

The Battle of Neighborhoods - Moscow

May 24, 2020

Segmenting and Clustering Neighborhoods in Moscow

0.1 Introduction

As a young professional living in Moscow I am currently renting an apartment, but I am considering buying my own apartment in the future. I like the neighborhood I currently live in and I want to find some similar neighborhoods for possible relocation.

0.2 Table of Contents

1. Description of Data
2. Methodology
3. Segmenting by price
4. Download and Explore Dataset
5. Explore and Pre-Select Neighborhoods in Moscow
6. Analyze Each Neighborhood
7. Cluster Neighborhoods
8. Examine Clusters
9. Conclusion

Before we get the data and start exploring it, let's download all the dependencies that we will need.

```
[1]: import numpy as np # library to handle data in a vectorized manner

import pandas as pd # library for data analysis
pd.set_option('display.max_columns', None)
pd.set_option('display.max_rows', None)

import requests # library to handle requests

# Matplotlib and associated plotting modules
import matplotlib.cm as cm
import matplotlib.colors as colors

# import k-means from clustering stage
```

```

from sklearn.cluster import KMeans

# import DBSCAN from clustering stage
from sklearn.cluster import DBSCAN
from sklearn import metrics
from sklearn.preprocessing import StandardScaler

#!conda install -c conda-forge folium=0.5.0 --yes
import folium # map rendering library

print('Libraries imported.')

```

Libraries imported.

0.3 Description of data

There are several sources of data used in this analysis. 1. Moscow geographical data on boroughs and neighborhoods Gis-LAB, an independent resource on Moscow geo data, can be used to obtain a CSV file with neighborhoods, boroughs and geo data on the borders of the neighborhoods in latitudes and longitudes. That data will then have to be adjusted as for the clustering I will need a single point within the borders of the neighborhood for segmentation.

2. Data on average price in different Moscow neighborhoods
IRN, analytical center for real estate, has the data on average price per sq.m for neighborhoods in Moscow. That data will be used for pre-selecting neighborhoods based on price.
3. Foursquare API data on venues. Having pre-selected neighborhoods data i will use a get request to get data on venues located near the neighborhood center. I will only need data on the venue type, as the neighborhoods will be clustered based on that.

0.4 Methodology

In this notebook, I will use Foursquare API to explore neighborhoods in Moscow to find appropriate neighborhoods for relocation. I will pre-select the neighborhoods using DBSCAN algorithm based on price and then use the **explore** function to get the most common venue categories in each neighborhood, and then use this feature to group the neighborhoods into clusters. I will use the *k*-means clustering algorithm to complete this task. Finally, i will use the Folium library to visualize the neighborhoods in Moscow and their emerging clusters and based on the clustering results choose the appropriate neighborhoods.

0.5 Download and Explore Dataset

In order to segment the neighborhoods and explore them, we will essentially need a dataset that contains the boroughs and the neighborhoods that exist in each borough as well as the the latitude and logitude coordinates of each neighborhood.

Let's first download the file with names of boroughs and neighborhoods in Moscow and their geographical coordinates and have a look on it.

```
[3]: moscow_df=pd.read_csv('mo.csv')
      moscow_df.head()
```

```
[3]:
```

		WKT	NAME \
0	MULTIPOLYGON (((36.8031012 55.4408329,36.80319...		
1	POLYGON ((37.4276499 55.7482092,37.4284863 55...		
2	POLYGON ((36.8035692 55.4516224,36.8045117 55...		
3	POLYGON ((36.9372397 55.2413907,36.9372604 55...		
4	POLYGON ((37.4395575 55.6273129,37.4401803 55...	" "	" "

	OKATO	OKTMO	NAME_AO	OKATO_AO	ABBREV_AO \
0	45298555	45945000		45298000	
1	45268595	45328000		45268000	
2	45298567	45954000		45298000	
3	45298575	45956000		45298000	
4	45297568	45953000		45297000	

	TYPE_MO
0	
1	
2	
3	
4	

To save space I can remove columns with unnecessary information.

```
[4]: moscow_df=moscow_df.drop(columns=['OKATO', 'OKTMO', 'NAME_AO', 'OKATO_AO', 'ABBREV_AO', 'TYPE_MO'])
```

Proceeding with tidying up the data i will rename the columns for clarity and analyze the results by boroughs.

```
[5]: moscow_df.rename(columns={'NAME': 'Neighborhood', 'ABBREV_AO': 'Borough'}, inplace=True)
      moscow_df['Borough'].value_counts()
```

```
[5]:
```

	17
	16
	16
	16
	13
	12
	12
11	
10	
	10
	8
	5

Name: Borough, dtype: int64

At this stage I understand, that I am mostly interested in certain boroughs, mostly located to the North or West of the City, closer to where I work. So let's remove the other boroughs from analysis.

```
[6]: boroughs=['Central', 'North', 'South', 'East', 'West']  
moscow_df = moscow_df[moscow_df.Borough.isin(boroughs)]
```

Checking the resulting shape of the dataframe.

```
[7]: moscow_df.shape
```

[7]: (64, 3)

Now we need to get a single central point for longitude and latitude for each neighborhood. Currently we have a list of all points representing the border of each neighborhood in the column WKT. I will get a mean of latitude and longitude for all of those points for my further analysis.

```
[8]: # define a function that takes mean of a string of coordinates divided by a
      ↪ comma
def coord(list):

    i=0
    row1=[]
    for coordinate in coordinates:
        r=coordinates[i].split(',')
        r=[float(r[0].replace('(',' ').replace(')',' '),float(r[1].
        ↪replace('(',' ').replace(')',' '))]
        row1=row1+r
        i=i+1
        latitude=sum((row1[:2]))/len(row1[:2])
        longitude=sum((row1[1:2]))/len(row1[1:2])
    return [latitude,longitude]
```

```
[9]: # create an empty table to store coordinates
df_lalg=pd.DataFrame(columns=['Latitude','Longitude'])
i=0
# run a cycle for all neighborhoods
while i<=moscow_df.shape[0]-1:
    coordinates=moscow_df['WKT'].iloc[i].split(' ')
    coordinates=coordinates[2:len(coordinates)-2]
    lg=coord(coordinates)
    df_lalg=df_lalg.append({'Latitude':lg[0],'Longitude':
    ↪lg[1]},ignore_index=True)
    i=i+1
```

```
[10]: # merge the two tables together
moscow_df.reset_index(drop=True,inplace=True)
moscow_df=moscow_df.merge(df_lalg,left_index=True,right_index=True)
# drop unnecessary columns
moscow_df=moscow_df.drop(columns=['WKT'])
```

Let's check the resulting dataframe.

```
[11]: moscow_df.head()
```

```
[11]:      Neighborhood Borough  Latitude  Longitude
0                55.748708  37.474149
1                55.612560  37.280081
2                55.798118  37.475900
3                55.886669  37.529184
4                55.865166  37.486793
```

```
[12]: moscow_df.shape
```

```
[12]: (64, 4)
```

0.6 Segmenting by price

Let's start with downloading the file with the data on average price per sq.m. for neighborhoods in Moscow and add it to the original dataframe.

```
[13]: moscow_price=pd.read_excel('Price.xlsx')
moscow_price.head()
```

```
[13]:      Neighborhood  Price per sq.m  Unnamed: 2  Unnamed: 3  Unnamed: 4  \
0                403710             NaN          NaN          NaN
1                384385             NaN          NaN          NaN
2                370425             NaN          NaN          NaN
3                361844             NaN          NaN          NaN
4                345097             NaN          NaN          NaN

      Unnamed: 5  Unnamed: 6  Unnamed: 7
0             NaN          NaN          NaN
1             NaN          NaN          NaN
2             NaN          NaN          NaN
3             NaN          NaN          NaN
4             NaN          NaN          NaN
```

```
[14]: moscow_df=moscow_df.join(moscow_price[['Neighborhood','Price per sq.m']]).
      ↪set_index('Neighborhood'), on='Neighborhood')
```

```
[15]: moscow_df.shape
```

```
[15]: (64, 5)
```

Further I will check, whether all neighborhoods have been assigned a price per sq.m.

```
[16]: moscow_df['Price per sq.m'].isna().value_counts()
```

```
[16]: False      62
      True       2
      Name: Price per sq.m, dtype: int64
```

Since some of the neighborhoods are missing a price, I will assign an average price to those neighborhoods.

```
[17]: moscow_df.describe()
```

```
[17]:
```

	Latitude	Longitude	Price per sq.m
count	64.000000	64.000000	62.000000
mean	55.802519	37.529175	196998.677419
std	0.073974	0.109717	58668.458329
min	55.612560	37.104304	137036.000000
25%	55.750567	37.473558	155005.750000
50%	55.806166	37.550749	183106.000000
75%	55.865264	37.610455	207106.000000
max	55.943791	37.701189	384385.000000

```
[18]: moscow_df.fillna(value=moscow_df['Price per sq.m'].mean(),inplace=True)
```

Let's check the results.

```
[19]: moscow_df['Price per sq.m'].isna().value_counts()
```

```
[19]: False      64
      Name: Price per sq.m, dtype: int64
```

All set, the data is ready for clustering.

```
[20]: # transform the data for segmenting into a NumPy array
```

```
price_clustering=moscow_df['Price per sq.m']
price_clustering=price_clustering.to_numpy()
price_clustering=np.reshape(price_clustering,(-1, 1))
```

```
[21]: # transform data for clustering
X= StandardScaler().fit_transform(price_clustering)

# compute DBSCAN
db = DBSCAN(eps=0.3, min_samples=10).fit(X)
core_samples_mask = np.zeros_like(db.labels_, dtype=bool)
core_samples_mask[db.core_sample_indices_] = True
```

```

labels = db.labels_

# number of clusters in labels, ignoring noise if present.
n_clusters_ = len(set(labels)) - (1 if -1 in labels else 0)
n_noise_ = list(labels).count(-1)

print('Estimated number of clusters: %d' % n_clusters_)
print('Estimated number of noise points: %d' % n_noise_)
print("Silhouette Coefficient: %0.3f"
      % metrics.silhouette_score(X, labels))

```

Estimated number of clusters: 1
 Estimated number of noise points: 13
 Silhouette Coefficient: 0.713

```

[22]: # add clustering labels
moscow_df.insert(0, 'Cluster Labels', labels)

moscow_df.head()

```

```

[22]:   Cluster Labels   Neighborhood Borough  Latitude  Longitude \
0              0              55.748708  37.474149
1              0              55.612560  37.280081
2              0              55.798118  37.475900
3              0              55.886669  37.529184
4              0              55.865166  37.486793

      Price per sq.m
0      207763.0
1      137036.0
2      188032.0
3      153984.0
4      164663.0

```

Let's examine the results of clustering.

```

[23]: moscow_df['Cluster Labels'].value_counts()

```

```

[23]: 0      51
      -1     13
      Name: Cluster Labels, dtype: int64

```

```

[24]: moscow_df.head()

```

```

[24]:   Cluster Labels   Neighborhood Borough  Latitude  Longitude \
0              0              55.748708  37.474149
1              0              55.612560  37.280081
2              0              55.798118  37.475900

```

3	0	55.886669	37.529184
4	0	55.865166	37.486793

	Price per sq.m
0	207763.0
1	137036.0
2	188032.0
3	153984.0
4	164663.0

DBSCAN algorithm identified 1 cluster and several noise points. Now I will examine the noise points.

```
[25]: moscow_df[moscow_df['Cluster Labels'] == -1]
```

```
[25]:
```

	Cluster Labels	Neighborhood	Borough	Latitude	Longitude	\
22	-1			55.696332	37.504112	
27	-1			55.738817	37.532958	
29	-1			55.781101	37.568215	
32	-1			55.762533	37.579689	
37	-1			55.740496	37.589018	
43	-1			55.751186	37.595816	
45	-1			55.769137	37.605589	
46	-1			55.729036	37.610110	
51	-1			55.776174	37.627336	
53	-1			55.730883	37.631416	
55	-1			55.765190	37.667150	
56	-1			55.773126	37.647874	
58	-1			55.742243	37.665034	

	Price per sq.m
22	244627.0
27	256013.0
29	242010.0
32	309195.0
37	330257.0
43	370425.0
45	345097.0
46	384385.0
51	276583.0
53	312005.0
55	233088.0
56	267042.0
58	254289.0

The results seem correct, as all of those neighborhoods are very expensive and should be excluded from further classification. So I will drop those noise points and irrelevant columns.


```
[26]: moscow_df = moscow_df[moscow_df['Cluster Labels'] == 0].reset_index(drop=True)
moscow_df.drop(columns=['Cluster Labels', 'Price per sq.m'], inplace=True)

moscow_df.head()
```

```
[26]:
```

	Neighborhood	Borough	Latitude	Longitude
0			55.748708	37.474149
1			55.612560	37.280081
2			55.798118	37.475900
3			55.886669	37.529184
4			55.865166	37.486793

```
[27]: print('The resulting dataframe has {} boroughs and {} neighborhoods.'.format(
        len(moscow_df['Borough'].unique()),
        moscow_df.shape[0]
    )
)
```

The resulting dataframe has 4 boroughs and 51 neighborhoods.

Let's view the selected neighborhoods on a map.

```
[28]: # create map of Moscow using latitude and longitude values
map_moscow = folium.Map(location=[55.751244, 37.618423], zoom_start=12)

# add markers to map
for lat, lng, borough, neighborhood in zip(moscow_df['Latitude'],
    ↪ moscow_df['Longitude'], moscow_df['Borough'], moscow_df['Neighborhood']):
    label = '{} , {}'.format(neighborhood, borough)
    label = folium.Popup(label, parse_html=True)
    folium.CircleMarker(
        [lat, lng],
        radius=5,
        popup=label,
        color='blue',
        fill=True,
        fill_color='#3186cc',
        fill_opacity=0.7,
        parse_html=False).add_to(map_moscow)

map_moscow
```

```
[28]: <folium.folium.Map at 0x7f7ef5dee390>
```

Define Foursquare Credentials and Version

```
[29]: CLIENT_ID = 'SK0IS030BOSVY4PBGQXI5SE2DEMFCTPBUZLHLAE3N3FONFNI' # your_
    ↪ Foursquare ID
```

```

CLIENT_SECRET = 'TKJXRTX3FZFXSNGZEULGGOLVYWAPZQWDCTQM2T4KTZJM4W3L' # your_
↳Foursquare Secret
VERSION = '20201605' # Foursquare API version

print('Your credentails:')
print('CLIENT_ID: ' + CLIENT_ID)
print('CLIENT_SECRET:' + CLIENT_SECRET)

```

Your credentails:

CLIENT_ID: SK0IS030B0SVY4PBGQXI5SE2DEMFC TPBUZLHLAE3N3FONFNI

CLIENT_SECRET:TKJXRTX3FZFXSNGZEULGGOLVYWAPZQWDCTQM2T4KTZJM4W3L

```

[30]: LIMIT = 100 # limit of number of venues returned by Foursquare API
radius = 500 # define radius

```

0.7 Explore Neighborhoods in Moscow

Let's create a function to repeat the same process to all the neighborhoods in Moscow

```

[31]: def getNearbyVenues(names, latitudes, longitudes, radius=500):

    venues_list=[]
    for name, lat, lng in zip(names, latitudes, longitudes):
        print(name)

        # create the API request URL
        url = 'https://api.foursquare.com/v2/venues/explore?
↳&client_id={} &client_secret={} &v={} &ll={},{} &radius={} &limit={}'.format(
            CLIENT_ID,
            CLIENT_SECRET,
            VERSION,
            lat,
            lng,
            radius,
            LIMIT)

        # make the GET request
        results = requests.get(url).json()["response"]['groups'][0]['items']

        # return only relevant information for each nearby venue
        venues_list.append([
            name,
            lat,
            lng,
            v['venue']['name'],
            v['venue']['location']['lat'],
            v['venue']['location']['lng'],

```

```

        v['venue']['categories'][0]['name']) for v in results])

    nearby_venues = pd.DataFrame([item for venue_list in venues_list for item_
    ↪in venue_list])
    nearby_venues.columns = ['Neighborhood',
                             'Neighborhood Latitude',
                             'Neighborhood Longitude',
                             'Venue',
                             'Venue Latitude',
                             'Venue Longitude',
                             'Venue Category']

    return(nearby_venues)

```

Now write the code to run the above function on each neighborhood and create a new dataframe called *moscow_venues*.

```

[32]: moscow_venues = getNearbyVenues(names=moscow_df['Neighborhood'],
                                     latitudes=moscow_df['Latitude'],
                                     longitudes=moscow_df['Longitude']
                                     )

```

-

-
-
-
-

-

Let's check the size of the resulting dataframe.

```
[33]: print(moscow_venues.shape)
      moscow_venues.head()
```

```
(854, 7)
```

```
[33]:
```

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	\
0		55.748708	37.474149	
1		55.748708	37.474149	
2		55.748708	37.474149	
3		55.748708	37.474149	
4		55.748708	37.474149	

	Venue	Venue Latitude	Venue Longitude	Venue Category
0		55.750919	37.478342	Athletics & Sports
1		55.747616	37.474336	Beach
2		55.745042	37.471349	BBQ Joint
3		55.750337	37.479428	Tea Room
4		55.748931	37.480362	Concert Hall

Let's check how many venues were returned for each neighborhood.

```
[34]: moscow_venues.groupby('Neighborhood').count()
```

```
[34]:
```

	Neighborhood	Latitude	Neighborhood	Longitude	Venue	\
		28		28	28	
		8		8	8	
		9		9	9	
		34		34	34	
		7		7	7	
		6		6	6	
		80		80	80	
		4		4	4	
		48		48	48	
		8		8	8	
		10		10	10	
		14		14	14	
		11		11	11	
		18		18	18	
		6		6	6	
		5		5	5	
		17		17	17	
		22		22	22	
		14		14	14	
		6		6	6	
		26		26	26	
		20		20	20	
		15		15	15	
		3		3	3	
-		4		4	4	
		33		33	33	
		16		16	16	
-		7		7	7	
-		20		20	20	
		12		12	12	
		16		16	16	
		26		26	26	
		20		20	20	
		24		24	24	
		11		11	11	
		3		3	3	
		54		54	54	
		6		6	6	
		15		15	15	
		6		6	6	
-		25		25	25	
-		9		9	9	
		5		5	5	

	Venue Latitude	Venue Longitude	Venue Category
Neighborhood			

	Venue Latitude	Venue Longitude	Venue Category
Neighborhood			

	6	6	6
	15	15	15
	6	6	6
-	25	25	25
-	9	9	9
	5	5	5
	22	22	22
-	21	21	21
	23	23	23
	30	30	30
	11	11	11
	11	11	11
	5	5	5

Let's find out how many unique categories can be curated from all the returned venues.

```
[35]: print('There are {} unique categories.'.format(len(moscow_venues['Venue_
      ↪Category'].unique())))
```

There are 205 unique categories.

```
[36]: moscow_venues['Venue Category'].value_counts()
```

```
[36]: Supermarket          37
      Coffee Shop          29
      Pharmacy             28
      Gym / Fitness Center  26
      Park                 25
      Pizza Place          24
      Café                 22
      Restaurant           20
      Bus Stop             20
      Clothing Store       18
      Cosmetics Shop       17
      Convenience Store    16
      Fast Food Restaurant  16
      Health Food Store    15
      Sushi Restaurant     13
      Food & Drink Shop     12
      Middle Eastern Restaurant 11
      Bakery               11
      Caucasian Restaurant  10
      Auto Workshop        10
      Sporting Goods Shop  10
      Flower Shop          10
      Pet Store            10
      Gym                  10
```

Bookstore	9
Salon / Barbershop	9
Hotel	8
Photography Studio	8
Farmers Market	7
Eastern European Restaurant	7
Electronics Store	7
Grocery Store	7
Athletics & Sports	7
Kids Store	7
Multiplex	7
Gourmet Shop	6
Wine Shop	6
Soccer Field	6
Arts & Crafts Store	6
Dance Studio	6
Yoga Studio	6
Shoe Store	6
Beer Bar	6
Shopping Mall	6
Japanese Restaurant	5
Blini House	5
Boutique	5
Martial Arts Dojo	5
Gastropub	5
Plaza	5
Liquor Store	5
Toy / Game Store	5
Playground	5
Science Museum	4
Arcade	4
Concert Hall	4
Lake	4
Mobile Phone Shop	4
Garden	4
Women's Store	4
Big Box Store	4
Asian Restaurant	4
Gym Pool	4
Hobby Shop	4
Baby Store	3
Pedestrian Plaza	3
Scenic Lookout	3
Pub	3
Noodle House	3
Stables	3
Italian Restaurant	3

Bath House	3
Paper / Office Supplies Store	3
Bar	3
Lingerie Store	3
Skate Park	3
Fountain	3
Dessert Shop	3
Ice Cream Shop	3
Other Repair Shop	2
Museum	2
Modern European Restaurant	2
Board Shop	2
Pool	2
Burger Joint	2
Vape Store	2
Tram Station	2
History Museum	2
Bus Line	2
Fabric Shop	2
Department Store	2
Hot Dog Joint	2
Beer Store	2
Recording Studio	2
Surf Spot	2
Music Venue	2
Health & Beauty Service	2
Dry Cleaner	2
Butcher	2
Market	2
Spa	2
Art Gallery	2
Automotive Shop	2
Dog Run	2
Chinese Restaurant	2
Climbing Gym	2
Medical Center	2
Theater	2
Soccer Stadium	2
Recreation Center	2
Accessories Store	2
Karaoke Bar	2
Snack Place	2
Tennis Court	2
Jewelry Store	2
Shipping Store	2
Train Station	1
Furniture / Home Store	1

Cocktail Bar	1
Gift Shop	1
Outdoor Sculpture	1
Tree	1
Pool Hall	1
Vegetarian / Vegan Restaurant	1
Stadium	1
Track	1
Track Stadium	1
Bowling Alley	1
Greek Restaurant	1
Hardware Store	1
Jewish Restaurant	1
Arts & Entertainment	1
Salad Place	1
Russian Restaurant	1
Bus Station	1
Tea Room	1
Warehouse Store	1
Gymnastics Gym	1
Dairy Store	1
Skating Rink	1
Sculpture Garden	1
Locksmith	1
Hockey Rink	1
Men's Store	1
Dumpling Restaurant	1
Outdoor Gym	1
Exhibit	1
Rock Climbing Spot	1
Breakfast Spot	1
Historic Site	1
Gun Range	1
Trail	1
Vietnamese Restaurant	1
Mexican Restaurant	1
Strip Club	1
Campground	1
Camera Store	1
Israeli Restaurant	1
Boxing Gym	1
Pelmeni House	1
Dive Spot	1
Tattoo Parlor	1
Nightclub	1
Supplement Shop	1
Belgian Restaurant	1

Comic Shop	1
Performing Arts Venue	1
Coworking Space	1
Adult Boutique	1
Rock Club	1
Frozen Yogurt Shop	1
Movie Theater	1
Aquarium	1
Indoor Play Area	1
Optical Shop	1
Lighthouse	1
Garden Center	1
Sports Club	1
Miscellaneous Shop	1
Indie Theater	1
Public Art	1
Forest	1
Smoke Shop	1
Design Studio	1
Brazilian Restaurant	1
Cantonese Restaurant	1
Rental Car Location	1
Print Shop	1
Food Court	1
College Gym	1
Botanical Garden	1
Seafood Restaurant	1
Volleyball Court	1
Shawarma Place	1
Donut Shop	1
Art Museum	1
Pie Shop	1
Beach	1
Film Studio	1
BBQ Joint	1
Massage Studio	1
Fish Market	1
Hookah Bar	1
Other Great Outdoors	1
Jazz Club	1
Name: Venue Category, dtype: int64	

0.7.1 Analyze each neighborhood

```
[37]: # one hot encoding
moscow_onehot = pd.get_dummies(moscow_venues[['Venue Category']], prefix="",
    ↪ prefix_sep="")

# add neighborhood column back to dataframe
moscow_onehot['Neighborhood'] = moscow_venues['Neighborhood']

# move neighborhood column to the first column
fixed_columns = [moscow_onehot.columns[-1]] + list(moscow_onehot.columns[:-1])
moscow_onehot = moscow_onehot[fixed_columns]

moscow_onehot.head()
```

```
[37]:
```

	Neighborhood	Accessories Store	Adult Boutique	Aquarium	Arcade	\
0		0	0	0	0	
1		0	0	0	0	
2		0	0	0	0	
3		0	0	0	0	
4		0	0	0	0	

	Art Gallery	Art Museum	Arts & Crafts Store	Arts & Entertainment	\
0	0	0	0	0	
1	0	0	0	0	
2	0	0	0	0	
3	0	0	0	0	
4	0	0	0	0	

	Asian Restaurant	Athletics & Sports	Auto Workshop	Automotive Shop	\
0	0	1	0	0	
1	0	0	0	0	
2	0	0	0	0	
3	0	0	0	0	
4	0	0	0	0	

	BBQ Joint	Baby Store	Bakery	Bar	Bath House	Beach	Beer Bar	\
0	0	0	0	0	0	0	0	
1	0	0	0	0	0	1	0	
2	1	0	0	0	0	0	0	
3	0	0	0	0	0	0	0	
4	0	0	0	0	0	0	0	

	Beer Store	Belgian Restaurant	Big Box Store	Blini House	Board Shop	\
0	0	0	0	0	0	
1	0	0	0	0	0	
2	0	0	0	0	0	

3	0	0	0	0	0
4	0	0	0	0	0

	Bookstore	Botanical Garden	Boutique	Bowling Alley	Boxing Gym	\
0	0	0	0	0	0	
1	0	0	0	0	0	
2	0	0	0	0	0	
3	0	0	0	0	0	
4	0	0	0	0	0	

	Brazilian Restaurant	Breakfast Spot	Burger Joint	Bus Line	Bus Station	\
0	0	0	0	0	0	
1	0	0	0	0	0	
2	0	0	0	0	0	
3	0	0	0	0	0	
4	0	0	0	0	0	

	Bus Stop	Butcher	Café	Camera Store	Campground	Cantonese Restaurant	\
0	0	0	0	0	0	0	
1	0	0	0	0	0	0	
2	0	0	0	0	0	0	
3	0	0	0	0	0	0	
4	0	0	0	0	0	0	

	Caucasian Restaurant	Chinese Restaurant	Climbing Gym	Clothing Store	\
0	0	0	0	0	
1	0	0	0	0	
2	0	0	0	0	
3	0	0	0	0	
4	0	0	0	0	

	Cocktail Bar	Coffee Shop	College Gym	Comic Shop	Concert Hall	\
0	0	0	0	0	0	
1	0	0	0	0	0	
2	0	0	0	0	0	
3	0	0	0	0	0	
4	0	0	0	0	1	

	Convenience Store	Cosmetics Shop	Coworking Space	Dairy Store	\
0	0	0	0	0	
1	0	0	0	0	
2	0	0	0	0	
3	0	0	0	0	
4	0	0	0	0	

	Dance Studio	Department Store	Design Studio	Dessert Shop	Dive Spot	\
0	0	0	0	0	0	

1	0	0	0	0	0
2	0	0	0	0	0
3	0	0	0	0	0
4	0	0	0	0	0

	Dog Run	Donut Shop	Dry Cleaner	Dumpling Restaurant	\
0	0	0	0	0	
1	0	0	0	0	
2	0	0	0	0	
3	0	0	0	0	
4	0	0	0	0	

	Eastern European Restaurant	Electronics Store	Exhibit	Fabric Shop	\
0	0	0	0	0	
1	0	0	0	0	
2	0	0	0	0	
3	0	0	0	0	
4	0	0	0	0	

	Farmers Market	Fast Food Restaurant	Film Studio	Fish Market	\
0	0	0	0	0	
1	0	0	0	0	
2	0	0	0	0	
3	0	0	0	0	
4	0	0	0	0	

	Flower Shop	Food & Drink Shop	Food Court	Forest	Fountain	\
0	0	0	0	0	0	
1	0	0	0	0	0	
2	0	0	0	0	0	
3	0	0	0	0	0	
4	0	0	0	0	0	

	Frozen Yogurt Shop	Furniture / Home Store	Garden	Garden Center	\
0	0	0	0	0	
1	0	0	0	0	
2	0	0	0	0	
3	0	0	0	0	
4	0	0	0	0	

	Gastropub	Gift Shop	Gourmet Shop	Greek Restaurant	Grocery Store	\
0	0	0	0	0	0	
1	0	0	0	0	0	
2	0	0	0	0	0	
3	0	0	0	0	0	
4	0	0	0	0	0	

	Gun Range	Gym	Gym / Fitness Center	Gym Pool	Gymnastics Gym	\
0	0	0	0	0	0	
1	0	0	0	0	0	
2	0	0	0	0	0	
3	0	0	0	0	0	
4	0	0	0	0	0	

	Hardware Store	Health & Beauty Service	Health Food Store	Historic Site	\
0	0	0	0	0	
1	0	0	0	0	
2	0	0	0	0	
3	0	0	0	0	
4	0	0	0	0	

	History Museum	Hobby Shop	Hockey Rink	Hookah Bar	Hot Dog Joint	Hotel	\
0	0	0	0	0	0	0	
1	0	0	0	0	0	0	
2	0	0	0	0	0	0	
3	0	0	0	0	0	0	
4	0	0	0	0	0	0	

	Ice Cream Shop	Indie Theater	Indoor Play Area	Israeli Restaurant	\
0	0	0	0	0	
1	0	0	0	0	
2	0	0	0	0	
3	0	0	0	0	
4	0	0	0	0	

	Italian Restaurant	Japanese Restaurant	Jazz Club	Jewelry Store	\
0	0	0	0	0	
1	0	0	0	0	
2	0	0	0	0	
3	0	0	0	0	
4	0	0	0	0	

	Jewish Restaurant	Karaoke Bar	Kids Store	Lake	Lighthouse	\
0	0	0	0	0	0	
1	0	0	0	0	0	
2	0	0	0	0	0	
3	0	0	0	0	0	
4	0	0	0	0	0	

	Lingerie Store	Liquor Store	Locksmith	Market	Martial Arts Dojo	\
0	0	0	0	0	0	
1	0	0	0	0	0	
2	0	0	0	0	0	
3	0	0	0	0	0	

4	0	0	0	0	0
---	---	---	---	---	---

	Massage Studio	Medical Center	Men's Store	Mexican Restaurant	\
0	0	0	0	0	
1	0	0	0	0	
2	0	0	0	0	
3	0	0	0	0	
4	0	0	0	0	

	Middle Eastern Restaurant	Miscellaneous Shop	Mobile Phone Shop	\
0	0	0	0	
1	0	0	0	
2	0	0	0	
3	0	0	0	
4	0	0	0	

	Modern European Restaurant	Movie Theater	Multiplex	Museum	Music Venue	\
0	0	0	0	0	0	
1	0	0	0	0	0	
2	0	0	0	0	0	
3	0	0	0	0	0	
4	0	0	0	0	0	

	Nightclub	Noodle House	Optical Shop	Other Great Outdoors	\
0	0	0	0	0	
1	0	0	0	0	
2	0	0	0	0	
3	0	0	0	0	
4	0	0	0	0	

	Other Repair Shop	Outdoor Gym	Outdoor Sculpture	\
0	0	0	0	
1	0	0	0	
2	0	0	0	
3	0	0	0	
4	0	0	0	

	Paper / Office Supplies Store	Park	Pedestrian Plaza	Pelmeni House	\
0	0	0	0	0	
1	0	0	0	0	
2	0	0	0	0	
3	0	0	0	0	
4	0	0	0	0	

	Performing Arts Venue	Pet Store	Pharmacy	Photography Studio	Pie Shop	\
0	0	0	0	0	0	
1	0	0	0	0	0	

2	0	0	0	0	0
3	0	0	0	0	0
4	0	0	0	0	0

	Pizza Place	Playground	Plaza	Pool	Pool Hall	Print Shop	Pub	\
0	0	0	0	0	0	0	0	
1	0	0	0	0	0	0	0	
2	0	0	0	0	0	0	0	
3	0	0	0	0	0	0	0	
4	0	0	0	0	0	0	0	

	Public Art	Recording Studio	Recreation Center	Rental Car Location	\
0	0	0	0	0	
1	0	0	0	0	
2	0	0	0	0	
3	0	0	0	0	
4	0	0	0	0	

	Restaurant	Rock Climbing Spot	Rock Club	Russian Restaurant	Salad Place	\
0	0	0	0	0	0	
1	0	0	0	0	0	
2	0	0	0	0	0	
3	0	0	0	0	0	
4	0	0	0	0	0	

	Salon / Barbershop	Scenic Lookout	Science Museum	Sculpture Garden	\
0	0	0	0	0	
1	0	0	0	0	
2	0	0	0	0	
3	0	0	0	0	
4	0	0	0	0	

	Seafood Restaurant	Shawarma Place	Shipping Store	Shoe Store	\
0	0	0	0	0	
1	0	0	0	0	
2	0	0	0	0	
3	0	0	0	0	
4	0	0	0	0	

	Shopping Mall	Skate Park	Skating Rink	Smoke Shop	Snack Place	\
0	0	0	0	0	0	
1	0	0	0	0	0	
2	0	0	0	0	0	
3	0	0	0	0	0	
4	0	0	0	0	0	

	Soccer Field	Soccer Stadium	Spa	Sporting Goods Shop	Sports Club	\
--	--------------	----------------	-----	---------------------	-------------	---

0	0	0	0	0	0	0
1	0	0	0	0	0	0
2	0	0	0	0	0	0
3	0	0	0	0	0	0
4	0	0	0	0	0	0

	Stables	Stadium	Strip Club	Supermarket	Supplement Shop	Surf Spot	\
0	0	0	0	0	0	0	
1	0	0	0	0	0	0	
2	0	0	0	0	0	0	
3	0	0	0	0	0	0	
4	0	0	0	0	0	0	

	Sushi Restaurant	Tattoo Parlor	Tea Room	Tennis Court	Theater	\
0	0	0	0	0	0	
1	0	0	0	0	0	
2	0	0	0	0	0	
3	0	0	1	0	0	
4	0	0	0	0	0	

	Toy / Game Store	Track	Track Stadium	Trail	Train Station	Tram Station	\
0	0	0	0	0	0	0	
1	0	0	0	0	0	0	
2	0	0	0	0	0	0	
3	0	0	0	0	0	0	
4	0	0	0	0	0	0	

	Tree Vape Store	Vegetarian / Vegan Restaurant	Vietnamese Restaurant	\
0	0	0	0	
1	0	0	0	
2	0	0	0	
3	0	0	0	
4	0	0	0	

	Volleyball Court	Warehouse Store	Wine Shop	Women's Store	Yoga Studio
0	0	0	0	0	0
1	0	0	0	0	0
2	0	0	0	0	0
3	0	0	0	0	0
4	0	0	0	0	0

And let's examine the new dataframe size.

```
[38]: moscow_onehot.shape
```

```
[38]: (854, 206)
```

Next, let's group rows by neighborhood and by taking the mean of the frequency of occurrence of

each category.

```
[39]: moscow_grouped = moscow_onehot.groupby('Neighborhood').mean().reset_index()
```

Let's confirm the new size.

```
[40]: moscow_grouped.head()
```

```
[40]:
```

	Neighborhood	Accessories Store	Adult Boutique	Aquarium	Arcade	\
0		0.0	0.0	0.0	0.0	
1		0.0	0.0	0.0	0.0	
2		0.0	0.0	0.0	0.0	
3		0.0	0.0	0.0	0.0	
4		0.0	0.0	0.0	0.0	

	Art Gallery	Art Museum	Arts & Crafts Store	Arts & Entertainment	\
0	0.0	0.0	0.000000	0.0	
1	0.0	0.0	0.000000	0.0	
2	0.0	0.0	0.000000	0.0	
3	0.0	0.0	0.029412	0.0	
4	0.0	0.0	0.000000	0.0	

	Asian Restaurant	Athletics & Sports	Auto Workshop	Automotive Shop	\
0	0.000000	0.0	0.107143	0.0	
1	0.000000	0.0	0.125000	0.0	
2	0.000000	0.0	0.000000	0.0	
3	0.000000	0.0	0.000000	0.0	
4	0.142857	0.0	0.000000	0.0	

	BBQ Joint	Baby Store	Bakery	Bar	Bath House	Beach	Beer Bar	\
0	0.0	0.000000	0.0	0.000000	0.0	0.0	0.0	
1	0.0	0.000000	0.0	0.000000	0.0	0.0	0.0	
2	0.0	0.000000	0.0	0.000000	0.0	0.0	0.0	
3	0.0	0.029412	0.0	0.000000	0.0	0.0	0.0	
4	0.0	0.000000	0.0	0.142857	0.0	0.0	0.0	

	Beer Store	Belgian Restaurant	Big Box Store	Blini House	Board Shop	\
0	0.0	0.0	0.0	0.000000	0.0	
1	0.0	0.0	0.0	0.000000	0.0	
2	0.0	0.0	0.0	0.000000	0.0	
3	0.0	0.0	0.0	0.029412	0.0	
4	0.0	0.0	0.0	0.000000	0.0	

	Bookstore	Botanical Garden	Boutique	Bowling Alley	Boxing Gym	\
0	0.000000	0.0	0.000000	0.0	0.0	
1	0.000000	0.0	0.000000	0.0	0.0	
2	0.000000	0.0	0.000000	0.0	0.0	
3	0.029412	0.0	0.029412	0.0	0.0	

4 0.000000 0.0 0.000000 0.0 0.0

	Brazilian Restaurant	Breakfast Spot	Burger Joint	Bus Line	Bus Station \
0	0.0	0.0	0.0	0.0	0.0
1	0.0	0.0	0.0	0.0	0.0
2	0.0	0.0	0.0	0.0	0.0
3	0.0	0.0	0.0	0.0	0.0
4	0.0	0.0	0.0	0.0	0.0

	Bus Stop	Butcher	Café	Camera Store	Campground \
0	0.035714	0.0	0.000000	0.0	0.0
1	0.125000	0.0	0.000000	0.0	0.0
2	0.000000	0.0	0.111111	0.0	0.0
3	0.029412	0.0	0.000000	0.0	0.0
4	0.142857	0.0	0.000000	0.0	0.0

	Cantonese Restaurant	Caucasian Restaurant	Chinese Restaurant \
0	0.0	0.035714	0.0
1	0.0	0.000000	0.0
2	0.0	0.000000	0.0
3	0.0	0.029412	0.0
4	0.0	0.000000	0.0

	Climbing Gym	Clothing Store	Cocktail Bar	Coffee Shop	College Gym \
0	0.0	0.000000	0.0	0.0	0.0
1	0.0	0.000000	0.0	0.0	0.0
2	0.0	0.111111	0.0	0.0	0.0
3	0.0	0.029412	0.0	0.0	0.0
4	0.0	0.000000	0.0	0.0	0.0

	Comic Shop	Concert Hall	Convenience Store	Cosmetics Shop \
0	0.0	0.0	0.000000	0.035714
1	0.0	0.0	0.000000	0.000000
2	0.0	0.0	0.000000	0.000000
3	0.0	0.0	0.029412	0.088235
4	0.0	0.0	0.000000	0.000000

	Coworking Space	Dairy Store	Dance Studio	Department Store \
0	0.0	0.0	0.035714	0.0
1	0.0	0.0	0.000000	0.0
2	0.0	0.0	0.000000	0.0
3	0.0	0.0	0.000000	0.0
4	0.0	0.0	0.000000	0.0

	Design Studio	Dessert Shop	Dive Spot	Dog Run	Donut Shop	Dry Cleaner \
0	0.0	0.0	0.0	0.035714	0.0	0.000
1	0.0	0.0	0.0	0.000000	0.0	0.125

2	0.0	0.0	0.0	0.000000	0.0	0.000
3	0.0	0.0	0.0	0.000000	0.0	0.000
4	0.0	0.0	0.0	0.000000	0.0	0.000

	Dumpling Restaurant	Eastern European Restaurant	Electronics Store	\
0	0.0		0.000000	
1	0.0		0.000000	
2	0.0		0.000000	
3	0.0		0.029412	
4	0.0		0.000000	

	Exhibit	Fabric Shop	Farmers Market	Fast Food Restaurant	Film Studio	\
0	0.0	0.000000	0.035714	0.000000	0.0	
1	0.0	0.000000	0.000000	0.000000	0.0	
2	0.0	0.000000	0.000000	0.000000	0.0	
3	0.0	0.029412	0.000000	0.029412	0.0	
4	0.0	0.000000	0.000000	0.142857	0.0	

	Fish Market	Flower Shop	Food & Drink Shop	Food Court	Forest	Fountain	\
0	0.0	0.0	0.000000	0.0	0.0	0.0	
1	0.0	0.0	0.000000	0.0	0.0	0.0	
2	0.0	0.0	0.000000	0.0	0.0	0.0	
3	0.0	0.0	0.029412	0.0	0.0	0.0	
4	0.0	0.0	0.000000	0.0	0.0	0.0	

	Frozen Yogurt Shop	Furniture / Home Store	Garden	Garden Center	\
0	0.0	0.0	0.0	0.0	
1	0.0	0.0	0.0	0.0	
2	0.0	0.0	0.0	0.0	
3	0.0	0.0	0.0	0.0	
4	0.0	0.0	0.0	0.0	

	Gastropub	Gift Shop	Gourmet Shop	Greek Restaurant	Grocery Store	\
0	0.0	0.035714	0.0	0.0	0.000	
1	0.0	0.000000	0.0	0.0	0.125	
2	0.0	0.000000	0.0	0.0	0.000	
3	0.0	0.000000	0.0	0.0	0.000	
4	0.0	0.000000	0.0	0.0	0.000	

	Gun Range	Gym	Gym / Fitness Center	Gym Pool	Gymnastics Gym	\
0	0.0	0.000000	0.071429	0.0	0.0	
1	0.0	0.000000	0.000000	0.0	0.0	
2	0.0	0.000000	0.000000	0.0	0.0	
3	0.0	0.029412	0.029412	0.0	0.0	
4	0.0	0.000000	0.000000	0.0	0.0	

Hardware Store	Health & Beauty Service	Health Food Store	Historic Site	\
----------------	-------------------------	-------------------	---------------	---

0	0.0	0.0	0.035714	0.0
1	0.0	0.0	0.000000	0.0
2	0.0	0.0	0.000000	0.0
3	0.0	0.0	0.029412	0.0
4	0.0	0.0	0.000000	0.0

	History Museum	Hobby Shop	Hockey Rink	Hookah Bar	Hot Dog Joint \
0	0.0	0.0	0.0	0.0	0.000000
1	0.0	0.0	0.0	0.0	0.000000
2	0.0	0.0	0.0	0.0	0.000000
3	0.0	0.0	0.0	0.0	0.029412
4	0.0	0.0	0.0	0.0	0.000000

	Hotel	Ice Cream Shop	Indie Theater	Indoor Play Area \
0	0.107143	0.0	0.0	0.0
1	0.000000	0.0	0.0	0.0
2	0.000000	0.0	0.0	0.0
3	0.000000	0.0	0.0	0.0
4	0.000000	0.0	0.0	0.0

	Israeli Restaurant	Italian Restaurant	Japanese Restaurant	Jazz Club \
0	0.0	0.0	0.000000	0.000000
1	0.0	0.0	0.000000	0.000000
2	0.0	0.0	0.000000	0.111111
3	0.0	0.0	0.029412	0.000000
4	0.0	0.0	0.000000	0.000000

	Jewelry Store	Jewish Restaurant	Karaoke Bar	Kids Store	Lake \
0	0.0	0.0	0.0	0.000000	0.000
1	0.0	0.0	0.0	0.000000	0.125
2	0.0	0.0	0.0	0.000000	0.000
3	0.0	0.0	0.0	0.029412	0.000
4	0.0	0.0	0.0	0.000000	0.000

	Lighthouse	Lingerie Store	Liquor Store	Locksmith	Market \
0	0.0	0.0	0.000000	0.0	0.0
1	0.0	0.0	0.000000	0.0	0.0
2	0.0	0.0	0.000000	0.0	0.0
3	0.0	0.0	0.029412	0.0	0.0
4	0.0	0.0	0.000000	0.0	0.0

	Martial Arts Dojo	Massage Studio	Medical Center	Men's Store \
0	0.000000	0.0	0.035714	0.0
1	0.000000	0.0	0.000000	0.0
2	0.111111	0.0	0.000000	0.0
3	0.000000	0.0	0.000000	0.0
4	0.000000	0.0	0.000000	0.0

	Mexican Restaurant	Middle Eastern Restaurant	Miscellaneous Shop	\
0	0.0	0.0	0.0	
1	0.0	0.0	0.0	
2	0.0	0.0	0.0	
3	0.0	0.0	0.0	
4	0.0	0.0	0.0	

	Mobile Phone Shop	Modern European Restaurant	Movie Theater	Multiplex	\
0	0.035714	0.0	0.0	0.0	
1	0.000000	0.0	0.0	0.0	
2	0.000000	0.0	0.0	0.0	
3	0.029412	0.0	0.0	0.0	
4	0.000000	0.0	0.0	0.0	

	Museum	Music Venue	Nightclub	Noodle House	Optical Shop	\
0	0.0	0.0	0.0	0.0	0.0	
1	0.0	0.0	0.0	0.0	0.0	
2	0.0	0.0	0.0	0.0	0.0	
3	0.0	0.0	0.0	0.0	0.0	
4	0.0	0.0	0.0	0.0	0.0	

	Other Great Outdoors	Other Repair Shop	Outdoor Gym	Outdoor Sculpture	\
0	0.0	0.0	0.0	0.0	
1	0.0	0.0	0.0	0.0	
2	0.0	0.0	0.0	0.0	
3	0.0	0.0	0.0	0.0	
4	0.0	0.0	0.0	0.0	

	Paper / Office Supplies Store	Park	Pedestrian Plaza	Pelmeni House	\
0	0.000000	0.035714	0.0	0.0	
1	0.000000	0.000000	0.0	0.0	
2	0.000000	0.000000	0.0	0.0	
3	0.029412	0.000000	0.0	0.0	
4	0.000000	0.142857	0.0	0.0	

	Performing Arts Venue	Pet Store	Pharmacy	Photography Studio	Pie Shop	\
0	0.0	0.000000	0.000000	0.0	0.0	
1	0.0	0.000000	0.000000	0.0	0.0	
2	0.0	0.000000	0.222222	0.0	0.0	
3	0.0	0.029412	0.058824	0.0	0.0	
4	0.0	0.000000	0.000000	0.0	0.0	

	Pizza Place	Playground	Plaza	Pool	Pool Hall	Print Shop	Pub	\
0	0.071429	0.0	0.000000	0.0	0.0	0.000000	0.000	
1	0.000000	0.0	0.000000	0.0	0.0	0.000000	0.125	
2	0.000000	0.0	0.000000	0.0	0.0	0.000000	0.000	

3	0.000000	0.0	0.029412	0.0	0.0	0.029412	0.000
4	0.000000	0.0	0.142857	0.0	0.0	0.000000	0.000

	Public Art	Recording Studio	Recreation Center	Rental Car Location	\
0	0.0	0.035714	0.0	0.0	0.0
1	0.0	0.000000	0.0	0.0	0.0
2	0.0	0.000000	0.0	0.0	0.0
3	0.0	0.000000	0.0	0.0	0.0
4	0.0	0.000000	0.0	0.0	0.0

	Restaurant	Rock Climbing Spot	Rock Club	Russian Restaurant	Salad Place	\
0	0.035714	0.0	0.0	0.0	0.0	0.0
1	0.000000	0.0	0.0	0.0	0.0	0.0
2	0.000000	0.0	0.0	0.0	0.0	0.0
3	0.000000	0.0	0.0	0.0	0.0	0.0
4	0.000000	0.0	0.0	0.0	0.0	0.0

	Salon / Barbershop	Scenic Lookout	Science Museum	Sculpture Garden	\
0	0.000000	0.0	0.0	0.0	0.0
1	0.125000	0.0	0.0	0.0	0.0
2	0.111111	0.0	0.0	0.0	0.0
3	0.000000	0.0	0.0	0.0	0.0
4	0.000000	0.0	0.0	0.0	0.0

	Seafood Restaurant	Shawarma Place	Shipping Store	Shoe Store	\
0	0.0	0.0	0.0	0.0	0.0
1	0.0	0.0	0.0	0.0	0.0
2	0.0	0.0	0.0	0.0	0.0
3	0.0	0.0	0.0	0.0	0.0
4	0.0	0.0	0.0	0.0	0.0

	Shopping Mall	Skate Park	Skating Rink	Smoke Shop	Snack Place	\
0	0.0	0.0	0.0	0.000000	0.035714	0.0
1	0.0	0.0	0.0	0.000000	0.000000	0.0
2	0.0	0.0	0.0	0.111111	0.000000	0.0
3	0.0	0.0	0.0	0.000000	0.000000	0.0
4	0.0	0.0	0.0	0.000000	0.000000	0.0

	Soccer Field	Soccer Stadium	Spa	Sporting Goods Shop	Sports Club	\
0	0.000	0.0	0.0	0.0	0.035714	0.0
1	0.125	0.0	0.0	0.0	0.000000	0.0
2	0.000	0.0	0.0	0.0	0.000000	0.0
3	0.000	0.0	0.0	0.0	0.000000	0.0
4	0.000	0.0	0.0	0.0	0.000000	0.0

	Stables	Stadium	Strip Club	Supermarket	Supplement Shop	Surf Spot	\
0	0.0	0.0	0.0	0.071429	0.0	0.0	0.0

1	0.0	0.0	0.0	0.000000	0.0	0.0
2	0.0	0.0	0.0	0.000000	0.0	0.0
3	0.0	0.0	0.0	0.029412	0.0	0.0
4	0.0	0.0	0.0	0.142857	0.0	0.0

	Sushi Restaurant	Tattoo Parlor	Tea Room	Tennis Court	Theater \
0	0.000000	0.0	0.0	0.0	0.0
1	0.000000	0.0	0.0	0.0	0.0
2	0.000000	0.0	0.0	0.0	0.0
3	0.029412	0.0	0.0	0.0	0.0
4	0.000000	0.0	0.0	0.0	0.0

	Toy / Game Store	Track	Track Stadium	Trail	Train Station	Tram Station \
0	0.000000	0.0	0.0	0.0	0.0	0.0
1	0.000000	0.0	0.0	0.0	0.0	0.0
2	0.000000	0.0	0.0	0.0	0.0	0.0
3	0.029412	0.0	0.0	0.0	0.0	0.0
4	0.000000	0.0	0.0	0.0	0.0	0.0

	Tree Vape Store	Vegetarian / Vegan Restaurant	Vietnamese Restaurant \
0	0.0	0.0	0.0
1	0.0	0.0	0.0
2	0.0	0.0	0.0
3	0.0	0.0	0.0
4	0.0	0.0	0.0

	Volleyball Court	Warehouse Store	Wine Shop	Women's Store	Yoga Studio
0	0.0	0.0	0.0	0.0	0.035714
1	0.0	0.0	0.0	0.0	0.000000
2	0.0	0.0	0.0	0.0	0.111111
3	0.0	0.0	0.0	0.0	0.029412
4	0.0	0.0	0.0	0.0	0.000000

Let's put that into a *pandas* dataframe.

```
[41]: moscow_grouped.drop(columns=["Bus Line", "Bus Stop"], inplace=True)
```

First, let's write a function to sort the venues in descending order.

```
[42]: def return_most_common_venues(row, num_top_venues):
    row_categories = row.iloc[1:]
    row_categories_sorted = row_categories.sort_values(ascending=False)

    return row_categories_sorted.index.values[0:num_top_venues]
```

Now let's create the new dataframe and display the top 10 venues for each neighborhood.

```
[43]: num_top_venues = 10

indicators = ['st', 'nd', 'rd']

# create columns according to number of top venues
columns = ['Neighborhood']
for ind in np.arange(num_top_venues):
    try:
        columns.append('{}{} Most Common Venue'.format(ind+1, indicators[ind]))
    except:
        columns.append('{}th Most Common Venue'.format(ind+1))

# create a new dataframe
neighborhoods_venues_sorted = pd.DataFrame(columns=columns)
neighborhoods_venues_sorted['Neighborhood'] = moscow_grouped['Neighborhood']

[44]: for ind in np.arange(moscow_grouped.shape[0]):
        neighborhoods_venues_sorted.iloc[ind, 1:] =
        ↪return_most_common_venues(moscow_grouped.iloc[ind, :], num_top_venues)

neighborhoods_venues_sorted.head(15)
```

```
[44]:
```

	Neighborhood	1st Most Common Venue	2nd Most Common Venue \
0		Hotel	Auto Workshop
1		Soccer Field	Pub
2		Pharmacy	Yoga Studio
3		Cosmetics Shop	Pharmacy
4		Park	Plaza
5		Italian Restaurant	Department Store
6		Coffee Shop	Photography Studio
7		Soccer Field	Fast Food Restaurant
8		Clothing Store	Shoe Store
9		Supermarket	Bookstore
10		Food & Drink Shop	Supermarket
11		Supermarket	Modern European Restaurant
12		Cosmetics Shop	Health Food Store
13		Café	Convenience Store
14		Pizza Place	Athletics & Sports

	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue \
0	Pizza Place	Supermarket	Gym / Fitness Center
1	Lake	Dry Cleaner	Auto Workshop
2	Martial Arts Dojo	Smoke Shop	Jazz Club
3	Yoga Studio	Electronics Store	Boutique
4	Supermarket	Asian Restaurant	Fast Food Restaurant
5	Skating Rink	Pharmacy	Gym
6	Kids Store	Martial Arts Dojo	Clothing Store

7	Scenic Lookout	Yoga Studio	Exhibit
8	Lingerie Store	Sporting Goods Shop	Coffee Shop
9	Pet Store	Miscellaneous Shop	Café
10	Park	Stadium	Playground
11	Gym	Pizza Place	Restaurant
12	Eastern European Restaurant	Pharmacy	Salon / Barbershop
13	Pharmacy	Farmers Market	Park
14	Auto Workshop	Track	Park

	6th Most Common Venue	7th Most Common Venue \
0	Yoga Studio	Farmers Market
1	Salon / Barbershop	Grocery Store
2	Café	Clothing Store
3	Caucasian Restaurant	Clothing Store
4	Bar	Yoga Studio
5	Bar	Farmers Market
6	Accessories Store	Sushi Restaurant
7	Forest	Food Court
8	Shopping Mall	Toy / Game Store
9	Gym	Baby Store
10	Pizza Place	Middle Eastern Restaurant
11	Pharmacy	Park
12	Café	Supermarket
13	Middle Eastern Restaurant	Gym / Fitness Center
14	Yoga Studio	Fabric Shop

	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Sports Club	Mobile Phone Shop	Snack Place
1	Exhibit	Food Court	Food & Drink Shop
2	Salon / Barbershop	Food Court	Food & Drink Shop
3	Convenience Store	Plaza	Pet Store
4	Fountain	Forest	Food Court
5	Forest	Food Court	Food & Drink Shop
6	Dance Studio	Music Venue	Boutique
7	Food & Drink Shop	Flower Shop	Fish Market
8	Seafood Restaurant	Garden Center	Bookstore
9	Dumpling Restaurant	Fish Market	Fountain
10	Gourmet Shop	Exhibit	Flower Shop
11	Eastern European Restaurant	Café	Soccer Field
12	Dance Studio	History Museum	Food & Drink Shop
13	Big Box Store	Hardware Store	Exhibit
14	Fountain	Forest	Food Court

```
[45]: moscow_grouped.head()
```

```
[45]:      Neighborhood  Accessories Store  Adult Boutique  Aquarium  Arcade \
0                0.0                0.0                0.0                0.0
```

1		0.0	0.0	0.0	0.0
2		0.0	0.0	0.0	0.0
3		0.0	0.0	0.0	0.0
4		0.0	0.0	0.0	0.0

	Art Gallery	Art Museum	Arts & Crafts Store	Arts & Entertainment	\
0	0.0	0.0	0.000000		0.0
1	0.0	0.0	0.000000		0.0
2	0.0	0.0	0.000000		0.0
3	0.0	0.0	0.029412		0.0
4	0.0	0.0	0.000000		0.0

	Asian Restaurant	Athletics & Sports	Auto Workshop	Automotive Shop	\
0	0.000000	0.0	0.107143		0.0
1	0.000000	0.0	0.125000		0.0
2	0.000000	0.0	0.000000		0.0
3	0.000000	0.0	0.000000		0.0
4	0.142857	0.0	0.000000		0.0

	BBQ Joint	Baby Store	Bakery	Bar	Bath House	Beach	Beer Bar	\
0	0.0	0.000000	0.0	0.000000	0.0	0.0		0.0
1	0.0	0.000000	0.0	0.000000	0.0	0.0		0.0
2	0.0	0.000000	0.0	0.000000	0.0	0.0		0.0
3	0.0	0.029412	0.0	0.000000	0.0	0.0		0.0
4	0.0	0.000000	0.0	0.142857	0.0	0.0		0.0

	Beer Store	Belgian Restaurant	Big Box Store	Blini House	Board Shop	\
0	0.0	0.0	0.0	0.000000		0.0
1	0.0	0.0	0.0	0.000000		0.0
2	0.0	0.0	0.0	0.000000		0.0
3	0.0	0.0	0.0	0.029412		0.0
4	0.0	0.0	0.0	0.000000		0.0

	Bookstore	Botanical Garden	Boutique	Bowling Alley	Boxing Gym	\
0	0.000000	0.0	0.000000	0.0	0.0	
1	0.000000	0.0	0.000000	0.0	0.0	
2	0.000000	0.0	0.000000	0.0	0.0	
3	0.029412	0.0	0.029412	0.0	0.0	
4	0.000000	0.0	0.000000	0.0	0.0	

	Brazilian Restaurant	Breakfast Spot	Burger Joint	Bus Station	Butcher	\
0	0.0	0.0	0.0	0.0	0.0	0.0
1	0.0	0.0	0.0	0.0	0.0	0.0
2	0.0	0.0	0.0	0.0	0.0	0.0
3	0.0	0.0	0.0	0.0	0.0	0.0
4	0.0	0.0	0.0	0.0	0.0	0.0

	Café	Camera Store	Campground	Cantonese Restaurant	\
0	0.000000	0.0	0.0	0.0	
1	0.000000	0.0	0.0	0.0	
2	0.111111	0.0	0.0	0.0	
3	0.000000	0.0	0.0	0.0	
4	0.000000	0.0	0.0	0.0	

	Caucasian Restaurant	Chinese Restaurant	Climbing Gym	Clothing Store	\
0	0.035714	0.0	0.0	0.000000	
1	0.000000	0.0	0.0	0.000000	
2	0.000000	0.0	0.0	0.111111	
3	0.029412	0.0	0.0	0.029412	
4	0.000000	0.0	0.0	0.000000	

	Cocktail Bar	Coffee Shop	College Gym	Comic Shop	Concert Hall	\
0	0.0	0.0	0.0	0.0	0.0	
1	0.0	0.0	0.0	0.0	0.0	
2	0.0	0.0	0.0	0.0	0.0	
3	0.0	0.0	0.0	0.0	0.0	
4	0.0	0.0	0.0	0.0	0.0	

	Convenience Store	Cosmetics Shop	Coworking Space	Dairy Store	\
0	0.000000	0.035714	0.0	0.0	
1	0.000000	0.000000	0.0	0.0	
2	0.000000	0.000000	0.0	0.0	
3	0.029412	0.088235	0.0	0.0	
4	0.000000	0.000000	0.0	0.0	

	Dance Studio	Department Store	Design Studio	Dessert Shop	Dive Spot	\
0	0.035714	0.0	0.0	0.0	0.0	
1	0.000000	0.0	0.0	0.0	0.0	
2	0.000000	0.0	0.0	0.0	0.0	
3	0.000000	0.0	0.0	0.0	0.0	
4	0.000000	0.0	0.0	0.0	0.0	

	Dog Run	Donut Shop	Dry Cleaner	Dumpling Restaurant	\
0	0.035714	0.0	0.000	0.0	
1	0.000000	0.0	0.125	0.0	
2	0.000000	0.0	0.000	0.0	
3	0.000000	0.0	0.000	0.0	
4	0.000000	0.0	0.000	0.0	

	Eastern European Restaurant	Electronics Store	Exhibit	Fabric Shop	\
0	0.0	0.000000	0.0	0.000000	
1	0.0	0.000000	0.0	0.000000	
2	0.0	0.000000	0.0	0.000000	
3	0.0	0.029412	0.0	0.029412	

4		0.0	0.000000	0.0	0.000000
---	--	-----	----------	-----	----------

	Farmers Market	Fast Food Restaurant	Film Studio	Fish Market	\
0	0.035714	0.000000	0.0	0.0	
1	0.000000	0.000000	0.0	0.0	
2	0.000000	0.000000	0.0	0.0	
3	0.000000	0.029412	0.0	0.0	
4	0.000000	0.142857	0.0	0.0	

	Flower Shop	Food & Drink Shop	Food Court	Forest	Fountain	\
0	0.0	0.000000	0.0	0.0	0.0	
1	0.0	0.000000	0.0	0.0	0.0	
2	0.0	0.000000	0.0	0.0	0.0	
3	0.0	0.029412	0.0	0.0	0.0	
4	0.0	0.000000	0.0	0.0	0.0	

	Frozen Yogurt Shop	Furniture / Home Store	Garden	Garden Center	\
0	0.0	0.0	0.0	0.0	
1	0.0	0.0	0.0	0.0	
2	0.0	0.0	0.0	0.0	
3	0.0	0.0	0.0	0.0	
4	0.0	0.0	0.0	0.0	

	Gastropub	Gift Shop	Gourmet Shop	Greek Restaurant	Grocery Store	\
0	0.0	0.035714	0.0	0.0	0.000	
1	0.0	0.000000	0.0	0.0	0.125	
2	0.0	0.000000	0.0	0.0	0.000	
3	0.0	0.000000	0.0	0.0	0.000	
4	0.0	0.000000	0.0	0.0	0.000	

	Gun Range	Gym	Gym / Fitness Center	Gym Pool	Gymnastics Gym	\
0	0.0	0.000000	0.071429	0.0	0.0	
1	0.0	0.000000	0.000000	0.0	0.0	
2	0.0	0.000000	0.000000	0.0	0.0	
3	0.0	0.029412	0.029412	0.0	0.0	
4	0.0	0.000000	0.000000	0.0	0.0	

	Hardware Store	Health & Beauty Service	Health Food Store	Historic Site	\
0	0.0	0.0	0.035714	0.0	
1	0.0	0.0	0.000000	0.0	
2	0.0	0.0	0.000000	0.0	
3	0.0	0.0	0.029412	0.0	
4	0.0	0.0	0.000000	0.0	

	History Museum	Hobby Shop	Hockey Rink	Hookah Bar	Hot Dog Joint	\
0	0.0	0.0	0.0	0.0	0.000000	
1	0.0	0.0	0.0	0.0	0.000000	

2	0.0	0.0	0.0	0.0	0.000000
3	0.0	0.0	0.0	0.0	0.029412
4	0.0	0.0	0.0	0.0	0.000000

	Hotel	Ice Cream Shop	Indie Theater	Indoor Play Area	\
0	0.107143	0.0	0.0	0.0	
1	0.000000	0.0	0.0	0.0	
2	0.000000	0.0	0.0	0.0	
3	0.000000	0.0	0.0	0.0	
4	0.000000	0.0	0.0	0.0	

	Israeli Restaurant	Italian Restaurant	Japanese Restaurant	Jazz Club	\
0	0.0	0.0	0.000000	0.000000	
1	0.0	0.0	0.000000	0.000000	
2	0.0	0.0	0.000000	0.111111	
3	0.0	0.0	0.029412	0.000000	
4	0.0	0.0	0.000000	0.000000	

	Jewelry Store	Jewish Restaurant	Karaoke Bar	Kids Store	Lake	\
0	0.0	0.0	0.0	0.000000	0.000	
1	0.0	0.0	0.0	0.000000	0.125	
2	0.0	0.0	0.0	0.000000	0.000	
3	0.0	0.0	0.0	0.029412	0.000	
4	0.0	0.0	0.0	0.000000	0.000	

	Lighthouse	Lingerie Store	Liquor Store	Locksmith	Market	\
0	0.0	0.0	0.000000	0.0	0.0	
1	0.0	0.0	0.000000	0.0	0.0	
2	0.0	0.0	0.000000	0.0	0.0	
3	0.0	0.0	0.029412	0.0	0.0	
4	0.0	0.0	0.000000	0.0	0.0	

	Martial Arts Dojo	Massage Studio	Medical Center	Men's Store	\
0	0.000000	0.0	0.035714	0.0	
1	0.000000	0.0	0.000000	0.0	
2	0.111111	0.0	0.000000	0.0	
3	0.000000	0.0	0.000000	0.0	
4	0.000000	0.0	0.000000	0.0	

	Mexican Restaurant	Middle Eastern Restaurant	Miscellaneous Shop	\
0	0.0	0.0	0.0	
1	0.0	0.0	0.0	
2	0.0	0.0	0.0	
3	0.0	0.0	0.0	
4	0.0	0.0	0.0	

Mobile Phone Shop	Modern European Restaurant	Movie Theater	Multiplex	\
-------------------	----------------------------	---------------	-----------	---

0	0.035714	0.0	0.0	0.0
1	0.000000	0.0	0.0	0.0
2	0.000000	0.0	0.0	0.0
3	0.029412	0.0	0.0	0.0
4	0.000000	0.0	0.0	0.0

	Museum	Music Venue	Nightclub	Noodle House	Optical Shop \
0	0.0	0.0	0.0	0.0	0.0
1	0.0	0.0	0.0	0.0	0.0
2	0.0	0.0	0.0	0.0	0.0
3	0.0	0.0	0.0	0.0	0.0
4	0.0	0.0	0.0	0.0	0.0

	Other Great Outdoors	Other Repair Shop	Outdoor Gym	Outdoor Sculpture \
0	0.0	0.0	0.0	0.0
1	0.0	0.0	0.0	0.0
2	0.0	0.0	0.0	0.0
3	0.0	0.0	0.0	0.0
4	0.0	0.0	0.0	0.0

	Paper / Office Supplies Store	Park	Pedestrian Plaza	Pelmeni House \
0	0.000000	0.035714	0.0	0.0
1	0.000000	0.000000	0.0	0.0
2	0.000000	0.000000	0.0	0.0
3	0.029412	0.000000	0.0	0.0
4	0.000000	0.142857	0.0	0.0

	Performing Arts Venue	Pet Store	Pharmacy	Photography Studio	Pie Shop \
0	0.0	0.000000	0.000000	0.0	0.0
1	0.0	0.000000	0.000000	0.0	0.0
2	0.0	0.000000	0.222222	0.0	0.0
3	0.0	0.029412	0.058824	0.0	0.0
4	0.0	0.000000	0.000000	0.0	0.0

	Pizza Place	Playground	Plaza	Pool	Pool Hall	Print Shop	Pub \
0	0.071429	0.0	0.000000	0.0	0.0	0.000000	0.000
1	0.000000	0.0	0.000000	0.0	0.0	0.000000	0.125
2	0.000000	0.0	0.000000	0.0	0.0	0.000000	0.000
3	0.000000	0.0	0.029412	0.0	0.0	0.029412	0.000
4	0.000000	0.0	0.142857	0.0	0.0	0.000000	0.000

	Public Art	Recording Studio	Recreation Center	Rental Car Location \
0	0.0	0.035714	0.0	0.0
1	0.0	0.000000	0.0	0.0
2	0.0	0.000000	0.0	0.0
3	0.0	0.000000	0.0	0.0
4	0.0	0.000000	0.0	0.0

	Restaurant	Rock Climbing Spot	Rock Club	Russian Restaurant	Salad Place \
0	0.035714	0.0	0.0	0.0	0.0
1	0.000000	0.0	0.0	0.0	0.0
2	0.000000	0.0	0.0	0.0	0.0
3	0.000000	0.0	0.0	0.0	0.0
4	0.000000	0.0	0.0	0.0	0.0

	Salon / Barbershop	Scenic Lookout	Science Museum	Sculpture Garden \
0	0.000000	0.0	0.0	0.0
1	0.125000	0.0	0.0	0.0
2	0.111111	0.0	0.0	0.0
3	0.000000	0.0	0.0	0.0
4	0.000000	0.0	0.0	0.0

	Seafood Restaurant	Shawarma Place	Shipping Store	Shoe Store \
0	0.0	0.0	0.0	0.0
1	0.0	0.0	0.0	0.0
2	0.0	0.0	0.0	0.0
3	0.0	0.0	0.0	0.0
4	0.0	0.0	0.0	0.0

	Shopping Mall	Skate Park	Skating Rink	Smoke Shop	Snack Place \
0	0.0	0.0	0.0	0.000000	0.035714
1	0.0	0.0	0.0	0.000000	0.000000
2	0.0	0.0	0.0	0.111111	0.000000
3	0.0	0.0	0.0	0.000000	0.000000
4	0.0	0.0	0.0	0.000000	0.000000

	Soccer Field	Soccer Stadium	Spa	Sporting Goods Shop	Sports Club \
0	0.000	0.0	0.0	0.0	0.035714
1	0.125	0.0	0.0	0.0	0.000000
2	0.000	0.0	0.0	0.0	0.000000
3	0.000	0.0	0.0	0.0	0.000000
4	0.000	0.0	0.0	0.0	0.000000

	Stables	Stadium	Strip Club	Supermarket	Supplement Shop	Surf Spot \
0	0.0	0.0	0.0	0.071429	0.0	0.0
1	0.0	0.0	0.0	0.000000	0.0	0.0
2	0.0	0.0	0.0	0.000000	0.0	0.0
3	0.0	0.0	0.0	0.029412	0.0	0.0
4	0.0	0.0	0.0	0.142857	0.0	0.0

	Sushi Restaurant	Tattoo Parlor	Tea Room	Tennis Court	Theater \
0	0.000000	0.0	0.0	0.0	0.0
1	0.000000	0.0	0.0	0.0	0.0
2	0.000000	0.0	0.0	0.0	0.0

3	0.029412	0.0	0.0	0.0	0.0
4	0.000000	0.0	0.0	0.0	0.0

	Toy / Game Store	Track	Track Stadium	Trail	Train Station	Tram Station \
0	0.000000	0.0	0.0	0.0	0.0	0.0
1	0.000000	0.0	0.0	0.0	0.0	0.0
2	0.000000	0.0	0.0	0.0	0.0	0.0
3	0.029412	0.0	0.0	0.0	0.0	0.0
4	0.000000	0.0	0.0	0.0	0.0	0.0

	Tree Vape Store	Vegetarian / Vegan Restaurant	Vietnamese Restaurant \
0	0.0 0.0	0.0	0.0
1	0.0 0.0	0.0	0.0
2	0.0 0.0	0.0	0.0
3	0.0 0.0	0.0	0.0
4	0.0 0.0	0.0	0.0

	Volleyball Court	Warehouse Store	Wine Shop	Women's Store	Yoga Studio
0	0.0	0.0	0.0	0.0	0.035714
1	0.0	0.0	0.0	0.0	0.000000
2	0.0	0.0	0.0	0.0	0.111111
3	0.0	0.0	0.0	0.0	0.029412
4	0.0	0.0	0.0	0.0	0.000000

0.8 Cluster Neighborhoods

Run *k*-means to cluster the neighborhood into 4 clusters.

```
[46]: # set number of clusters
kclusters = 4

moscow_grouped_clustering = moscow_grouped.drop('Neighborhood', 1)

# run k-means clustering
kmeans = KMeans(n_clusters=kclusters, random_state=0).
    ↪ fit(moscow_grouped_clustering)

# check cluster labels generated for each row in the dataframe
kmeans.labels_[0:10]
```

```
[46]: array([2, 2, 3, 0, 2, 0, 0, 2, 0, 0], dtype=int32)
```

Let's create a new dataframe that includes the cluster as well as the top 10 venues for each neighborhood.

```
[47]: # add clustering labels
neighborhoods_venues_sorted.insert(0, 'Cluster Labels', kmeans.labels_)
```

```

moscow_merged = moscow_df

# merge moscow_grouped with moscow_data to add latitude/longitude for each
↳ neighborhood
moscow_merged = moscow_merged.join(neighborhoods_venues_sorted.
↳ set_index('Neighborhood'), on='Neighborhood')

moscow_merged

```

```

[47]:

```

	Neighborhood	Borough	Latitude	Longitude	Cluster	Labels \
0		55.748708	37.474149		0.0	
1		55.612560	37.280081		2.0	
2		55.798118	37.475900		0.0	
3		55.886669	37.529184		0.0	
4		55.865166	37.486793		0.0	
5		55.704913	37.383375		2.0	
6	-	55.650576	37.354337		2.0	
7		55.797015	37.392717		0.0	
8		55.943791	37.381299		0.0	
9		55.855051	37.360428		2.0	
10		55.892015	37.378868		1.0	
11		55.765999	37.416811		2.0	
12		55.642583	37.398145		2.0	
13		55.848161	37.503539		2.0	
14		55.844646	37.421538		0.0	
15		55.855324	37.423243		3.0	
16	-	55.826513	37.446410		2.0	
17	-	55.774477	37.475388		0.0	
18	-	55.706315	37.452668		2.0	
19	-	55.666194	37.478399		0.0	
20		55.867730	37.465845		0.0	
21	-	55.717706	37.471783		1.0	
22		55.827699	37.498243		0.0	
23		55.804283	37.509413		0.0	
24		55.875493	37.519462		0.0	
25		55.678449	37.494415		0.0	
26		55.778313	37.529088		0.0	
27		55.830160	37.525658		3.0	
28		55.801448	37.549368		3.0	
29		55.942667	37.552130		1.0	
30		55.866847	37.552988		2.0	
31		55.829539	37.559509		0.0	
32		55.901305	37.568914		0.0	
33		55.873777	37.560752		0.0	
34		55.799518	37.566503		0.0	
35		55.881136	37.589701		2.0	
36		55.835338	37.589365		3.0	

37	55.865559	37.614190	0.0
38	55.808049	37.586794	0.0
39	55.836305	37.617413	0.0
40	55.889685	37.611491	0.0
41	55.790632	37.608042	0.0
42	55.855019	37.644431	2.0
43	55.873908	37.630261	0.0
44	55.885997	37.654836	0.0
45	55.834181	37.652304	2.0
46	55.814850	37.652527	2.0
47	55.870001	37.665404	0.0
48	55.859651	37.701189	2.0
49	55.879053	37.683302	0.0
50	55.745892	37.104304	NaN

	1st Most Common Venue	2nd Most Common Venue \
0	Concert Hall	Athletics & Sports
1	Soccer Field	Fast Food Restaurant
2	Plaza	Coffee Shop
3	Supermarket	Modern European Restaurant
4	Supermarket	Gym / Fitness Center
5	Fast Food Restaurant	Clothing Store
6	Gym Pool	Tennis Court
7	Supermarket	Department Store
8	Restaurant	Other Great Outdoors
9	Japanese Restaurant	Pedestrian Plaza
10	Park	Restaurant
11	Pizza Place	Athletics & Sports
12	Recreation Center	Convenience Store
13	Food & Drink Shop	Supermarket
14	Park	Cosmetics Shop
15	Convenience Store	Pharmacy
16	Fast Food Restaurant	Soccer Stadium
17	Coffee Shop	Caucasian Restaurant
18	Hotel	Fast Food Restaurant
19	Coffee Shop	Cosmetics Shop
20	Sporting Goods Shop	Convenience Store
21	Park	Italian Restaurant
22	Clothing Store	Shoe Store
23	Bakery	Electronics Store
24	Cosmetics Shop	Health Food Store
25	Bookstore	Sushi Restaurant
26	Supermarket	Coffee Shop
27	Café	Convenience Store
28	Pharmacy	Yoga Studio
29	Fountain	Park
30	Park	Plaza

31	Stables	Garden
32	Playground	Theater
33	Supermarket	Bookstore
34	Pharmacy	Supermarket
35	Soccer Field	Pub
36	Botanical Garden	Gym / Fitness Center
37	Pharmacy	Eastern European Restaurant
38	Coffee Shop	Photography Studio
39	Café	Museum
40	Italian Restaurant	Department Store
41	Caucasian Restaurant	Health Food Store
42	Pizza Place	Gym / Fitness Center
43	Bakery	Restaurant
44	Pharmacy	Flower Shop
45	Gym / Fitness Center	Middle Eastern Restaurant
46	Hotel	Auto Workshop
47	Cosmetics Shop	Pharmacy
48	Breakfast Spot	Hotel
49	Restaurant	Supermarket
50	NaN	NaN

	3rd Most Common Venue	4th Most Common Venue \
0	Beach	BBQ Joint
1	Scenic Lookout	Yoga Studio
2	Dessert Shop	Sporting Goods Shop
3	Gym	Pizza Place
4	Café	Pizza Place
5	Cosmetics Shop	Big Box Store
6	Gym / Fitness Center	Spa
7	Tram Station	Bakery
8	Lake	Yoga Studio
9	Gym Pool	Forest
10	Scenic Lookout	Yoga Studio
11	Auto Workshop	Track
12	Food & Drink Shop	Pizza Place
13	Park	Stadium
14	Gastropub	Gym
15	Gym / Fitness Center	Wine Shop
16	Auto Workshop	Farmers Market
17	Italian Restaurant	Sushi Restaurant
18	Lighthouse	Auto Workshop
19	Bakery	Gym / Fitness Center
20	Café	Sushi Restaurant
21	Gym / Fitness Center	Hotel
22	Lingerie Store	Sporting Goods Shop
23	Beer Bar	Bookstore
24	Eastern European Restaurant	Pharmacy

25	Gourmet Shop	Gym / Fitness Center
26	Clothing Store	Multiplex
27	Pharmacy	Farmers Market
28	Martial Arts Dojo	Smoke Shop
29	Eastern European Restaurant	Yoga Studio
30	Supermarket	Asian Restaurant
31	Gym	Frozen Yogurt Shop
32	Flower Shop	Skate Park
33	Pet Store	Miscellaneous Shop
34	Pet Store	Cosmetics Shop
35	Lake	Dry Cleaner
36	Pool	Convenience Store
37	Beer Store	Electronics Store
38	Kids Store	Martial Arts Dojo
39	Science Museum	Garden
40	Skating Rink	Pharmacy
41	Restaurant	Hookah Bar
42	Pool Hall	Movie Theater
43	Gym / Fitness Center	Caucasian Restaurant
44	Bakery	Paper / Office Supplies Store
45	Outdoor Sculpture	Supermarket
46	Pizza Place	Supermarket
47	Yoga Studio	Electronics Store
48	Concert Hall	Grocery Store
49	Park	Convenience Store
50	NaN	NaN

5th Most Common Venue	6th Most Common Venue \
0	Tea Room
1	Exhibit
2	Food & Drink Shop
3	Restaurant
4	Coffee Shop
5	Blini House
6	Yoga Studio
7	Pedestrian Plaza
8	Exhibit
9	Shopping Mall
10	Exhibit
11	Park
12	Supermarket
13	Playground
14	Supermarket
15	Beer Store
16	Gun Range
17	Kids Store
18	Park

19	Multiplex	Noodle House
20	Film Studio	Liquor Store
21	Restaurant	Dog Run
22	Coffee Shop	Shopping Mall
23	Pharmacy	Cosmetics Shop
24	Salon / Barbershop	Café
25	Supermarket	Middle Eastern Restaurant
26	Blini House	Gastropub
27	Park	Middle Eastern Restaurant
28	Jazz Club	Café
29	Fabric Shop	Forest
30	Fast Food Restaurant	Bar
31	Fountain	Forest
32	Café	Pizza Place
33	Café	Gym
34	Automotive Shop	Garden
35	Auto Workshop	Salon / Barbershop
36	Arts & Entertainment	Sushi Restaurant
37	Park	Soccer Field
38	Clothing Store	Accessories Store
39	Surf Spot	Restaurant
40	Gym	Bar
41	College Gym	Shawarma Place
42	Sporting Goods Shop	Board Shop
43	Supermarket	Athletics & Sports
44	Gym Pool	Sushi Restaurant
45	Park	Track Stadium
46	Gym / Fitness Center	Yoga Studio
47	Boutique	Caucasian Restaurant
48	Auto Workshop	Yoga Studio
49	Shipping Store	Pharmacy
50	NaN	NaN

	7th Most Common Venue	8th Most Common Venue \
0	Fast Food Restaurant	Frozen Yogurt Shop
1	Food Court	Food & Drink Shop
2	Fast Food Restaurant	Boxing Gym
3	Park	Eastern European Restaurant
4	Farmers Market	Sushi Restaurant
5	Sushi Restaurant	Gym / Fitness Center
6	Food Court	Food & Drink Shop
7	Restaurant	Rental Car Location
8	Food Court	Food & Drink Shop
9	Middle Eastern Restaurant	Campground
10	Food Court	Food & Drink Shop
11	Fabric Shop	Fountain
12	Yoga Studio	Fabric Shop

13	Middle Eastern Restaurant	Gourmet Shop
14	Arts & Crafts Store	Convenience Store
15	Grocery Store	Park
16	Gym / Fitness Center	Blini House
17	Gourmet Shop	Modern European Restaurant
18	Yoga Studio	Fabric Shop
19	Bookstore	Middle Eastern Restaurant
20	Fish Market	Supermarket
21	Athletics & Sports	Food Court
22	Toy / Game Store	Seafood Restaurant
23	Pizza Place	Coffee Shop
24	Supermarket	Dance Studio
25	Rock Club	Health & Beauty Service
26	Salon / Barbershop	Salad Place
27	Gym / Fitness Center	Big Box Store
28	Clothing Store	Salon / Barbershop
29	Food Court	Food & Drink Shop
30	Yoga Studio	Fountain
31	Food Court	Food & Drink Shop
32	Stables	Supermarket
33	Baby Store	Dumpling Restaurant
34	Sushi Restaurant	Middle Eastern Restaurant
35	Grocery Store	Exhibit
36	Fountain	Forest
37	Gym	Supermarket
38	Sushi Restaurant	Dance Studio
39	Eastern European Restaurant	Flower Shop
40	Farmers Market	Forest
41	Pizza Place	Middle Eastern Restaurant
42	Fast Food Restaurant	Farmers Market
43	Pharmacy	Middle Eastern Restaurant
44	Farmers Market	Grocery Store
45	Fabric Shop	Athletics & Sports
46	Farmers Market	Sports Club
47	Clothing Store	Convenience Store
48	Fabric Shop	Forest
49	Gym / Fitness Center	Food & Drink Shop
50	NaN	NaN

	9th Most Common Venue	10th Most Common Venue
0	Fountain	Forest
1	Flower Shop	Fish Market
2	Fountain	Furniture / Home Store
3	Café	Soccer Field
4	Grocery Store	Salon / Barbershop
5	Boutique	Hot Dog Joint
6	Flower Shop	Fish Market

7	Gym / Fitness Center	Pet Store
8	Flower Shop	Fish Market
9	Martial Arts Dojo	Sushi Restaurant
10	Flower Shop	Fish Market
11	Forest	Food Court
12	Forest	Food Court
13	Exhibit	Flower Shop
14	Wine Shop	Bath House
15	Yoga Studio	Food Court
16	Clothing Store	Park
17	Fast Food Restaurant	Park
18	Forest	Food Court
19	Blini House	Mexican Restaurant
20	Climbing Gym	Big Box Store
21	Food & Drink Shop	Flower Shop
22	Garden Center	Bookstore
23	Food & Drink Shop	Café
24	History Museum	Food & Drink Shop
25	Pedestrian Plaza	Historic Site
26	Greek Restaurant	Shopping Mall
27	Hardware Store	Exhibit
28	Food Court	Food & Drink Shop
29	Flower Shop	Fish Market
30	Forest	Food Court
31	Flower Shop	Fish Market
32	Restaurant	Liquor Store
33	Fish Market	Fountain
34	Liquor Store	Flower Shop
35	Food Court	Food & Drink Shop
36	Food Court	Food & Drink Shop
37	Gymnastics Gym	Sushi Restaurant
38	Music Venue	Boutique
39	Fountain	Caucasian Restaurant
40	Food Court	Food & Drink Shop
41	Electronics Store	Gastropub
42	Belgian Restaurant	Caucasian Restaurant
43	Café	Pizza Place
44	Multiplex	Café
45	Food & Drink Shop	Women's Store
46	Mobile Phone Shop	Snack Place
47	Plaza	Pet Store
48	Food Court	Food & Drink Shop
49	Wine Shop	Exhibit
50	NaN	NaN

Finally, let's visualize the resulting clusters

```
[48]: moscow_merged['Cluster Labels'].value_counts()
```

```
[48]: 0.0    28
      2.0    15
      3.0     4
      1.0     3
      Name: Cluster Labels, dtype: int64
```

Now let's create map of Moscow with clustered neighborhoods marked in different colors.

```
[49]: moscow_merged['Cluster Labels'].isna().value_counts()
```

```
[49]: False    50
      True     1
      Name: Cluster Labels, dtype: int64
```

```
[50]: moscow_merged.dropna(inplace=True)
```

```
[51]: moscow_merged=moscow_merged.astype({'Cluster Labels':'int32'})
```

```
[62]: # create map
map_clusters = folium.Map(location=[55.7, 37.4], zoom_start=11)

# set color scheme for the clusters
x = np.arange(kclusters)
ys = [i + x + (i*x)**2 for i in range(kclusters)]
colors_array = cm.rainbow(np.linspace(0, 1, len(ys)))
rainbow = [colors.rgb2hex(i) for i in colors_array]

# add markers to the map
markers_colors = []
for lat, lon, poi, cluster in zip(moscow_merged['Latitude'],
    ↪ moscow_merged['Longitude'], moscow_merged['Neighborhood'],
    ↪ moscow_merged['Cluster Labels']):
    label = folium.Popup(str(poi) + ' Cluster ' + str(cluster), parse_html=True)
    folium.CircleMarker(
        [lat, lon],
        radius=5,
        popup=label,
        color=rainbow[cluster-1],
        fill=True,
        fill_color=rainbow[cluster-1],
        fill_opacity=0.7).add_to(map_clusters)

map_clusters
```

```
[62]: <folium.folium.Map at 0x7f7ef41c70b8>
```

0.9 Examine Clusters

Now, we can examine each cluster and determine the discriminating venue categories that distinguish each cluster.

Cluster 1

```
[53]: moscow_merged.loc[moscow_merged['Cluster Labels'] == 0, moscow_merged.  
      ↪ columns[[0] + list(range(5, moscow_merged.shape[1]))]]
```

```
[53]:
```

	Neighborhood	1st Most Common Venue	2nd Most Common Venue \
0		Concert Hall	Athletics & Sports
2		Plaza	Coffee Shop
3		Supermarket	Modern European Restaurant
4		Supermarket	Gym / Fitness Center
7		Supermarket	Department Store
8		Restaurant	Other Great Outdoors
14		Park	Cosmetics Shop
17	-	Coffee Shop	Caucasian Restaurant
19	-	Coffee Shop	Cosmetics Shop
20		Sporting Goods Shop	Convenience Store
22		Clothing Store	Shoe Store
23		Bakery	Electronics Store
24		Cosmetics Shop	Health Food Store
25		Bookstore	Sushi Restaurant
26		Supermarket	Coffee Shop
31		Stables	Garden
32		Playground	Theater
33		Supermarket	Bookstore
34		Pharmacy	Supermarket
37		Pharmacy	Eastern European Restaurant
38		Coffee Shop	Photography Studio
39		Café	Museum
40		Italian Restaurant	Department Store
41		Caucasian Restaurant	Health Food Store
43		Bakery	Restaurant
44		Pharmacy	Flower Shop
47		Cosmetics Shop	Pharmacy
49		Restaurant	Supermarket

	3rd Most Common Venue	4th Most Common Venue \
0	Beach	BBQ Joint
2	Dessert Shop	Sporting Goods Shop
3	Gym	Pizza Place
4	Café	Pizza Place
7	Tram Station	Bakery
8	Lake	Yoga Studio
14	Gastropub	Gym
17	Italian Restaurant	Sushi Restaurant

19	Bakery	Gym / Fitness Center
20	Café	Sushi Restaurant
22	Lingerie Store	Sporting Goods Shop
23	Beer Bar	Bookstore
24	Eastern European Restaurant	Pharmacy
25	Gourmet Shop	Gym / Fitness Center
26	Clothing Store	Multiplex
31	Gym	Frozen Yogurt Shop
32	Flower Shop	Skate Park
33	Pet Store	Miscellaneous Shop
34	Pet Store	Cosmetics Shop
37	Beer Store	Electronics Store
38	Kids Store	Martial Arts Dojo
39	Science Museum	Garden
40	Skating Rink	Pharmacy
41	Restaurant	Hookah Bar
43	Gym / Fitness Center	Caucasian Restaurant
44	Bakery	Paper / Office Supplies Store
47	Yoga Studio	Electronics Store
49	Park	Convenience Store

5th Most Common Venue	6th Most Common Venue \
0 Tea Room	Yoga Studio
2 Food & Drink Shop	Multiplex
3 Restaurant	Pharmacy
4 Coffee Shop	Pet Store
7 Pedestrian Plaza	Toy / Game Store
8 Exhibit	Forest
14 Supermarket	Karaoke Bar
17 Kids Store	Massage Studio
19 Multiplex	Noodle House
20 Film Studio	Liquor Store
22 Coffee Shop	Shopping Mall
23 Pharmacy	Cosmetics Shop
24 Salon / Barbershop	Café
25 Supermarket	Middle Eastern Restaurant
26 Blini House	Gastropub
31 Fountain	Forest
32 Café	Pizza Place
33 Café	Gym
34 Automotive Shop	Garden
37 Park	Soccer Field
38 Clothing Store	Accessories Store
39 Surf Spot	Restaurant
40 Gym	Bar
41 College Gym	Shawarma Place
43 Supermarket	Athletics & Sports

44	Gym Pool	Sushi Restaurant
47	Boutique	Caucasian Restaurant
49	Shipping Store	Pharmacy
	7th Most Common Venue	8th Most Common Venue \
0	Fast Food Restaurant	Frozen Yogurt Shop
2	Fast Food Restaurant	Boxing Gym
3	Park	Eastern European Restaurant
4	Farmers Market	Sushi Restaurant
7	Restaurant	Rental Car Location
8	Food Court	Food & Drink Shop
14	Arts & Crafts Store	Convenience Store
17	Gourmet Shop	Modern European Restaurant
19	Bookstore	Middle Eastern Restaurant
20	Fish Market	Supermarket
22	Toy / Game Store	Seafood Restaurant
23	Pizza Place	Coffee Shop
24	Supermarket	Dance Studio
25	Rock Club	Health & Beauty Service
26	Salon / Barbershop	Salad Place
31	Food Court	Food & Drink Shop
32	Stables	Supermarket
33	Baby Store	Dumpling Restaurant
34	Sushi Restaurant	Middle Eastern Restaurant
37	Gym	Supermarket
38	Sushi Restaurant	Dance Studio
39	Eastern European Restaurant	Flower Shop
40	Farmers Market	Forest
41	Pizza Place	Middle Eastern Restaurant
43	Pharmacy	Middle Eastern Restaurant
44	Farmers Market	Grocery Store
47	Clothing Store	Convenience Store
49	Gym / Fitness Center	Food & Drink Shop
	9th Most Common Venue	10th Most Common Venue
0	Fountain	Forest
2	Fountain	Furniture / Home Store
3	Café	Soccer Field
4	Grocery Store	Salon / Barbershop
7	Gym / Fitness Center	Pet Store
8	Flower Shop	Fish Market
14	Wine Shop	Bath House
17	Fast Food Restaurant	Park
19	Blini House	Mexican Restaurant
20	Climbing Gym	Big Box Store
22	Garden Center	Bookstore
23	Food & Drink Shop	Café

24	History Museum	Food & Drink Shop
25	Pedestrian Plaza	Historic Site
26	Greek Restaurant	Shopping Mall
31	Flower Shop	Fish Market
32	Restaurant	Liquor Store
33	Fish Market	Fountain
34	Liquor Store	Flower Shop
37	Gymnastics Gym	Sushi Restaurant
38	Music Venue	Boutique
39	Fountain	Caucasian Restaurant
40	Food Court	Food & Drink Shop
41	Electronics Store	Gastropub
43	Café	Pizza Place
44	Multiplex	Café
47	Plaza	Pet Store
49	Wine Shop	Exhibit

Cluster 2

```
[54]: moscow_merged.loc[moscow_merged['Cluster Labels'] == 1, moscow_merged.
      ↪columns[[0] + list(range(5, moscow_merged.shape[1]))]]
```

```
[54]: Neighborhood 1st Most Common Venue 2nd Most Common Venue \
10                Park                Restaurant
21 -                Park    Italian Restaurant
29                Fountain                Park

        3rd Most Common Venue 4th Most Common Venue 5th Most Common Venue \
10                Scenic Lookout                Yoga Studio                Exhibit
21                Gym / Fitness Center                Hotel                Restaurant
29 Eastern European Restaurant                Yoga Studio                Fabric Shop

        6th Most Common Venue 7th Most Common Venue 8th Most Common Venue \
10                Forest                Food Court    Food & Drink Shop
21                Dog Run    Athletics & Sports                Food Court
29                Forest                Food Court    Food & Drink Shop

        9th Most Common Venue 10th Most Common Venue
10                Flower Shop                Fish Market
21    Food & Drink Shop                Flower Shop
29                Flower Shop                Fish Market
```

Cluster 3

```
[55]: moscow_merged.loc[moscow_merged['Cluster Labels'] == 2, moscow_merged.
      ↪columns[[0] + list(range(5, moscow_merged.shape[1]))]]
```

[55]:	Neighborhood	1st Most Common Venue	2nd Most Common Venue	\
1		Soccer Field	Fast Food Restaurant	
5		Fast Food Restaurant	Clothing Store	
6	-	Gym Pool	Tennis Court	
9		Japanese Restaurant	Pedestrian Plaza	
11		Pizza Place	Athletics & Sports	
12		Recreation Center	Convenience Store	
13		Food & Drink Shop	Supermarket	
16	-	Fast Food Restaurant	Soccer Stadium	
18	-	Hotel	Fast Food Restaurant	
30		Park	Plaza	
35		Soccer Field	Pub	
42		Pizza Place	Gym / Fitness Center	
45		Gym / Fitness Center	Middle Eastern Restaurant	
46		Hotel	Auto Workshop	
48		Breakfast Spot	Hotel	
	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	\
1	Scenic Lookout	Yoga Studio	Exhibit	
5	Cosmetics Shop	Big Box Store	Blini House	
6	Gym / Fitness Center	Spa	Yoga Studio	
9	Gym Pool	Forest	Shopping Mall	
11	Auto Workshop	Track	Park	
12	Food & Drink Shop	Pizza Place	Supermarket	
13	Park	Stadium	Playground	
16	Auto Workshop	Farmers Market	Gun Range	
18	Lighthouse	Auto Workshop	Park	
30	Supermarket	Asian Restaurant	Fast Food Restaurant	
35	Lake	Dry Cleaner	Auto Workshop	
42	Pool Hall	Movie Theater	Sporting Goods Shop	
45	Outdoor Sculpture	Supermarket	Park	
46	Pizza Place	Supermarket	Gym / Fitness Center	
48	Concert Hall	Grocery Store	Auto Workshop	
	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	\
1	Forest	Food Court	Food & Drink Shop	
5	Mobile Phone Shop	Sushi Restaurant	Gym / Fitness Center	
6	Forest	Food Court	Food & Drink Shop	
9	Fast Food Restaurant	Middle Eastern Restaurant	Campground	
11	Yoga Studio	Fabric Shop	Fountain	
12	Auto Workshop	Yoga Studio	Fabric Shop	
13	Pizza Place	Middle Eastern Restaurant	Gourmet Shop	
16	Market	Gym / Fitness Center	Blini House	
18	Pet Store	Yoga Studio	Fabric Shop	
30	Bar	Yoga Studio	Fountain	
35	Salon / Barbershop	Grocery Store	Exhibit	
42	Board Shop	Fast Food Restaurant	Farmers Market	

45	Track Stadium	Fabric Shop	Athletics & Sports
46	Yoga Studio	Farmers Market	Sports Club
48	Yoga Studio	Fabric Shop	Forest

	9th Most Common Venue	10th Most Common Venue
1	Flower Shop	Fish Market
5	Boutique	Hot Dog Joint
6	Flower Shop	Fish Market
9	Martial Arts Dojo	Sushi Restaurant
11	Forest	Food Court
12	Forest	Food Court
13	Exhibit	Flower Shop
16	Clothing Store	Park
18	Forest	Food Court
30	Forest	Food Court
35	Food Court	Food & Drink Shop
42	Belgian Restaurant	Caucasian Restaurant
45	Food & Drink Shop	Women's Store
46	Mobile Phone Shop	Snack Place
48	Food Court	Food & Drink Shop

Cluster 4

```
[56]: moscow_merged.loc[moscow_merged['Cluster Labels'] == 3, moscow_merged.
      ↪columns[[0] + list(range(5, moscow_merged.shape[1]))]]
```

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	\
15		Convenience Store	Pharmacy	
27		Café	Convenience Store	
28		Pharmacy	Yoga Studio	
36		Botanical Garden	Gym / Fitness Center	

	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	\
15	Gym / Fitness Center	Wine Shop	Beer Store	
27	Pharmacy	Farmers Market	Park	
28	Martial Arts Dojo	Smoke Shop	Jazz Club	
36	Pool	Convenience Store	Arts & Entertainment	

	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	\
15	Café	Grocery Store	Park	
27	Middle Eastern Restaurant	Gym / Fitness Center	Big Box Store	
28	Café	Clothing Store	Salon / Barbershop	
36	Sushi Restaurant	Fountain	Forest	

	9th Most Common Venue	10th Most Common Venue
15	Yoga Studio	Food Court
27	Hardware Store	Exhibit

28	Food Court	Food & Drink Shop
36	Food Court	Food & Drink Shop

0.10 Conclusion

The neighborhood I currently reside in is located in the Cluster 0. This cluster is characterized by a lot of shops, cafes and restaurants. So, as possible neighborhoods for relocation I could select other neighborhoods from that cluster, for example or - .

[]: