Searing of safe, quiet district at the best possible price in Moscow

Applied Data Science Capstone by IBM/Coursera

Mikhail Lunov

Background of problem

Moscow is a fairly large city and the richest city in Russia:

➡ People from all regions of the country and neighboring countries are coming here

➡ Social and cultural stratification of the population naturally leads to crime

it would be good to find the area that is the least criminal, and also not super expensive for purchasing a home

Data

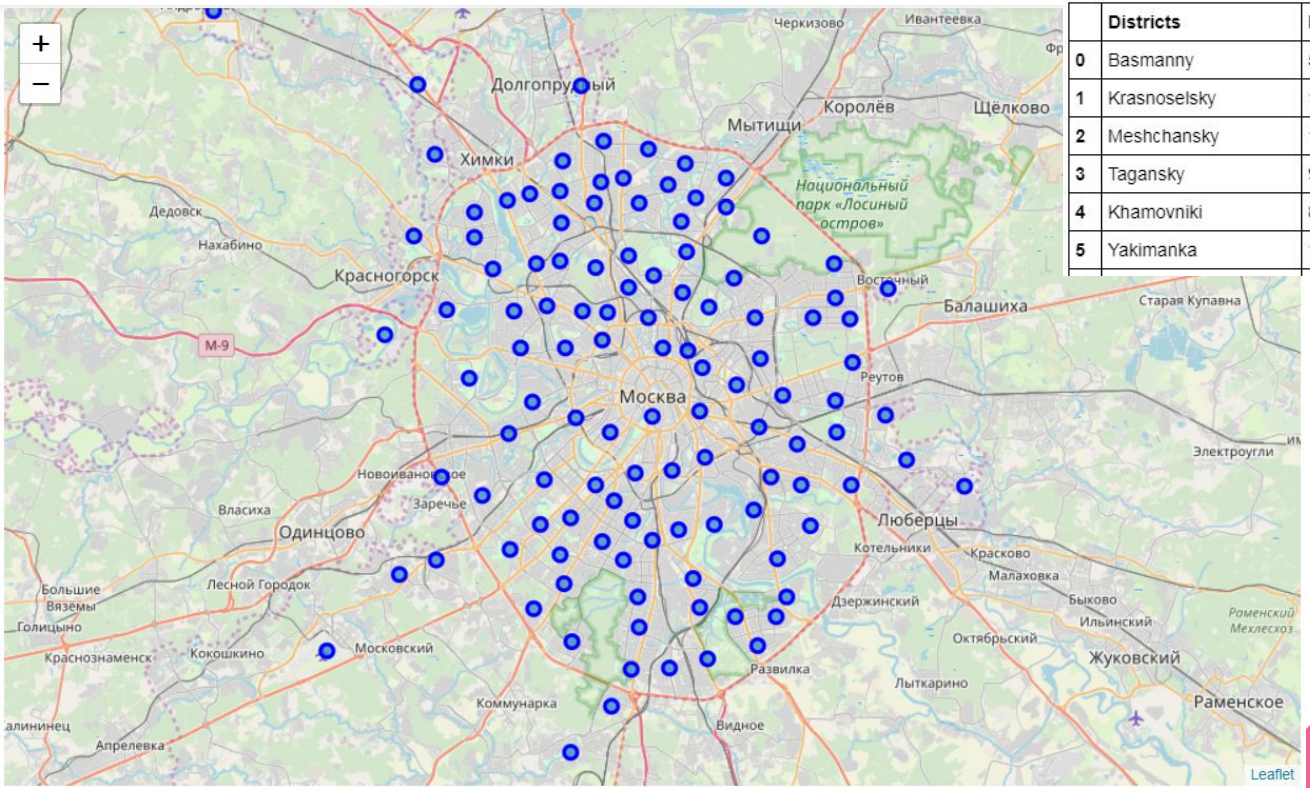
Based on the definition of the problem, factors that will influence decision are:

- Crime rate - number of crimes per the proportion of the population
- House prices - how much one square metre costs
- Distance of district from city center

Following data sources will be needed to extract/generate the required information:

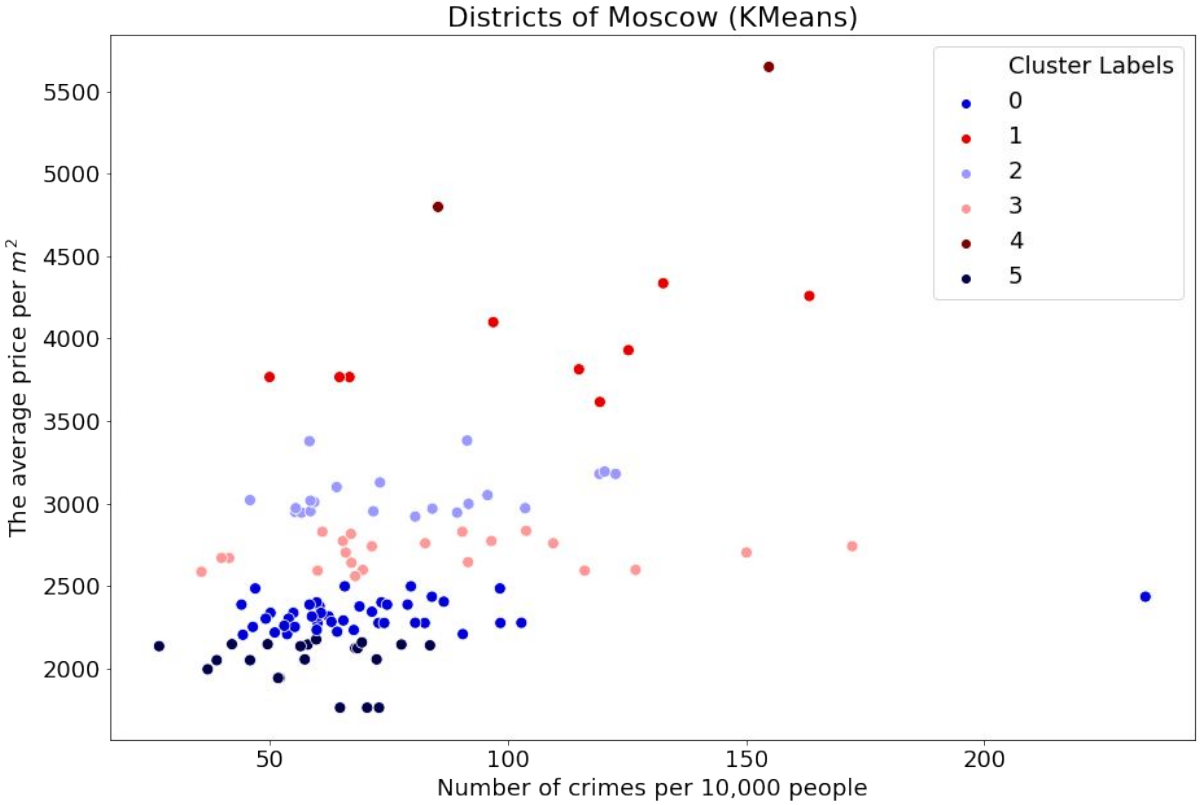
- The cost in USD per square meter of housing in Moscow by district at the time of October 19, 2019
- The number of crimes per 10,000 residents in the districts of Moscow
- The coordinates of Moscow center and its Districts will be obtained using Foursquare API

Using Foursquare API



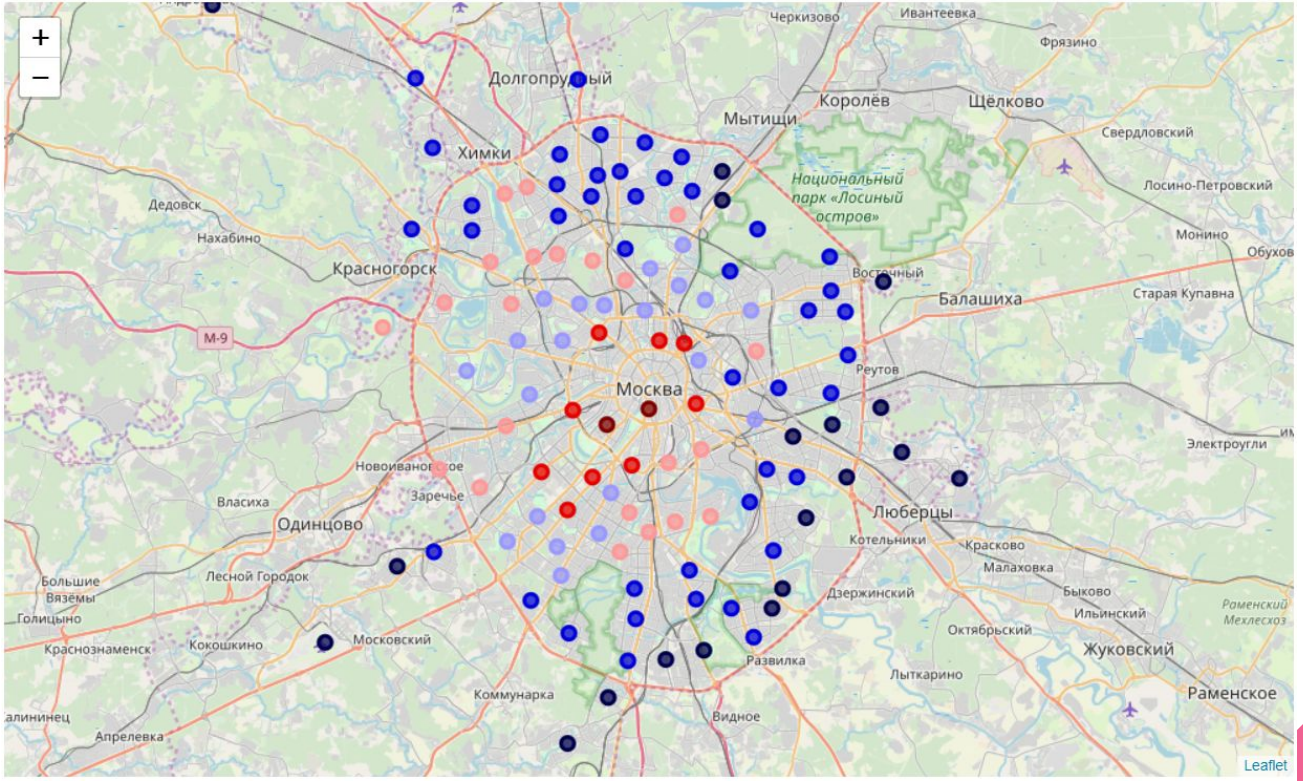
	Districts	Num_of_crimes	USD per sq.m	Latitude	Longitude
0	Basmanny	58.3	3 377	55.767281	37.669773
1	Krasnoselsky	132.6	4 335	55.777447	37.654160
2	Meshchansky	163.3	4 258	55.779172	37.627989
3	Tagansky	96.9	4 099	55.741572	37.667063
4	Khamovniki	85.3	4 798	55.729229	37.572948
5	Yakimanka	154.8	5 647	55.738374	37.616913

k-mean clustering



	Num_of_crimes			USD per sq.m
	amin	amax	mean	mean
Cluster Labels				
5	26.7	83.6	58.185714	2044.190476
0	44.0	233.9	70.211628	2325.744186
2	45.8	122.6	78.804762	3056.904762
3	35.6	172.3	84.163636	2699.681818
1	49.9	163.3	103.722222	3927.444444
4	85.3	154.8	120.050000	5222.500000

Map with k-mean clustering



- **cluster 5** - the calmest districts - **black color**
- **cluster 0** - the second - **dark blue color**
- **cluster 2** - the third - **grey color**
- **cluster 3** - the fourth - **salmon color**
- **cluster 1** - the fifth - **red color**
- **cluster 4** - the most criminal districts - **dark red color**

Conclusion

- The average cost of housing increases in proportion to the number of crimes
- The crime rate is growing when approaching the city center
- The most optimal clusters are 5, 0 and 3 - the districts of these clusters are the safest and cheapest