

Assignment 4: Becoming an Independent Data Scientist

I believe that many people who are taking this online certificate want to build a career in Data Science. So I decided to study some important information related to that in the United States: I chose six cities in order to know more about the number of data science job openings there and their average annual income, as well as the city annual max and min temperature and the values of housing offers.

1 Region and Domain

- Austin, TX
- Boston, MA
- Detroit, MI (instead of Ann Arbor, as I couldn't find housing data for Ann Arbor)
- New York, NY
- San Francisco, CA
- Seattle, WA
- Economic activity or measures

2 Research Question

Which city would be a good choice to live in if you want to work as a Data Scientist? I decided to compare the six cities above in terms of job offers, salaries, annual temperatures, and house prices (for renting and for sale).

I know it's a very tough question that is related to many different types of features, but for the purpose of this assignment, I decided to simplify the problem only looking at the data I chose.

3 Links

3.1 Ziprecruiter data regarding the number of job openings and annual income (web scraping)

<https://www.ziprecruiter.com/Jobs/Data-Scientist/-in-San-Francisco,CA>
<https://www.ziprecruiter.com/Jobs/Data-Scientist/-in-Boston,MA>
<https://www.ziprecruiter.com/Jobs/Data-Scientist/-in-Detroit,MI>
<https://www.ziprecruiter.com/Jobs/Data-Scientist/-in-New-York,NY>
<https://www.ziprecruiter.com/Jobs/Data-Scientist/-in-San-Francisco,CA>
<https://www.ziprecruiter.com/Jobs/Data-Scientist/-in-Seattle,WA>

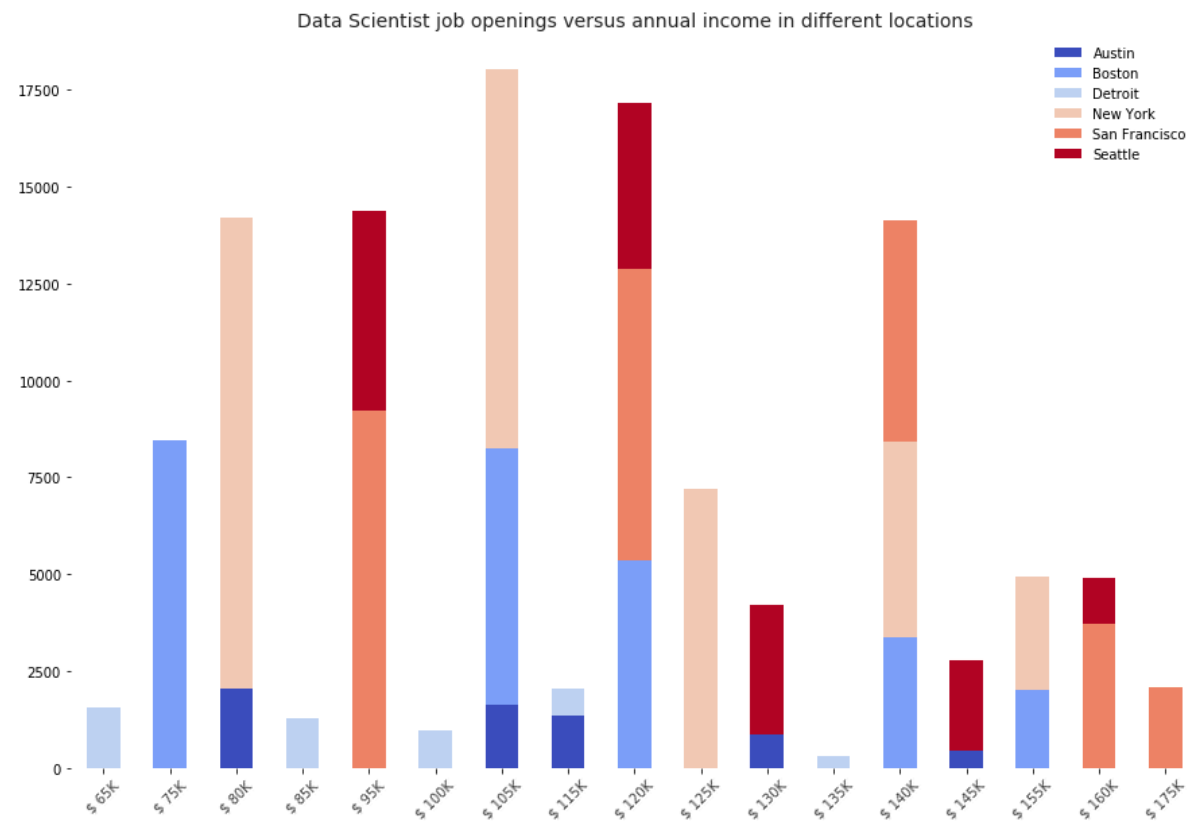
3.2 Wikipedia data regarding max and min annual temperatures in each city (web scraping)

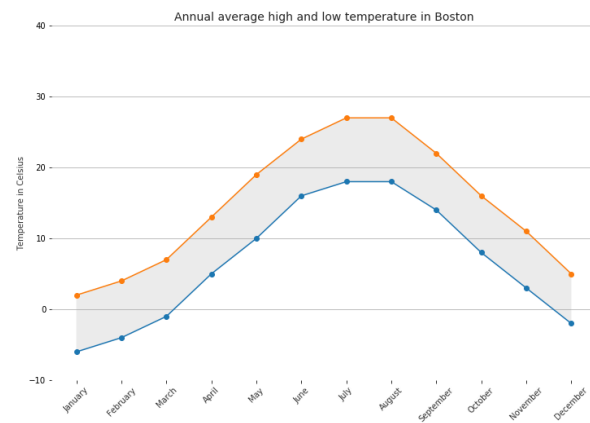
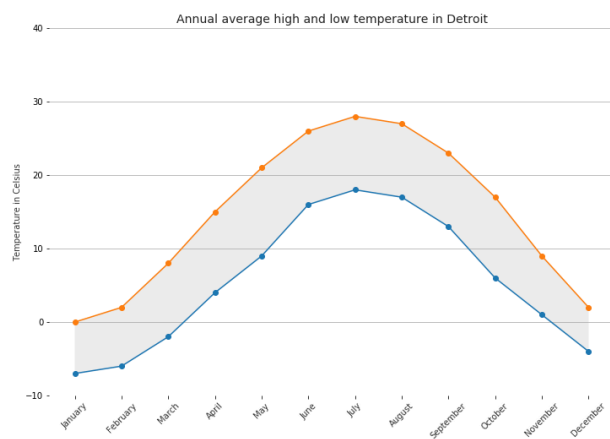
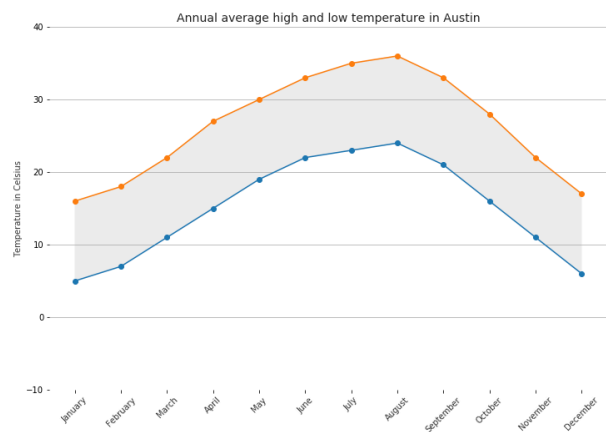
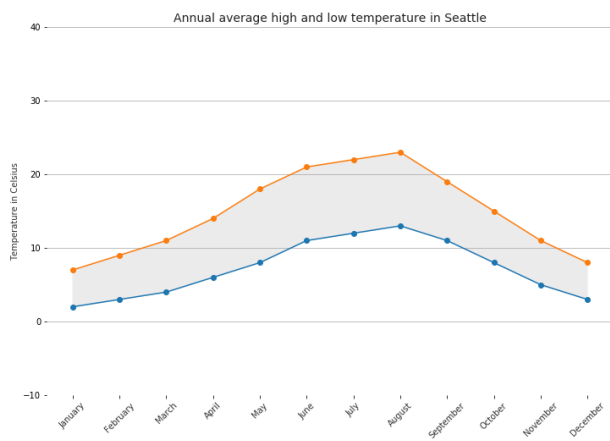
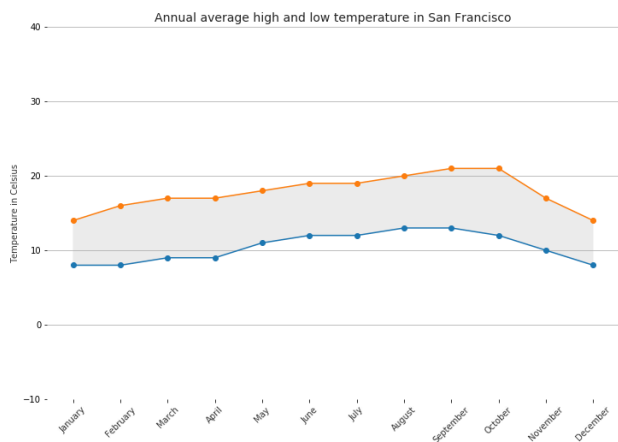
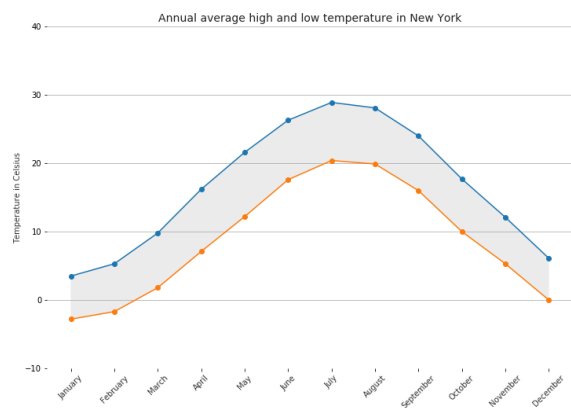
https://en.wikipedia.org/wiki/San_Francisco
<https://en.wikipedia.org/wiki/Boston>
<https://en.wikipedia.org/wiki/Detroit>
<https://en.wikipedia.org/wiki/Austin>
<https://en.wikipedia.org/wiki/Seattle>

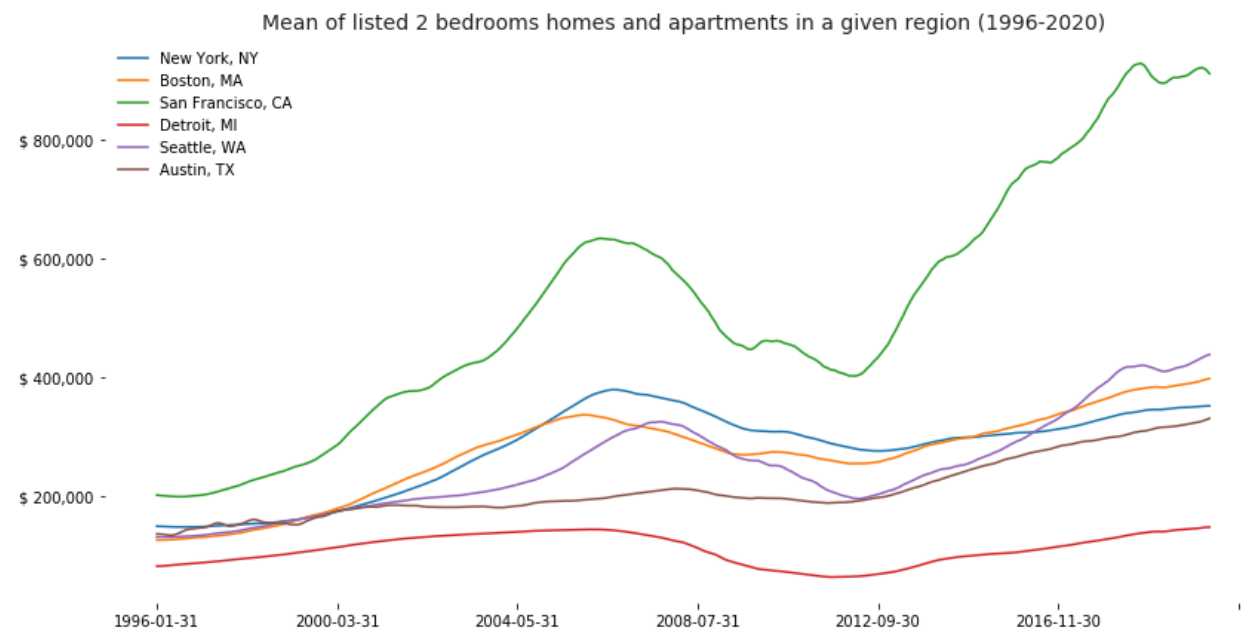
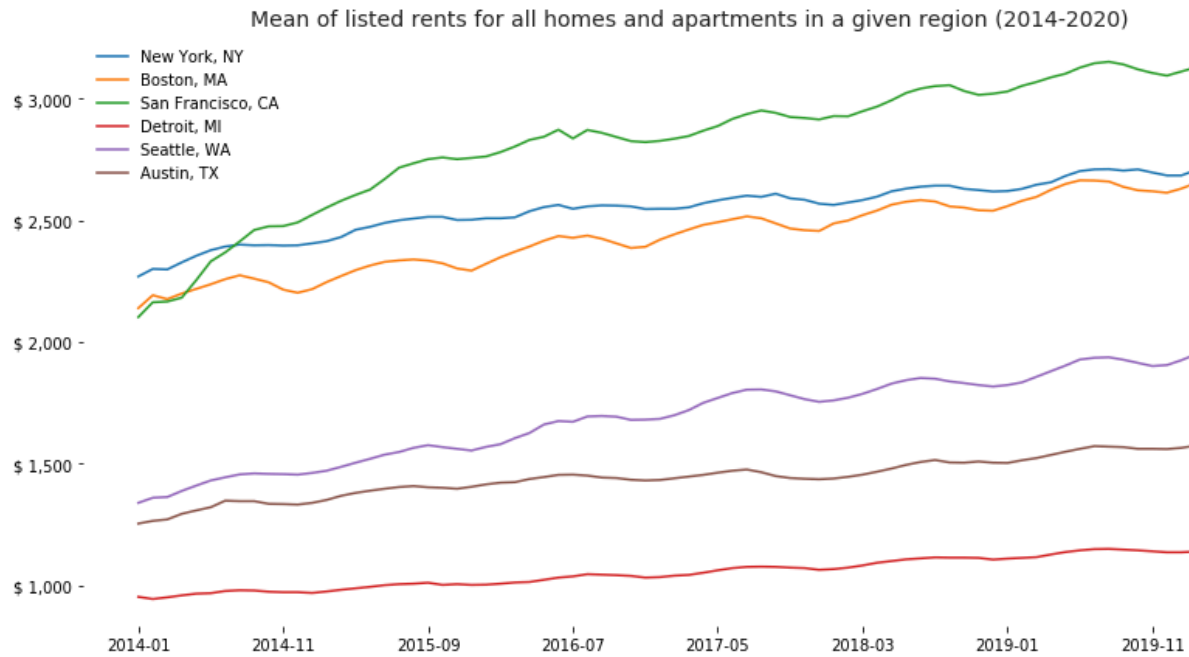
3.3 Zillow data regarding rentals and selling houses and apartments (CSV files)

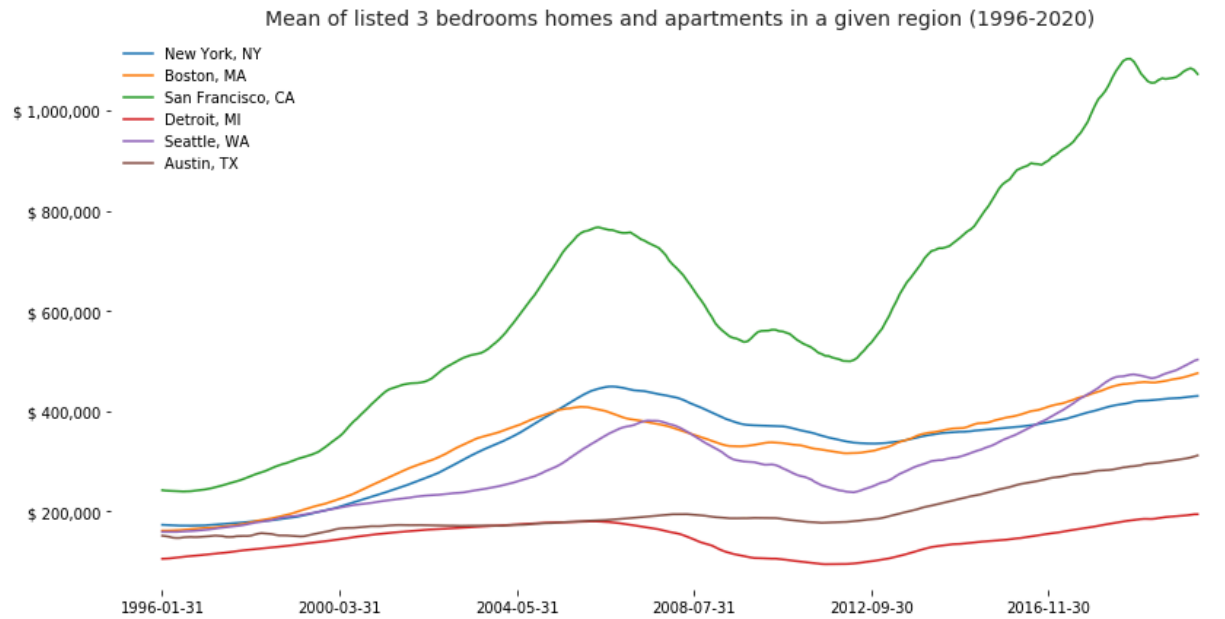
<https://www.zillow.com/research/data>

4 Images









5 Discussion

It can be seen in the first image that New York, Seattle, San Francisco, and Boston are the cities which offer the greatest number of job openings as a Data Scientist. New York and San Francisco are the cities that concentrate the highest annual incomes. Therefore, a data scientist job hunter would be willing to choose one of both cities.

Looking at the graphics about annual temperatures, it is visible that San Francisco is the city with the smallest temperature variation throughout the year. New York, Boston, and Detroit, on the other hand, are the cities in which the temperature variation is the highest. So, if a person doesn't really like extreme weather, like too cold or too warm, would prefer to pick San Francisco to start a career.

Finally, from the last three graphics, about housing prices (for renting and for sale), it is really clear that San Francisco has by far the highest prices. So, although it has the highest average annual incomes and great weather throughout the year, it also has the highest cost in terms of housing expenses. As I mentioned above, the question is not an easy one, many points should be evaluated and I just wanted to show some information in a way that it would be easier to help someone looking for a new city to establish as a data scientist.