

Final report of the Battle of the neighborhoods

Business problem

In my final project we want to explore the Moscow and Yerevan city neighborhoods.

We have formed following task: suppose we are representing a consulting agency that helps businesses and individuals to find right location for their business and living.

We had an order from a person who wants to relocate from Moscow (Russian Federation) to Yerevan (Republic of Armenia) and want to find relevant location for buying an apartment. The best location for him will be place that is maximum similar with his Moscow apartment neighborhoods (Marino - district of South-Eastern Administrative region of the federal city of Moscow).

To solve this task we have analyzed the South regions of Moscow city and all administrative districts of Yerevan city.

For Moscow we decided to take the metro stations as control points of our analysis, since the metro is the basis for the city's general transport infrastructure and provides accessibility to all parts of the city.

For Yerevan as a control points we have chosen the centers of administrative regions.

In this capstone project we have gone with these 5 steps:

1. Explore the neighborhoods of south of Moscow and all neighborhoods of Yerevan (Foursquare API)
2. Join this datasets each other,
3. With k-means clustering method divide explored areas to 5 clusters,
4. Find region in Yerevan city that includes in the same cluster with Moscow apartment of our client.

Data

For our analyze we need data from two cities – Moscow and Yerevan.

Moscow data

As we put emphasis on metro station locations and neighborhoods, we need data about all stations of Moscow. This type of information we can get from Wikipedia page (https://en.wikipedia.org/wiki/List_of_Moscow_Metro_stations) with “Beautiful Soup”.

After gathering information we will clean it and filter stations that included in South regions.

Yerevan data

Information about Yerevan`s administrative regions we also can gather from Wikipedia page.

There are 12 administrative districts in Yerevan and for our analysis we will use only 7 out of them. We will use only that districts which is more close to city center.

Methodology

To compare the similarities of two cities, we decided to explore neighborhoods, segment them, and group them into clusters to find similar neighborhoods in a city like Moscow and Yerevan. To be able to do that, we need to cluster data which is a form of unsupervised machine learning.

We will firstly find correct locations of chosen points with Google geocoding API.

Next we will explore neighborhoods in both cities on Foursquare API and get the most common venue categories in each neighborhood.

Next we will use K-means clustering method to divide dataset to clusters and find region in Yerevan that includes in the same cluster with Moscow apartment of our client.

Result

The clusters results of two cities are shown on Figure 1.

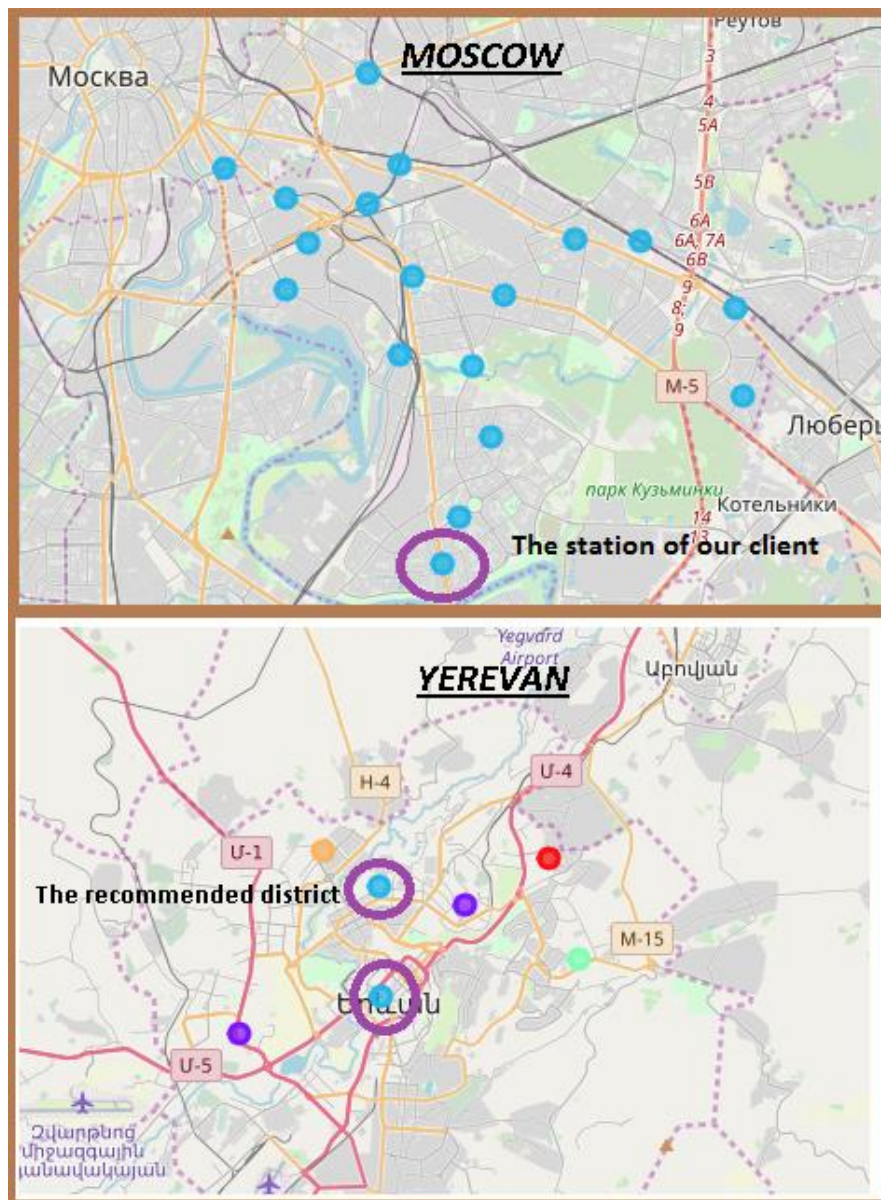


Figure 1

IBM Capstone Project

As we can see all 17 neighborhoods of Moscow city are included in the one cluster (Cluster 2) and from Yerevan city in the same cluster included only two districts (Kentron and Arabkir).

This means that in Yerevan we can recommend at least two districts to our client.

For a more thorough recommendation, we also studied the following factors:

1. The price of the real estate in Yerevan

For chosen neighborhoods in Yerevan real estate price is (for square meter):

- Kentron -1000 \$,
- Arabkir - 750 \$.

2. Transport accessibility

Since our client does not drive a car, transport accessibility is very important for him. we evaluated the selected districts from the standpoint of transport accessibility and found that there are metro stations in both districts and a developed transport network.

It should be noted that in the area of Kentron more loaded transport traffic.

Conclusion

In this project, through a k-means cluster algorithm we separate the neighborhood into 5 clusters, which have similar neighborhoods around them. We discover that we can recommend our client at least 2 district in Yerevan city to relocate.

This two districts have similar neighborhoods with our client Moscow apartment neighborhoods (Marino - district of South-Eastern Administrative region of the federal city of Moscow).

Thus, based on our research, we can recommend our client to buy an apartment in the **Arabkir** area. Our choice is based on three factors - neighborhood similarity, real estate price and transport accessibility.

Karen Davtyan

27/09/2018