The Battle of Neighborhoods

Relocation Decision Assistance

Relocation Decision Assistance

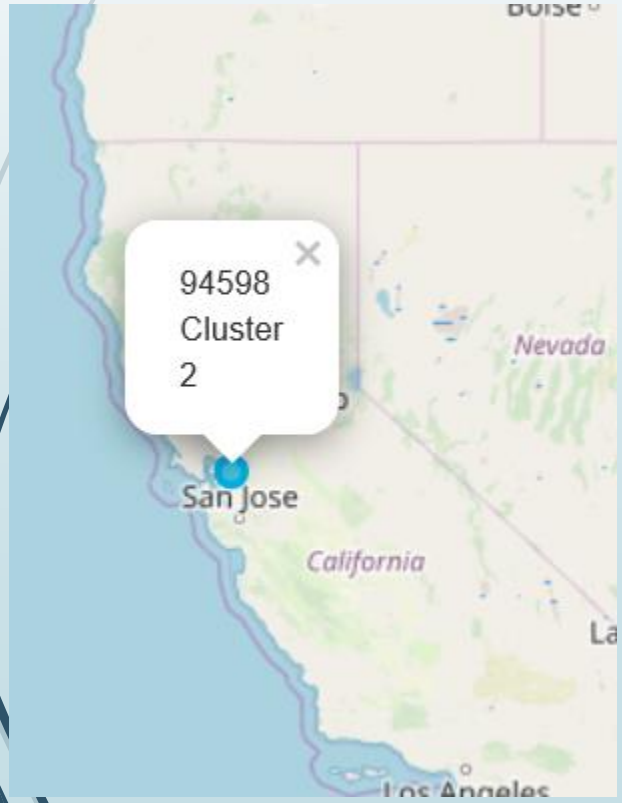
- About 40 million people (or 14% of US population) move every year.
- Questions to answer:
  - Is the housing price in the new location affordable?
  - Will I have same or better amenities in my new neighborhood?
  - What is the quality of public schools?
- Project uses data science to help answer those questions.
- The results will be equally useful for real estate agents as well as their clients.

Data Acquisition & Cleaning

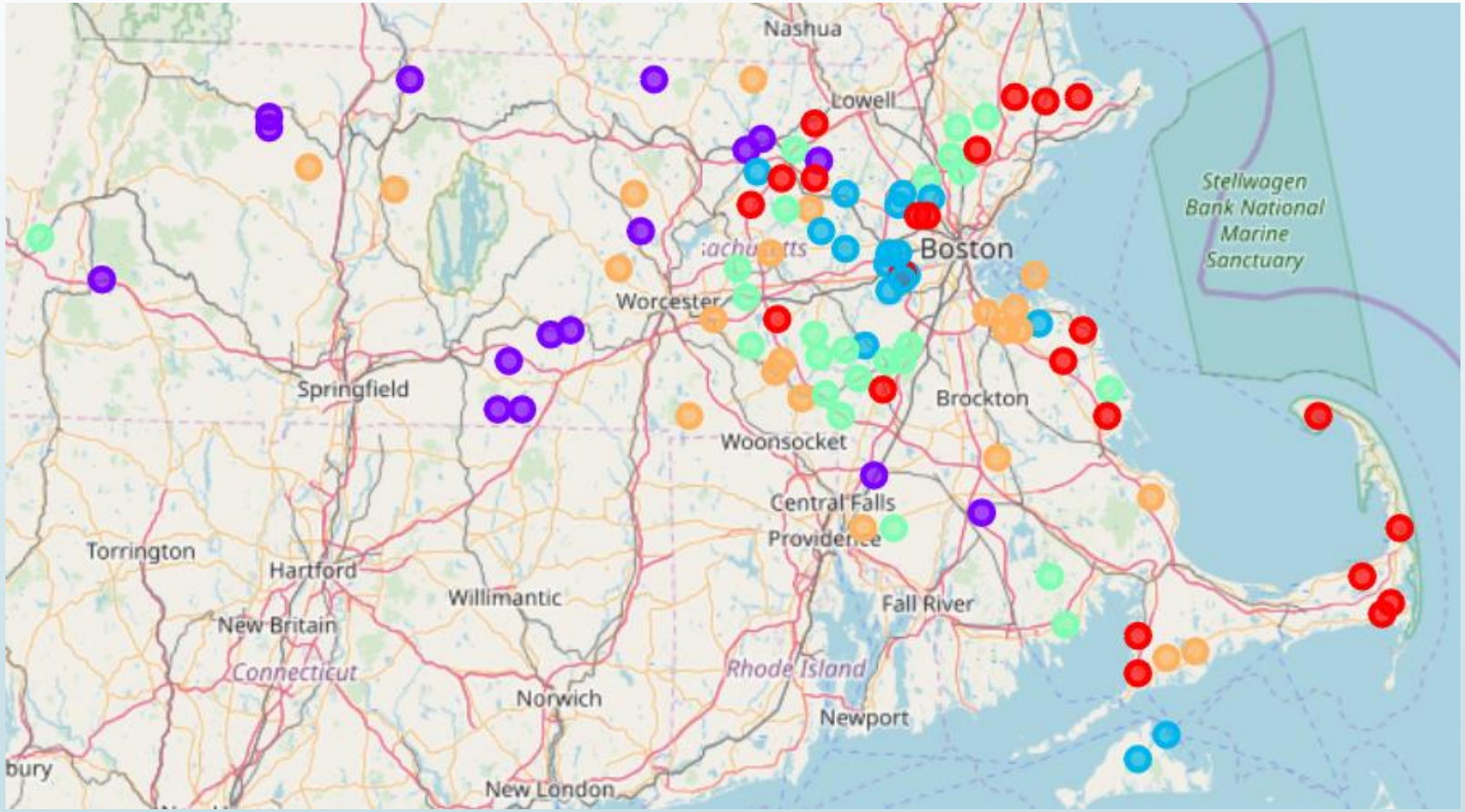
- USZipCode database available on PyPI. The database provides geographical, statistical, real-estate data for every US zip code.
- FourSquare data for information on amenities. FourSquare API is used to collect the data.
- GreatSchools web site for public schools rating. I wrote a function to scrape the web site and to get the average schools rating for the provided zip code.
- Locations with higher median home prices and lower school rating than desired are remove.
- The resulting data set consists of 103 rows and 323 features.

Using K-Means Clustering

Original Location

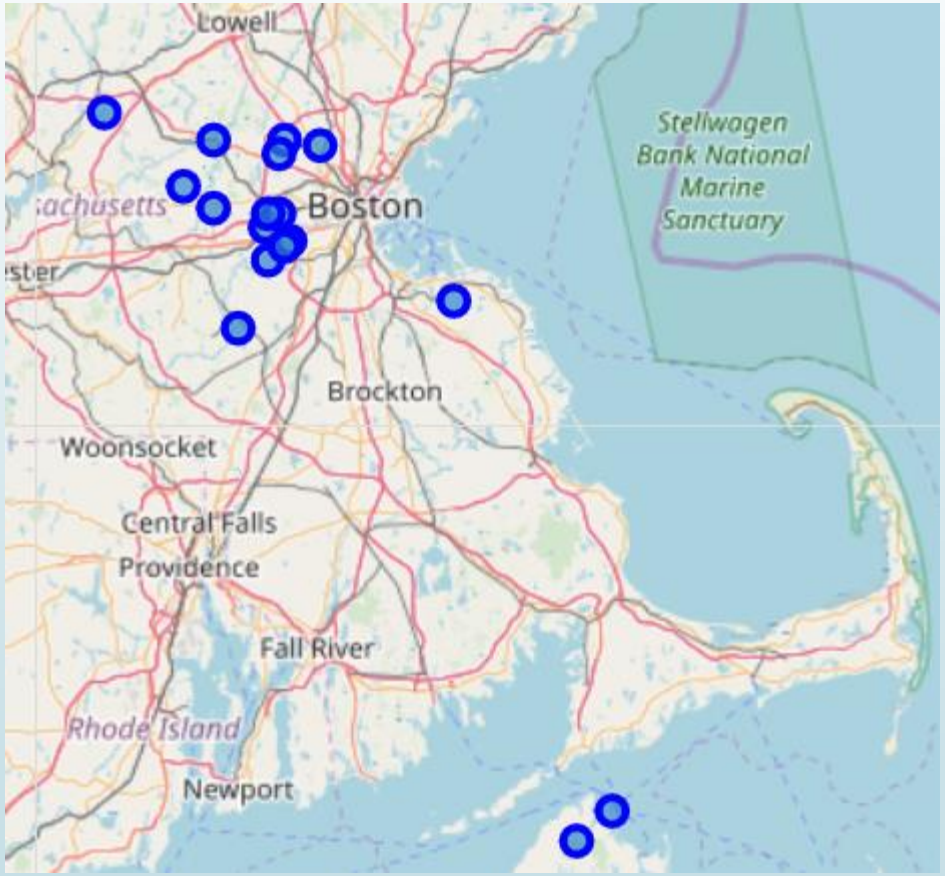


Potential Relocation Area



Results

- Current location ended up in the cluster #2. That cluster contains 17 potential relocation cities that are similar to the current location.



Conclusion

- Out of 17 similar cities,

3 stands up because of higher school rating and lower house price. These three locations might be potential relocation cities for the client.

	post_office_city	median_home_value	e_school
12	Harvard, MA	590900	6
30	Concord, MA	686700	7
38	Sudbury, MA	632800	6
39	Wayland, MA	589700	6
45	Winchester, MA	676800	6
54	Hingham, MA	646700	6
57	Medfield, MA	582800	6
74	Lexington, MA	688500	6
75	Lexington, MA	718300	6
76	Newton Highlands, MA	682600	6
77	Newton Lower Falls, MA	643300	6
79	West Newton, MA	646700	6
80	Auburndale, MA	633600	6
83	Needham, MA	678500	7
84	Needham Heights, MA	631000	6
88	Oak Bluffs, MA	627300	6
89	Vineyard Haven, MA	630600	7