

The Battle of Neighborhoods

1. Introduction

Sometimes we have to change our location because of many reasons: work, study, family... This project will help someone to understand how similar two areas from different cities are. As an example, we will try to compare two areas of two cities: Moscow and New York. Similarity of two areas can help to make a decision about migration from one city to another or about business expansion or just can provide some interesting information about two areas from different sides of the globe.

2. Data

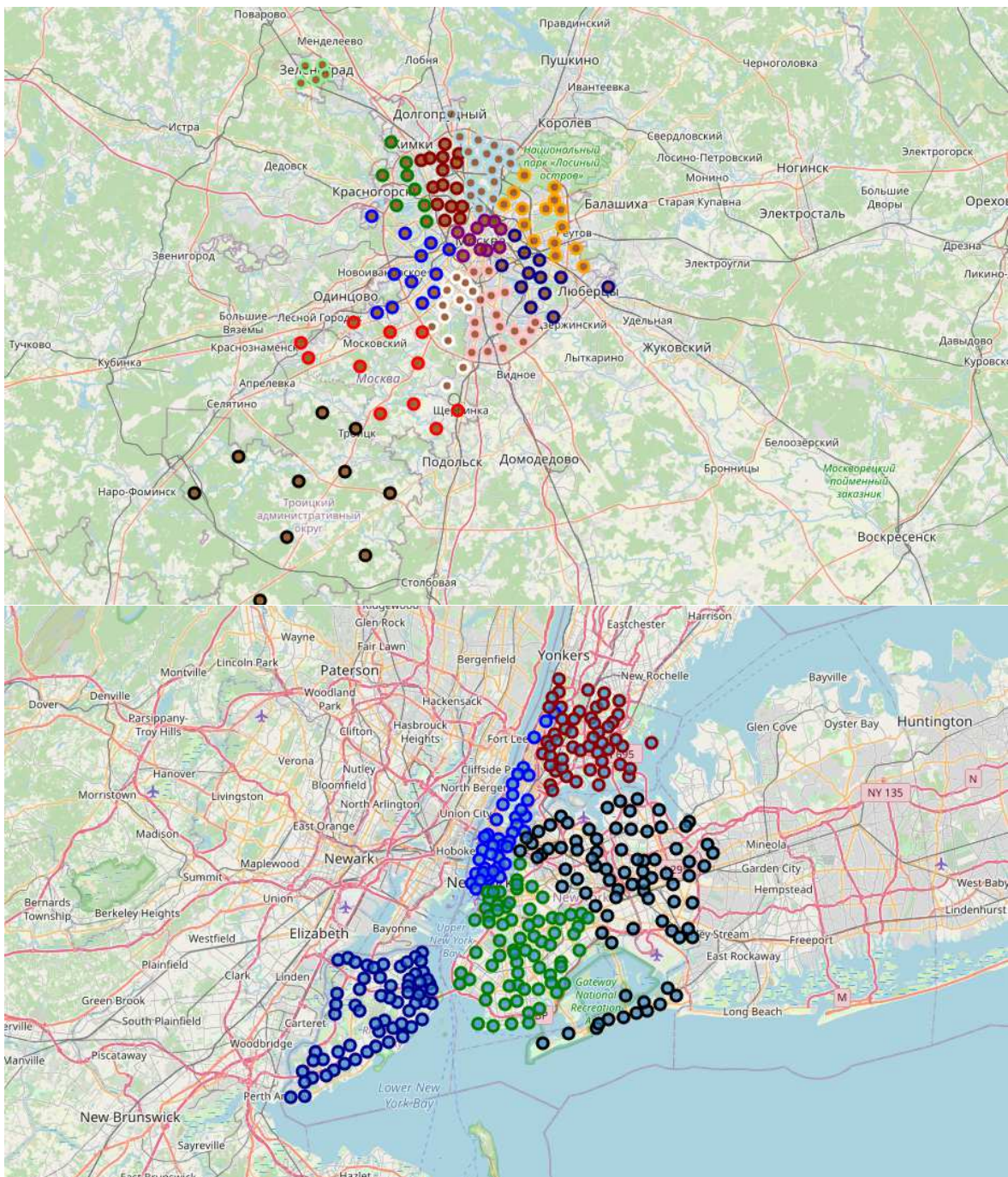
Most of initial information will be scraped from Wikipedia - free and open data source. Most detailed information about every location will be obtained using Foursquare API. Based on definition of our problem following data sources will be needed to extract/generate the required information:

- geographical coordinates of the studied cities
- territory division
- initial information about every area

New York City (NYC) is the most populous city in the United States. With an estimated 2018 population of 8,398,748 distributed over a land area of about 302.6 square miles (784 km²), New York is also the most densely populated major city in the United States. A global power city, New York City has been described as the cultural, financial, and media capital of the world, and exerts a significant impact upon commerce, entertainment, research, technology, education, politics, tourism, art, fashion, and sports. Situated on one of the world's largest natural harbors, New York City consists of five boroughs, each of which is a separate county of the State of New York. The five boroughs – Brooklyn, Queens, Manhattan, The Bronx, and Staten Island – were consolidated into a single city in 1898.

Moscow is the capital and most populous city of Russia, with approximately 12.6 million residents within the city limits. Moscow is the northernmost and coldest megacity on the Earth. Moscow is a major political, economic, cultural, and scientific center of Russia and Eastern Europe. It is the second-most populous city in Europe, the most populous city entirely within Europe, as well as the largest city (by area) on the European continent. Moscow has been ranked as the ninth most expensive city in the world and has one of the world's largest urban economies, being ranked as an alpha global city, and is also one of the fastest growing tourist destinations in the world. Moscow is home to the third-highest number of billionaires of any city in the world, and has the highest number of billionaires of any city in Europe. The city of Moscow is divided into twelve administrative okrugs. By its territorial expansion on July 1, 2012 southwest into the Moscow Oblast, the area of the capital more than doubled, going from 1,091 to 2,511 square kilometers (421 to 970 sq mi), resulting in Moscow becoming the largest city on the European continent by area; it also gained an additional population of 233,000 people.

First, we have to understand territory division of two cities to understand which areas should be explored. Then every area should be explored to define initial criteria of similarity. In the end - detailed exploration of chosen areas. Let's look at two maps of Moscow and NYC with territory division (different colors – different Administrative Okrug (in Moscow) and boroughs (in NYC):



Now let's have a look at numbers, which describes colored areas:

	Borough	Land area km ²	% of all area	Place by area	Population	% of all population	Place by population	Density pers/km ²	Place by density
0	The Bronx	109.04	13.91	4	1418207	17.01	4	13006	3
1	Brooklyn	183.42	23.40	2	2559903	30.71	1	13957	2
2	Manhattan	59.13	7.54	5	1628706	19.54	3	27544	1
3	Queens	281.09	35.86	1	2253858	27.03	2	8018	4
4	Staten Island	151.18	19.29	3	476143	5.71	5	3150	5
5	City of New York	783.83	100.00		8336817	100.00		10636	

	Administrative Okrug	Land area km ²	% of all area	Place by area	Population	% of all population	Place by population	Density pers/km ²	Place by density
0	Central	66.18	22678	11	783886	43252	9	11845.56	5
1	Northern	113.73	18354	7	1188312	13759	7	10448.90	7
2	North-Eastern	101.88	43894	9	1434842	11994	4	14083.23	1
3	Eastern	154.84	41426	3	1527316	43963	2	9864.12	8
4	South-Eastern	117.56	23833	6	1433828	11628	5	12196.59	4
5	Southern	131.77	44682	5	1796267	14.17	1	13631.54	2
6	South-Western	111.36	15067	8	1448130	15646	3	13003.78	3
7	Wester	153.03	43988	4	1397114	43872	6	9129.42	9
8	North-Western	93.28	25263	10	1012949	36342	8	10859.11	6
9	Zelenogradsky	37.20	17168	12	250453	35796	11	6732.63	10
#	Troitsky	1084.34	42.92	1	130812	43891	12	120.64	12
#	Novomoskovsky	361.36	14.30	2	274140	42401	10	758.63	11
#	All Moscow	2526.53	100.00		12678079	100.00		4993.10	

If we will use such features as land area, population and density to find similar areas, then we will stop on two candidates for further study:

Name	Land area km ²	% of all area	Place by area	Population	% of all population	Place by population	Density pers/km ²	Place by density
South-Eastern	117.56	23833	6	1433828	11628	5	12196.59	4
The Bronx	109.04	13.91	4	1418207	43847	4	13006.00	3

And now let's inspect neighborhoods of selected areas:

New York City:

	Borough	Neighborhood	Latitude	Longitude
0	Bronx	Wakefield	40.894705	-73.847201
1	Bronx	Co-op City	40.874294	-73.829939
2	Bronx	Eastchester	40.887556	-73.827806
3	Bronx	Fieldston	40.895437	-73.905643
4	Bronx	Riverdale	40.890834	-73.912585
5	Bronx	Kingsbridge	40.881687	-73.902818
6	Bronx	Woodlawn	40.898273	-73.867315
7	Bronx	Norwood	40.877224	-73.879391
8	Bronx	Williamsbridge	40.881039	-73.857446
9	Bronx	Baychester	40.866858	-73.835798
10	Bronx	Pelham Parkway	40.857413	-73.854756
11	Bronx	City Island	40.847247	-73.786488
12	Bronx	Bedford Park	40.870185	-73.885512
13	Bronx	University Heights	40.855727	-73.910416
14	Bronx	Morris Heights	40.847898	-73.919672
15	Bronx	Fordham	40.860997	-73.896427
16	Bronx	East Tremont	40.842696	-73.887356

17	Bronx	West Farms	40.839475	-73.877745
18	Bronx	High Bridge	40.836623	-73.926102
19	Bronx	Melrose	40.819754	-73.909422
20	Bronx	Mott Haven	40.806239	-73.916100
21	Bronx	Port Morris	40.801664	-73.913221
22	Bronx	Longwood	40.815099	-73.895788
23	Bronx	Hunts Point	40.809730	-73.883315
24	Bronx	Morrisania	40.823592	-73.901506
25	Bronx	Soundview	40.821012	-73.865746
26	Bronx	Clason Point	40.806551	-73.854144
27	Bronx	Throgs Neck	40.815109	-73.816350
28	Bronx	Country Club	40.844246	-73.824099
29	Bronx	Parkchester	40.837938	-73.856003
30	Bronx	Westchester Square	40.840619	-73.842194
31	Bronx	Van Nest	40.843608	-73.866299
32	Bronx	Morris Park	40.847549	-73.850402
33	Bronx	Belmont	40.857277	-73.888452
34	Bronx	Spuyten Duyvil	40.881395	-73.917190
35	Bronx	North Riverdale	40.908543	-73.904531
36	Bronx	Pelham Bay	40.850641	-73.832074
37	Bronx	Schuylerville	40.826580	-73.826203
38	Bronx	Edgewater Park	40.821986	-73.813885
39	Bronx	Castle Hill	40.819014	-73.848027
40	Bronx	Olinville	40.871371	-73.863324
41	Bronx	Pelham Gardens	40.862966	-73.841612
42	Bronx	Concourse	40.834284	-73.915589
43	Bronx	Unionport	40.829774	-73.850535
44	Bronx	Edenwald	40.884561	-73.848083
45	Bronx	Claremont Village	40.831428	-73.901199
46	Bronx	Concourse Village	40.824780	-73.915847
47	Bronx	Mount Eden	40.843826	-73.916556
48	Bronx	Mount Hope	40.848842	-73.908299
49	Bronx	Bronxdale	40.852723	-73.861726
50	Bronx	Allerton	40.865788	-73.859319
51	Bronx	Kingsbridge Heights	40.870392	-73.901523

Moscow:

	Okrug	District	Latitude	Longitude
0	South-Eastern Administrative Okrug	Kuzminki District	55.698318	37.771701
1	South-Eastern Administrative Okrug	Tekstilshchiki District	55.703181	37.740592
2	South-Eastern Administrative Okrug	Lyublino District	55.674310	37.782028
3	South-Eastern Administrative Okrug	Kapotnya District	55.640648	37.804027
4	South-Eastern Administrative Okrug	Maryino District	55.654828	37.747228
5	South-Eastern Administrative Okrug	Pechatniki District	55.684054	37.722773
6	South-Eastern Administrative Okrug	Vykhino-Zhulebino District	55.698397	37.824332

7	South-Eastern Administrative Okrug	Ryazansky District	55.722183	37.768830
8	South-Eastern Administrative Okrug	Nizhegorodsky District	55.732842	37.729088
9	South-Eastern Administrative Okrug	Lefortovo District	55.757399	37.704932
10	South-Eastern Administrative Okrug	Yuzhnoportovy District	55.714477	37.672121
11	South-Eastern Administrative Okrug	Nekrasovka District	55.683056	37.943889

And here we faced with another problem: South-Eastern area in Moscow has 12 districts and Bronx Area from New York has 52 neighborhoods. Comparison of 12 and 52 areas will not be fair. So, let's use postal offices coordinates from these 12 districts. It will let us to increase number of points inside study area. Let's take some data from

<https://data.mos.ru/opendata/1095> :

	Borough	Neighborhood	Latitude	Longitude
0	South-Eastern Administrative Okrug	Lefortovo District_OPS_20	55.766902	37.716897
1	South-Eastern Administrative Okrug	Lefortovo District_OPS_24	55.750886	37.717184
2	South-Eastern Administrative Okrug	Lefortovo District_OPS_33	55.758889	37.687112
3	South-Eastern Administrative Okrug	Nizhegorodsky District_OPS_52	55.730655	37.721188
4	South-Eastern Administrative Okrug	Yuzhnoportovy District_OPS_88	55.716392	37.676188
5	South-Eastern Administrative Okrug	Lefortovo District_OPS_94	55.774971	37.705207
6	South-Eastern Administrative Okrug	Lefortovo District_OPS_116	55.757716	37.713793
7	South-Eastern Administrative Okrug	Kuzminki District_OPS_117	55.711558	37.755294
8	South-Eastern Administrative Okrug	Tekstilshchiki District_OPS_125	55.712107	37.744124
9	South-Eastern Administrative Okrug	Tekstilshchiki District_OPS_129	55.704555	37.743256
10	South-Eastern Administrative Okrug	Maryino District_OPS_144	55.646297	37.737567
11	South-Eastern Administrative Okrug	Vykhino-Zhulebino District_OPS_145	55.699512	37.842421
12	South-Eastern Administrative Okrug	Vykhino-Zhulebino District_OPS_153	55.699180	37.854432
13	South-Eastern Administrative Okrug	Vykhino-Zhulebino District_OPS_156	55.689681	37.854647
14	South-Eastern Administrative Okrug	Nizhegorodsky District_OPS_202	55.736268	37.752902
15	South-Eastern Administrative Okrug	Lefortovo District_OPS_229	55.772342	37.706777
16	South-Eastern Administrative Okrug	Pechatniki District_OPS_235	55.654217	37.701454
17	South-Eastern Administrative Okrug	Lefortovo District_OPS_250	55.758164	37.697742
18	South-Eastern Administrative Okrug	Tekstilshchiki District_OPS_263	55.698670	37.746531
19	South-Eastern Administrative Okrug	Maryino District_OPS_341	55.658006	37.748041
20	South-Eastern Administrative Okrug	Maryino District_OPS_369	55.651824	37.731223
21	South-Eastern Administrative Okrug	Ryazansky District_OPS_377	55.715121	37.789442
22	South-Eastern Administrative Okrug	Kuzminki District_OPS_378	55.707734	37.781858
23	South-Eastern Administrative Okrug	Lyublino District_OPS_380	55.673368	37.810454
24	South-Eastern Administrative Okrug	Lyublino District_OPS_382	55.672172	37.740573
25	South-Eastern Administrative Okrug	Pechatniki District_OPS_383	55.673078	37.721765
26	South-Eastern Administrative Okrug	Lyublino District_OPS_384	55.679066	37.776422
27	South-Eastern Administrative Okrug	Lyublino District_OPS_386	55.679189	37.759394
28	South-Eastern Administrative Okrug	Lyublino District_OPS_387	55.682897	37.738899
29	South-Eastern Administrative Okrug	Pechatniki District_OPS_388	55.679596	37.726240
30	South-Eastern Administrative Okrug	Pechatniki District_OPS_389	55.680131	37.717033
31	South-Eastern Administrative Okrug	Tekstilshchiki District_OPS_390	55.699592	37.736062
32	South-Eastern Administrative Okrug	Nizhegorodsky District_OPS_391	55.731236	37.749364
33	South-Eastern Administrative Okrug	Vykhino-Zhulebino District_OPS_417	55.708493	37.825720
34	South-Eastern Administrative Okrug	Ryazansky District_OPS_428	55.723722	37.780064

35	South-Eastern Administrative Okrug	Kapotnya District_OPS_429	55.634246	37.801711
36	South-Eastern Administrative Okrug	Vykhino-Zhulebino District_OPS_431	55.685988	37.850475
37	South-Eastern Administrative Okrug	Yuzhnoportovy District_OPS_432	55.703189	37.672776
38	South-Eastern Administrative Okrug	Kuzminki District_OPS_439	55.702855	37.781148
39	South-Eastern Administrative Okrug	Kuzminki District_OPS_443	55.704431	37.769198
40	South-Eastern Administrative Okrug	Vykhino-Zhulebino District_OPS_444	55.703970	37.817886
41	South-Eastern Administrative Okrug	Maryino District_OPS_451	55.659505	37.759592
42	South-Eastern Administrative Okrug	Ryazansky District_OPS_456	55.718949	37.786010
43	South-Eastern Administrative Okrug	Kuzminki District_OPS_457	55.711495	37.773155
44	South-Eastern Administrative Okrug	Kuzminki District_OPS_462	55.701718	37.762494
45	South-Eastern Administrative Okrug	Maryino District_OPS_469	55.654143	37.763581
46	South-Eastern Administrative Okrug	Vykhino-Zhulebino District_OPS_472	55.694850	37.803088
47	South-Eastern Administrative Okrug	Vykhino-Zhulebino District_OPS_507	55.694053	37.817304
48	South-Eastern Administrative Okrug	Tekstilshchiki District_OPS_518	55.719273	37.732725
49	South-Eastern Administrative Okrug	Vykhino-Zhulebino District_OPS_542	55.714755	37.808859
50	South-Eastern Administrative Okrug	Pechatniki District_OPS_548	55.693651	37.719664
51	South-Eastern Administrative Okrug	Lyublino District_OPS_559	55.677403	37.763260
52	South-Eastern Administrative Okrug	Kapotnya District_OPS_649	55.630564	37.798656
53	South-Eastern Administrative Okrug	Maryino District_OPS_651	55.647640	37.720248
54	South-Eastern Administrative Okrug	Maryino District_OPS_652	55.651113	37.749052
55	South-Eastern Administrative Okrug	Nekrasovka District_OPS_674	55.683134	37.926576

Now we have 56 points, which is close to number of neighborhoods in New York.

3. Methodology and analysis

We are trying to compare two areas of biggest cities to understand how similar these areas are. Such information will be useful when you are trying to change living location to be sure that familiar things are nearby, or for example to understand is it possible to expand your business there - open a cafe, bar or gym.

By now we have collected some data about cities. We have learned about land area and population of districts inside cities. Based on this knowledge we have chosen two districts similar by land area and population.

Next, we will try to use Foursquare API to get more knowledge about chosen areas - venues around every point.

Similarity of these areas will be evaluated after venue clusterisation after getting data from Foursquare API.

After using Foursquare API for 500 meters around of all of our point we get 243 unique categories. These categories then go to one hot encoding:

	Neighborhood	Accessories Store	African Restaurant	American Restaurant	Women's Store	Yoga Studio
0	Lefortovo District_OPS_20	0	0	0	0	0
1	Lefortovo District_OPS_20	0	0	0	0	0
2	Lefortovo District_OPS_20	0	0	0	0	0
3	Lefortovo District_OPS_20	0	0	0	0	0
4	Lefortovo District_OPS_20	0	0	0	0	0

and group rows by neighborhood by taking the mean of the frequency of occurrence of each category:

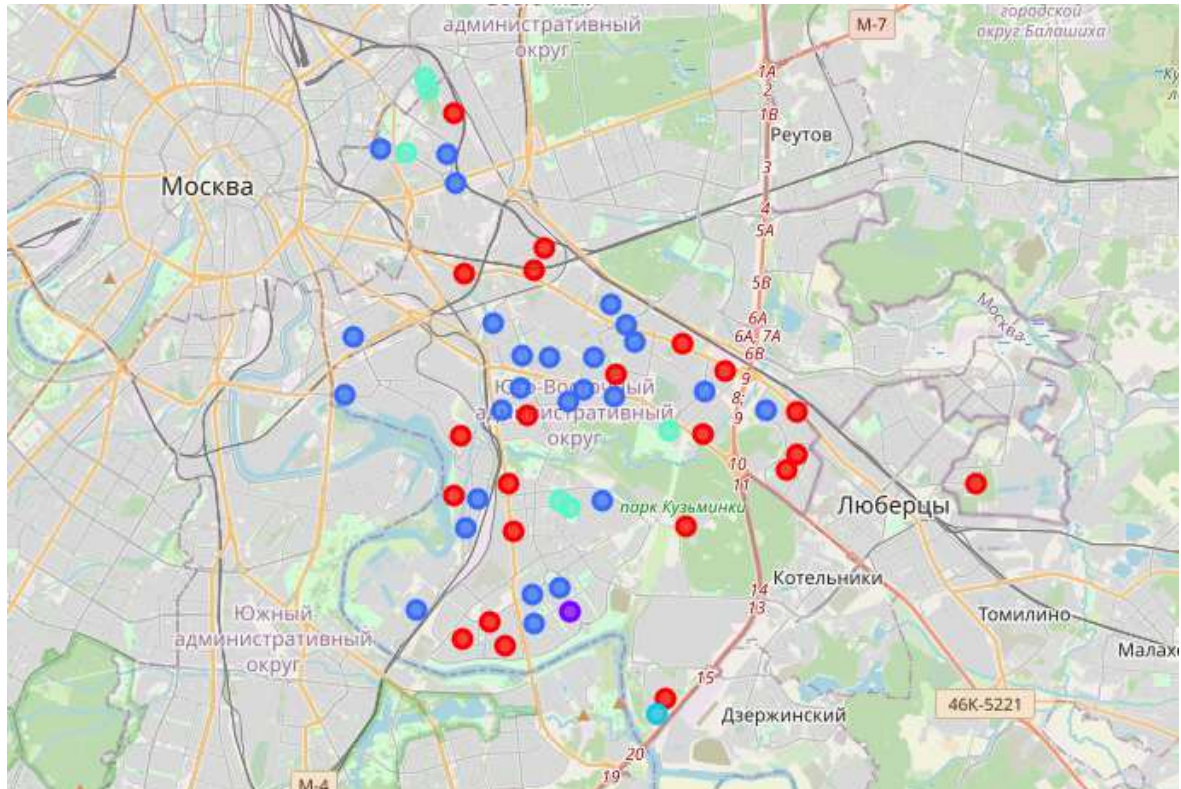
	Neighborhood	American Restaurant	Arcade	Arepa Restaurant	Art Gallery	Women's Store	Yoga Studio
0	Allerton	0.0	0.000000	0.0	0.0	0.0	0.0
1	Baychester	0.0	0.047619	0.0	0.0	0.0	0.0
2	Bedford Park	0.0	0.000000	0.0	0.0	0.0	0.0

After taking each neighborhood along with the top 5 most common venues and putting them together in pandas dataframe we get:

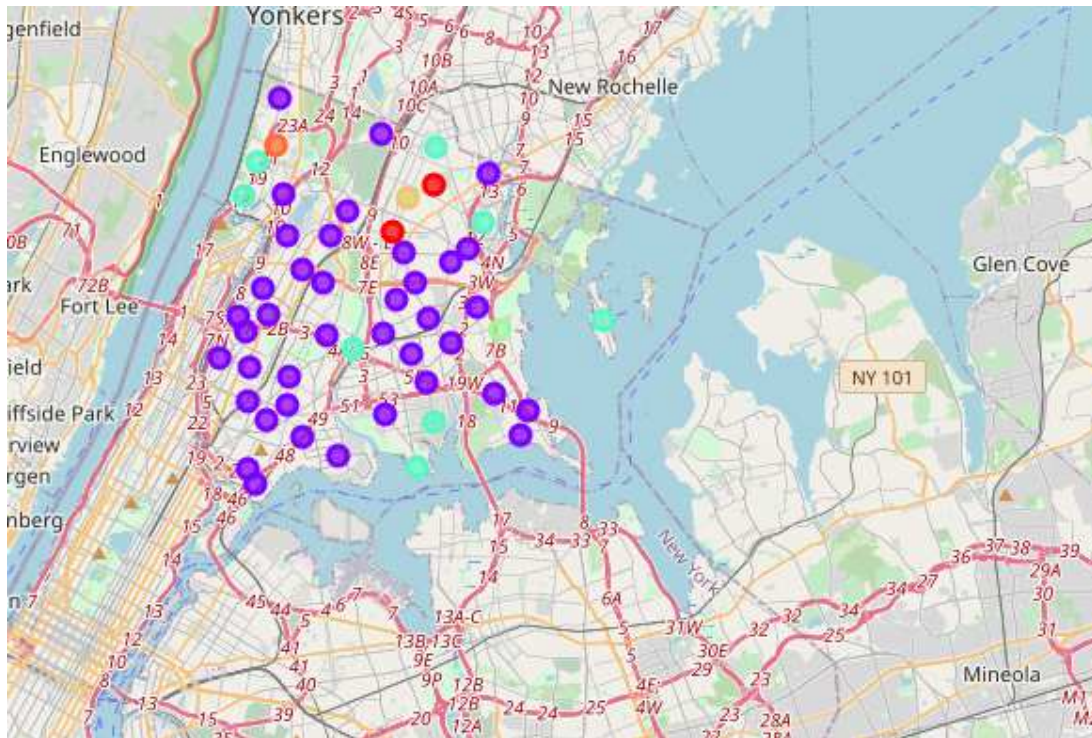
	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Allerton	Pizza Place	Cosmetics Shop	Deli / Bodega	Bus Station	Supermarket
1	Baychester	Donut Shop	Spanish Restaurant	Gym / Fitness Center	Bus Station	Mattress Store
2	Bedford Park	Mexican Restaurant	Diner	Deli / Bodega	Pizza Place	Spanish Restaurant
3	Belmont	Italian Restaurant	Pizza Place	Deli / Bodega	Bakery	Bank
4	Bronxdale	Pizza Place	Bank	Performing Arts Venue	Paper / Office Supplies Store	Chinese Restaurant

Clusterization will be done by 8 clusters. After clusterization we get draw two maps with colored clusters in each city:

Moscow



New York City



If we will try to compare results by colors - we can see that these two areas have nothing similar. Let's inspect frequency of each cluster in both cities:

New York City clusters and frequency:

1 39
4 8
0 2
7 1
6 1
5 1

Moscow clusters and frequency:

2 26
0 22
4 6
3 1
1 1

Most popular clusters in NY are 1 and 4 (and it's 39+8=47 places ($47/52 \cdot 100\% = 90,4\%$ of all)) and most popular clusters in Moscow are 2 and 0 (it's 26+22=48 places (85,7% of all), very similar quantity to NY).

If inspect 5 most common venues for every point in each of these four clusters: 1 and 4 for NY and 2 and 0 for Moscow, then we count frequency of each venue in two clusters for every city and show first 15 rows – we get complete comparison:

	Moscow Place	Moscow counts	NY Place	NY counts
0	Supermarket	29	Pizza Place	28
1	Pizza Place	11	Deli / Bodega	14
2	Gym / Fitness Center	11	Pharmacy	13
3	Food & Drink Shop	9	Donut Shop	12
4	Cosmetics Shop	7	Bus Station	11
5	Fast Food Restaurant	7	Grocery Store	11
6	Park	7	Bank	10
7	Bus Stop	7	Spanish Restaurant	7
8	Pharmacy	6	Italian Restaurant	6

9	Convenience Store	6	Sandwich Place	6
10	Mobile Phone Shop	5	Park	6
11	Sushi Restaurant	5	Chinese Restaurant	6
12	Café	5	Fast Food Restaurant	6
13	Clothing Store	4	Supermarket	5
14	Soccer Field	4	Mexican Restaurant	5

4. Results and discussion

Achieved results show us that two areas are very different with only 30% of similar places in top 15 places of each area. They are different after comparison of most common places in almost the same quantity of small areas inside of each big area. Though some places have different types, but they have similar ideas – fast food and sandwich place, Spanish/Italian/Mexican restaurants and Sushi restaurants, supermarkets and deli/bodega with grocery store... It's not the same, but it's somewhere similar anyway.

5. Conclusion

As we can see from previous research - two areas with similar land area, population and density from two biggest cities from different sides of the globe are very different. From top 15 places of more than 85% of all common places for these areas only pizza places, pharmacy, bus stops, grocery stores and parks are present in both areas in different proportions. Such similarity will allow you to eat pizza, buy pills and go to park if you will decide to change your location. But you will not be able to visit fitness center and cosmetics shop or even play football if you will migrate from Moscow to NYC, and you will not be able to buy donuts and visit Spanish restaurant in case of NYC to Moscow migration. In conclusion I would like to say that location change from South-Eastern part of Moscow to The Bronx in New York or vice versa will not be the easiest thing with familiar places in not familiar distance.