The Battle of the Neighborhoods

Content

Introduction & Business Problem :

Problem Background:

I want to open a dance studio in Moscow. It is important to understand that Moscow is a large city with a developed infrastructure. This means that the market is highly competitive. As it is highly developed city so cost of doing business is also one of the highest. Thus, any new business venture or expansion needs to be analyzed carefully. The insights derived from analysis will give good understanding of the business environment which help in strategically targeting the market. This will help in reduction of risk. And the Return on Investment will be reasonable.

It is required to choose a place for a new studio.

It is logical to assume that the studio should satisfy the following criteria:

1. Not located next to other studios. (gyms, yoga centers are allowed)

2. Located in a fairly populated area (that is, it will be convenient for people to visit the studio)

3. Free buildings can be taken from the cian website, in the absence of suitable places, just look at the areas

4. The place should be nearby metro (subway station, as subway is a key transport)

Data :

One city will be analyzed in this project : *Moscow*

We will be using the below datasets for analyzing this city:

Data 1. *Borough (Neighborhoods) of Moscow.*

Moscow has a total 125 neighborhoods and its shape in GEOJSON format. In order to explore neighborhoods, we will essentially need a dataset that contains the as well as the the latitude and longitude coordinates of each neighborhood.

List of Moscow District and they Boroughs were downloaded from the page Moscow Boroughs

Data 2. Shape of every Borough

Shape of the each Moscow Borough in GEOJSON format was downloaded from the page Moscow Boroughs GEOJSON

Data 3. *Density of every Neighborhood*.

Second data which will be used is the density of population dataset. We will use the data from Wikipedia.

Website: https://ru.wikipedia.org/wiki/%D0%A1%D0%BF%D0%B8%D1%81%D0%BE%D0%BA_%D1%80%D0%B0%

Data 4. Available places for rent.

The data was taken from a popular in Russia web-site cian. It includes data about price, time from this place to nearby metro-station, area of a place, address. Url: [https://www.cian.ru/cat.php?deal_type=rent&engine_version=2&offer_type=offices&office_type%5B0%5D=5&region=1&specialty_types%5B0%5D=7015&specialty_types%5B1%5D=7019&specialty_types%5B2%5D=7114&specialty_types%5B3%5D=7119&wp=1](#)

Data 5 . List of venues placed nearby places for rent.

Places for rent geographical coordinates data will be utilized as input for the Foursquare API, that will be leveraged to provision venues information for each potential place for rent. We will use the Foursquare API to explore neighbor of a place for rent in Moscow.

Data 6. Name for type of Dance venues.

In order not to open a new dance studio nearby existing studios we should filter available places for the absence of dance clubs nearby. From Foursquare I found venues_types dedicated to dance.

Data cleaning

As data for Moscow Boroughs dataset were downloaded from HTTP page it was necessary to perform a data cleaning. Such as:

- remove some unused colums
- replace some Borough_Name as of russian letters "e" and "ё"
- change places of some words in Borough_Name
- clear Borough Name from additional information, such as ', поселение ', ', городской округ '
- replace ' ↗' and '↘' in some columns
- delete extra spaces in numeric columns
- replace ',' to '.' for float columns

Data about available places also was cleaned:

- remove free-text from price column
- filtered on time and way how to get to the subway (it should contain «пешком»). Was extected time for getting to the subway
- address did not suit to Nominatim So it was cleaned
- replace some address as of russian letters "e" and "ё"

Methodology

Map of Moscow places with rental places and cluster of venues and density of population.

On this map showed only places for rent that are not situated nearby dance studios. The bigger diameter of a circle the less time you need to get to a metro station

```
35
36  #moscow_merged.drop_duplicates(take_last=True)
37  moscow_merged = moscow_merged[~moscow_merged['Cluster Labels'].isna()]
38  moscow_merged['Cluster Labels']=moscow_merged['Cluster Labels'].astype(int)
39  moscow_merged['price_per_meter']=moscow_merged['price']/ moscow_merged['area']
40  moscow_merged.head()
41
42  # create map
43  #map_clusters = folium.Map(location=[latitude, longitude], zoom_start=11)
44
45  # set color scheme for the clusters
46  x = np.arange(kclusters)
47  print(x)
48  ys = [i + x + (i*x)**2 for i in range(kclusters)]
49  colors_array = cm.rainbow(np.linspace(0, 1, len(ys)))
50  print(colors_array)
51  #colors_array=colors_array.astype(int)
52  print(colors_array)
53
54  rainbow = [colors.rgb2hex(i) for i in colors_array]
55  print(rainbow)
56  # add markers to the map
57  markers_colors = []
58  for adr, price, lat, lon, poi, cluster, time_to_metro_for_circle in zip(moscow_merged['address'], moscow_merged['price'], mo
59      #label = folium.Popup(str(poi) + ' Cluster ' + str(cluster), parse_html=True)
60      print(adr)
61      label = folium.Popup(" Address: " + translit(adr) +   " Price: "+ (str(price))+   ' Cluster ' + str(cluster), parse_html
62      print(" Address: " + adr+   " Price: "+ (str(price))+   ' Cluster ' + str(cluster))
63      #popup_text = "{}<br> ALAND: {:,}<br> AWATER: {:,}"
64
65      folium.CircleMarker(
66          [lat, lon],
67          radius=time_to_metro_for_circle,
68          popup=label,
69          color=rainbow[int(cluster)-1],
70          fill=True,
71          fill_color=rainbow[cluster-1],
72          fill_opacity=0.7).add_to(map_clusters)
73
74  map_clusters
```
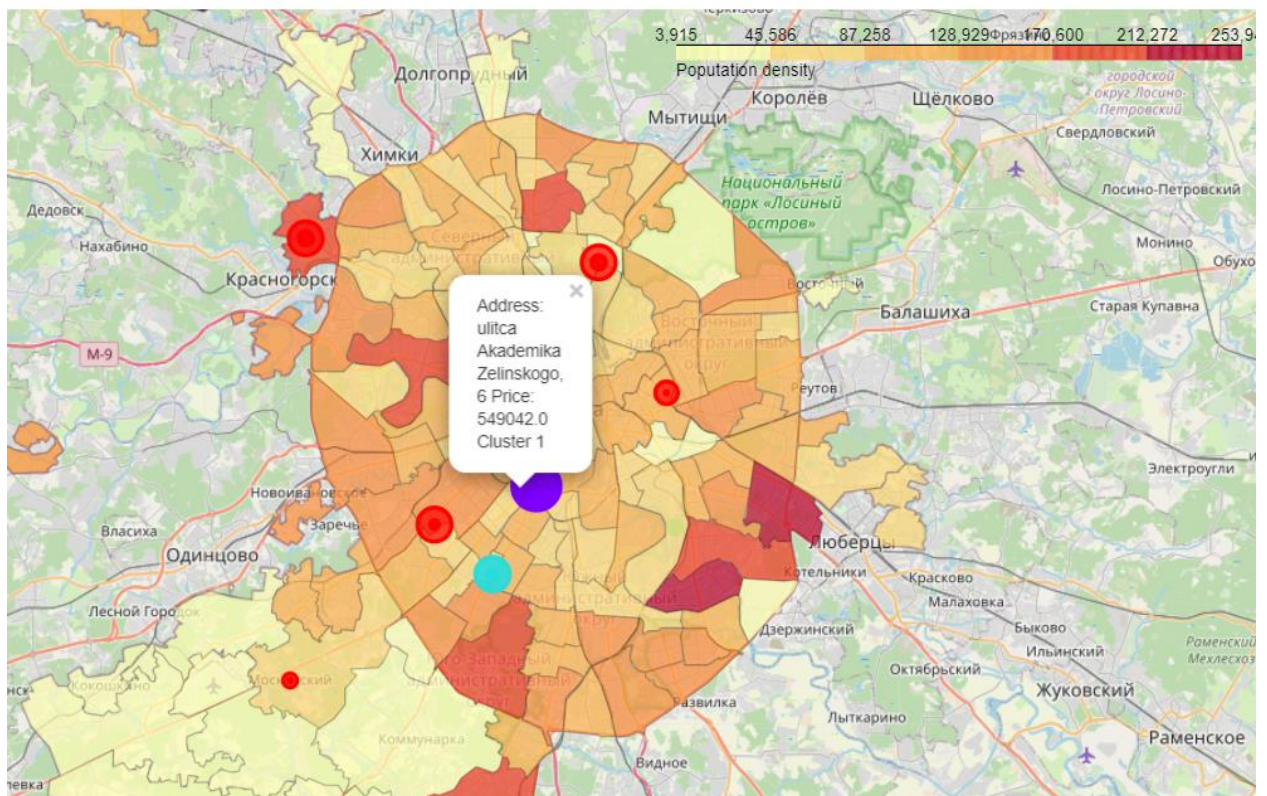

Results

After examining, I have chosen two locations that meet the requirements which will assess to make a choice:

First one place. The first venue has a good description and position.

The price is good and situated nearby metro. The population density is high as it could be seen on map

```
1  moscow_merged.loc[moscow_merged['Cluster Labels'] == 0]
```

6]:

geo_adress	time_to_metro	time_to_metro_un	time_to_metro_for_circle	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	price_per_meter
2-я Магистральная улица, Хорошёвский район, Мо...	8	8	8	0	Coffee Shop	Auto Workshop	Café	Cosmetics Shop	Dance Studio	1200.000000
60, улица Удальцова, Проспект Вернадского, рай...	5	12	12	0	Coffee Shop	Auto Workshop	Park	Gourmet Shop	Sushi Restaurant	800.000000
4 к1, Авиамоторная улица, Квартал Свобода, Анд...	10	8	8	0	Gym / Fitness Center	Park	Café	Cosmetics Shop	Bath House	1587.000000
Московский, поселение Московский, Новомосковск...	13	5	5	0	Clothing Store	Supermarket	Boutique	Sushi Restaurant	Park	2640.000000
2, улица Сергея Эйзенштейна, Ростокино, район ...	2	12	12	0	Park	Clothing Store	Fountain	Gym / Fitness Center	Coffee Shop	334.000000
25 к1 с1, Митинская улица, Митино, район Митин...	2	12	12	0	Gym	Pedestrian Plaza	Supermarket	Sushi Restaurant	Cosmetics Shop	723.684211

Second place. The second venue has also a good description and position.

The price is high, but situated nearby metro. There is no indoor gym nearby.

```
3  #There is no Dance studios around the place. The price affordable
4  moscow_merged.loc[moscow_merged['Cluster Labels'] == 1]
```

t[317]:

Long	geo_adress	time_to_metro	time_to_metro_un	time_to_metro_for_circle	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	price_per_meter
37.586483	Новый Арт Театр, Ленинский проспект, Донской р...	0	17	17	1	Park	Health Food Store	Wine Shop	Furniture / Home Store	Outdoor Gym	1716.586562
37.572352	6, улица Академика Зелинского, Гагаринский, Га...	13	5	5	1	Soccer Field	Park	Furniture / Home Store	Health Food Store	Outdoor Gym	1487.918699
37.594507	5-й Донской проезд, Донской, Донской район, Юж...	10	8	8	1	Cosmetics Shop	Boutique	Park	Gym / Fitness Center	Coffee Shop	1250.000000
37.576231	45 с1, Ленинский проспект, Гагаринский, Гагари...	13	5	5	1	Park	Outdoor Gym	Gourmet Shop	Soccer Field	Coffee Shop	587.503106

```
[271]:  1  moscow_merged.loc[moscow_merged['Cluster Labels'] == 2]
```

Selection

Using the "one map" above, I was able to explore all possibilities since the popups provide the information needed for a good decision.¶

Place 1 has area 152 m*m and rent cost is 110000 Rub per month. Place 1 is located in 2 minutes from metro station. Nearby venues shows that there is gym nearby. According to website of this gym there is no dance classes in this gym.

Place 2 has area 203 m*m and rent cost is 350000 Rub per month. It situated in a less populated borough. Nearby venues shows that there is now popular indoor gym and dance studio.

Based on research, Place 1 is a better choice since the extra monthly rent is worth the conveniences it provides, it has better conditions.

Discussion

In general, I grateful to Coursera IBM Certification Course.

I feel this project presented me a great opportunity to use knowledge

I have created my first project working during and I a bit proud of it, cause it helped me to believe in myself.

Conclusion

This project has shown me a practical application to resolve a real situation that has impacting personal and financial impact using Data Science tools.