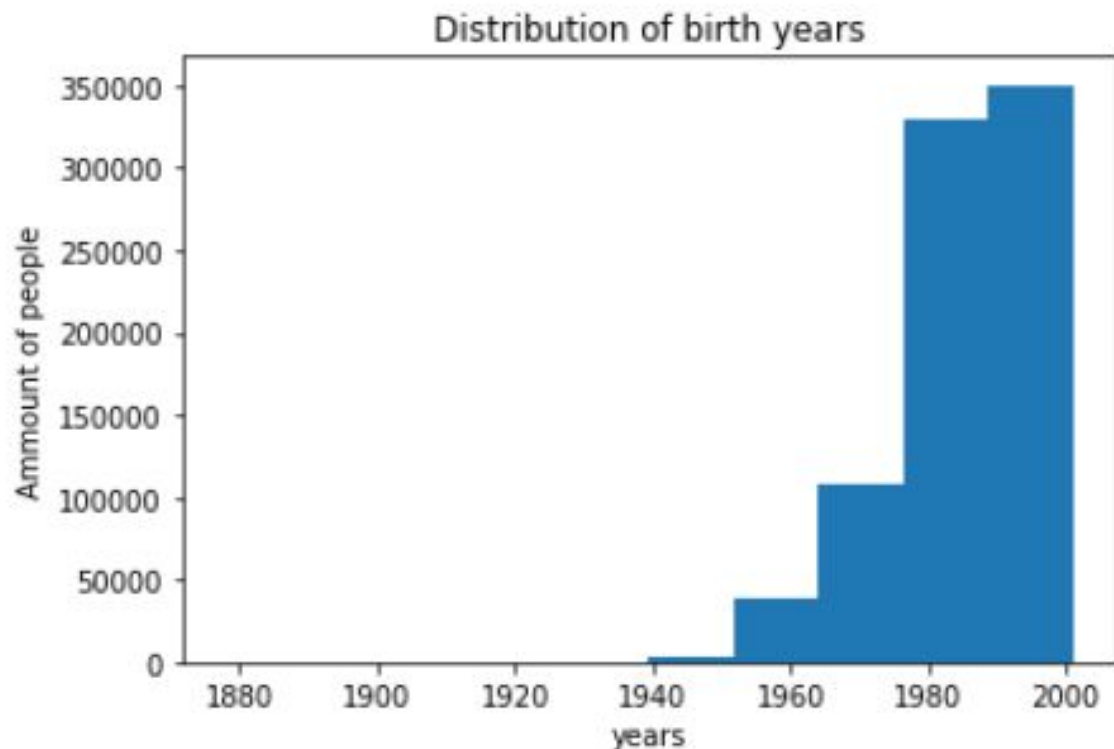


Ford - GoBike

The dataset that I chose was from Ford GoBike. I tried to think questions that are related to how to make more profit out of the data that we have already collected, that is, questions like how to target your main audience with advertisement or in which place we should offer our services (since there is more demand in them)

Distribution of born
years of clients

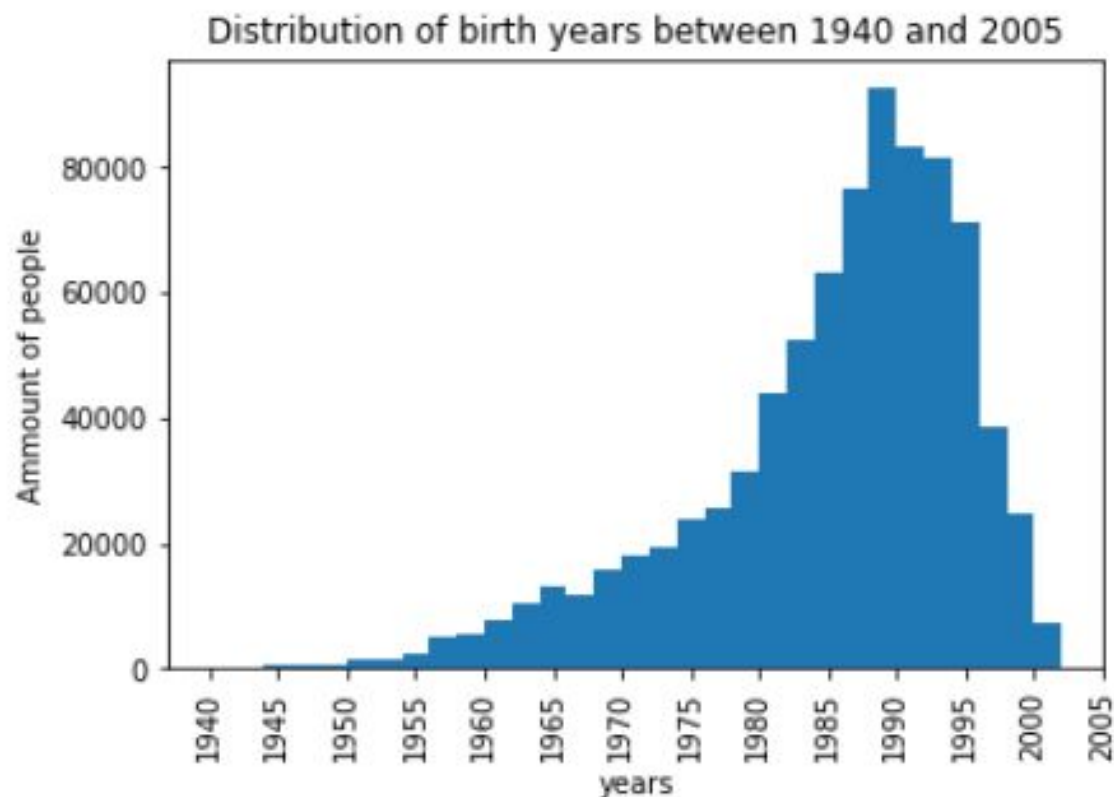
```
x_ = df_.member_birth_year.values  
plt.hist(x = x_);  
plt.title('Distribution of birth years')  
plt.xlabel('years')  
plt.ylabel('Ammount of people');
```



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).

Distribution of born
years of clients around 1940
And 2005

```
gr(2,5)  
plt.title('Distribution of birth years between 1940 and 2005')  
plt.xlabel('years')  
plt.ylabel('Ammount of people');
```



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).

Distribution of start and end
point location

```
plt.figure(figsize= [15,6])

plt.subplot(1,2,1)

lat = df_aux.start_station_latitude.values
long = df_aux.start_station_longitude.values

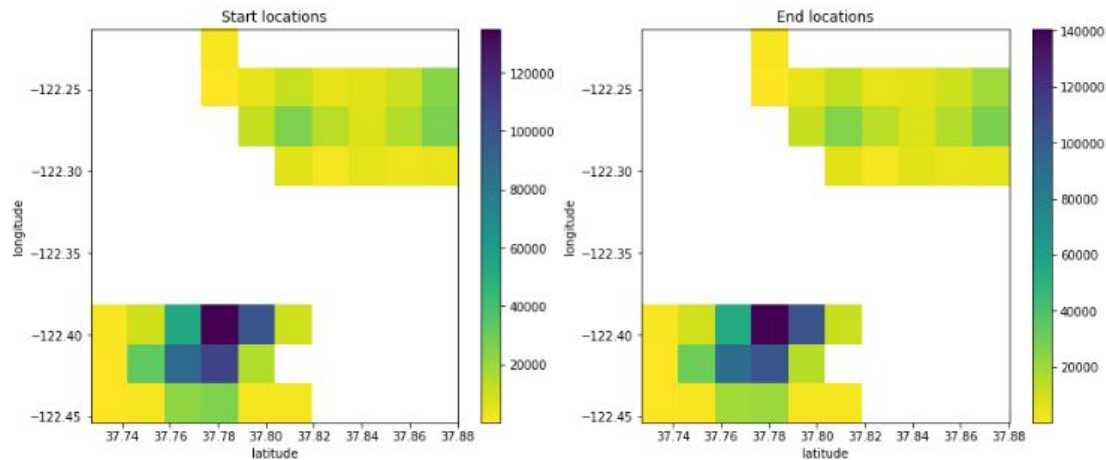
plt.hist2d(x=lat,y = long, cmap = 'viridis_r', cmin = 0.5);
plt.colorbar();
plt.title('Start locations')
plt.xlabel('latitude')
plt.ylabel('longitude')

plt.subplot(1,2,2)

plt.title('End locations')
plt.xlabel('latitude')
plt.ylabel('longitude')

end_lat = df_aux.end_station_latitude
end_long = df_aux.end_station_longitude

plt.hist2d(x=end_lat,y = end_long, cmap = 'viridis_r', cmin = 0.5);
plt.colorbar();
```



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