Capstone Project

The Battle of Neighborhoods

Investigation of Moscow regions

Introduction

Moscow is one of the fastest growing European cities and the largest Russian city. Currently, the population of Moscow is more than 12 million people, and the population growth exceeds 100,000 people per year. Residents of the city and official authorities are constantly faced with problems associated with urbanization. These are environmental problems, transport, overpopulation, uneven population density, tax regulation, financing, problems of recreation and entertainment.

Business Problem

One of the areas of regulation is to eliminate the asymmetry in the location of urban infrastructure centers and to promote the development of venues in certain areas where their absence is obvious. And vice versa, closing and moving venues, the number of which exceeds the necessary indicators. Significant help in managing these areas can be provided by Data Science and special methods for detecting hidden patterns. Thus, the problem is to find non-obvious patterns in the distribution of urban infrastructure centers in the context of each individual district of the city.

Data descriptiion

Using the Foursquare service allows you to get a list of objects for each of 146 districts of the city of Moscow. A subsequent analysis of the distribution of objects by district will allow us to find hidden dependencies. The predominance of objects belonging to a certain category over objects of other categories will allow us to build a certain rating of districts and combine them into groups. At the same time, segmentation and clustering (K-means) will allow us to quickly identify unknown patterns. A certain difficulty with this approach will lie in the process of verifying that a particular venue belongs to a specific area of the city. To solve this problem, we will need to describe the boundaries of each region in the form of a polygon. In the future, having calculated the location of the central point for each region, as well as the radius of the surrounding circle, it will be possible to filter out places that are included exclusively in a specific area. After additional transformations, an array of sites can be processed using the K-means algorithm. Then we can analyze the formed clusters.

Fortunately, we can get geo-points for the polygons of each district of the city on the website [http://gis-lab.info (http://gis-lab.info)](http://gis-lab.info). File format is CSV and every row is geo data for a district.

Sample data:

WKT,NAME,OKATO,OKTMO,NAME_AO,OKATO_AO,ABBREV_AO,TYPE_MO "MULTIPOLYGON (((36.8031012 55.4408329,36.8031903 55.4416007,36.8035692
55.4516224,36.812528 55.4513994,36.8274471 55.4513398,36.8333688 55.4513764,36.8338034 55.4516439,36.8345763 55.4512558,36.8348594 55.4514247,36.8349932
55.4514931,36.8358013 55.4511173,36.8360591 55.4511632,36.8461554 55.4510412,36.8602864 55.4508946,36.8649423 55.4506415,36.8608407 55.4492656,36.8582649
55.4478456,36.8582898 55.447659,36.8600008 55.4466656,36.8611076 55.4473042,36.8622805 55.4467128,36.8638768 55.4472018,36.8694408 55.4489425,36.8724625
55.4502245,36.8749845 55.4513839,36.8773319 55.453135,36.8804877 55.4548177,36.8822676 55.455771,36.8833225 55.4551478,36.8837761 55.4554817,36.8846681
55.4551548,36.8855977 55.4548773,36.8863281 55.4552539,36.8952465 55.450465,36.8875942 55.4460139,36.8818703 55.4426547,36.8912361 55.437452,36.8915818
55.4377769,36.893413 55.4368788,36.8948019 55.4377928,36.8963369 55.4389852,36.8968637 55.4392912,36.8968237 55.4387451,36.8964101 55.4380589,36.8959156
55.4369562,36.8931806 55.4285387,36.8930385 55.4281378,36.892214 55.4255804,36.8920661 55.4251213,36.8921556 55.4250135,36.8929218 55.4250569,36.8939156
55.4236071,36.8938567 55.4229336,36.8931251 55.4224349,36.8916934 55.4211121,36.8919249 55.4163998,36.892068 55.4106145,36.892487 55.4102303,36.8953439
55.4078434,36.8974602 55.4060326,36.897289 55.4057843,36.8978411 55.4053155,36.9014402 55.402228,36.9016975 55.4019726,36.9016599 55.4018918,36.9037445
55.4002463,36.9042105 55.3993939,36.9042939 55.3991925,36.9043756 55.3989367,36.9046011 55.3984985,36.9049361 55.3978212,36.905001 55.3975771,36.9050214
55.3974678,36.9051339 55.3973911,36.9051533 55.3973512,36.9052789 55.3973497,36.9052676 55.3971375,36.9056722 55.3966414,36.9058245 55.392973,36.9057008
55.3921492,36.9057503 55.3910598,36.9057594 55.3908605,36.9057755 55.3906733,36.9058681 55.3896065,36.9056813 55.3894557,36.9059927 55.3892795,36.9060527
55.3891022,36.9060885 55.3889279,36.9064487 55.3884621,36.9070191 55.3877243,36.9072097 55.3875265,36.908038 55.3866655,36.9085849 55.3860978,36.9086165
55.3858614,36.9099096 55.3858468,36.9123641 55.385847,36.913446 55.3846883,36.9145318 55.3849421,36.9148241 55.3852584,36.9154068 55.385838,36.9148314
55.3859811,36.9144995 55.3864376,36.9144027 55.3866182,36.914388 55.3873263,36.9237377 55.3857522,36.9237051 55.3855538,36.9235417 55.3845608,36.9232996
55.3833698,36.92286 55.381672,36.9243451 55.3813831,36.9256639 55.3811944,36.9268142 55.3810358,36.9285658 55.3808373,36.9294326 55.3807241,36.9333296
55.3785258,36.9368952 55.3778962,36.9372385 55.3778355,36.9375251 55.3779406,36.9378493 55.3780886,36.9380218 55.378223,36.938135 55.378259,36.9383523
55.378264,36.9384511 55.378274,36.9387604 55.3784687,36.9389054 55.3784744,36.938998 55.3784883,36.9390912 55.3785209,36.939216 55.3786952,36.9393834
55.3787832,36.9395425 55.378815,36.9397523 55.3788012,36.9398512 55.3788038,36.9399508 55.3788363,36.9400467 55.3789325,36.9401388 55.3791223,36.9402611
55.3792293,36.9404404 55.3792797,36.9406203 55.3793451,36.9407605 55.379579,36.9408409 55.3796229,36.9409539 55.3796516,36.9412515 55.3797079,36.94148

55.3797464,36.9417081 55.3797849,36.9418876 55.3798389,36.9419976 55.3799723,36.9420574 55.3801699,36.9421094 55.3803303,36.9422175 55.3804038,36.9424643
55.3805094,36.9427682 55.3805469,36.9431181 55.3805878,36.9434391 55.3807373,36.9436847 55.3807979,36.9439298 55.3808476,36.9440911 55.3809429,36.9441597
55.3810245,36.9444425 55.3812943,36.9456781 55.3813769,36.9463375 55.3814214,36.9463628 55.3819338,36.9462919 55.3837501,36.9476495 55.3840183,36.9480387
55.3840585,36.9489448 55.383842,36.950458 55.3835148,36.9505983 55.3833953,36.9523564 55.3815177,36.9540126 55.3816732,36.9566428 55.3813449,36.9536141
55.3797706,36.953013 55.3762218,36.9564788 55.3734597,36.9637211 55.3678044,36.9612597 55.3641726,36.9640825 55.360593,36.9686263 55.3613102,36.9737167
55.3621136,36.9735128 55.3603249,36.9717514 55.3578985,36.9710819 55.3571299,36.9703153 55.3564,36.9700765 55.3548495,36.9702181 55.3523218,36.9724179
55.3518163,36.9740716 55.3518001,36.9768397 55.350907,36.9920089 55.3464455,37.0056066 55.3423165,37.0109934 55.3407438,37.0055734 55.3370369,36.9992522
55.3330735,36.9983642 55.3320398,36.9949851 55.3335969,36.9939059 55.3328537,36.9922185 55.3326556,36.9921832 55.3317475,36.9919601 55.3303241,36.9911649
55.3303145,36.9909735 55.3289845,36.9912774 55.3275025,36.9897115 55.3276675,36.9888423 55.3265641,36.9886949 55.325333,36.9902152 55.3252964,36.9884623
55.3214103,36.9864415 55.3155888,36.9838644 55.3164982,36.9810291 55.3175455,36.9776942 55.3186451,36.9700473 55.3216217,36.9636827 55.3240984,36.9604066
55.3219045,36.94473 55.3293694,36.9466292 55.3306699,36.9442325 55.3316,36.9422175 55.3325707,36.9406172 55.3342645,36.9402706 55.3342711,36.9391706
55.3366222,36.9368876 55.3365924,36.9361926 55.3388126,36.9361867 55.3390532,36.9357751 55.3411599,36.936134 55.3412569,36.935493 55.3415637,36.9194896
55.3493119,36.9162405 55.3507965,36.9156603 55.3517006,36.9161441 55.3525398,36.9151419 55.3526258,36.915176 55.3528629,36.91557 55.3540344,36.9122387
55.3543218,36.909457 55.3544928,36.9069832 55.3547649,36.9047299 55.3550779,36.9029439 55.3556966,36.9031535 55.3564134,36.9040847 55.3570066,36.9016328
55.3583115,36.9014019 55.3583798,36.8879522 55.3648761,36.8721601 55.3724908,36.8560564 55.3803316,36.8629346 55.3847874,36.8482657 55.3920817,36.8549887
55.3965523,36.8618745 55.4010347,36.8455799 55.4087413,36.8342961 55.4139868,36.840391 55.4175758,36.8419298 55.4186124,36.8420879 55.4186994,36.8327628
55.4203749,36.8325854 55.4204076,36.8291041 55.4209979,36.8265397 55.4214312,36.8258145 55.4214344,36.8248774 55.4220134,36.8246559 55.4222541,36.8242945
55.4226464,36.8242744 55.4228545,36.8244641 55.4231067,36.8245509 55.4233368,36.8237061 55.4234926,36.8230683 55.4236753,36.8224488 55.4240415,36.8216071
55.4247244,36.8211988 55.4252289,36.8226024 55.4248663,36.8241 55.4262054,36.8252549 55.427376,36.8260484 55.4281616,36.8265865 55.4280778,36.8267325
55.4282417,36.8272452 55.4281283,36.8277826 55.428051,36.8282363 55.4286426,36.8247808 55.4290285,36.8242994 55.4285751,36.8239519 55.4286686,36.8219237
55.4292788,36.8191279 55.426764,36.8167848 55.429926,36.8164114 55.429923,36.8161343 55.4302984,36.812585 55.4351069,36.8113604 55.436299,36.8031012
55.4408329)),((36.9007522 55.3148568,36.9032546 55.3161103,36.9039411 55.3172336,36.9061728 55.3167452,36.909091 55.3199197,36.9096059 55.322215,36.9101209
55.3261704,36.9095202 55.3270494,36.908721 55.3277717,36.9088721 55.3280627,36.9105483 55.3312901,36.9158469 55.3303256,36.9156142 55.3299301,36.9132967
55.3259751,36.9160432 55.3254868,36.9225665 55.3245102,36.9226717 55.3238557,36.9194531 55.3187037,36.9190258 55.3186713,36.9188093 55.3186548,36.9176076
55.3167744,36.9161292 55.3169894,36.9134683 55.3171847,36.9120093 55.3167452,36.9110651 55.3135704,36.9101209 55.3130331,36.9078893 55.3117143,36.9066878
55.3112747,36.9029971 55.3128866,36.9012804 55.3139611,36.9007522 55.3148568)))",Киевский,45298555,45945000,Троицкий,45298000,Троицкий,Поселение "POLYGON
((37.4276499 55.7482092,37.4284863 55.7487502,37.429577 55.7493896,37.4305757 55.7497621,37.4316188 55.7500469,37.4318048 55.7500785,37.4320603
55.7501238,37.4322408 55.7501396,37.4326517 55.7501862,37.4329607 55.7501186,37.4351269 55.7496421,37.436979 55.7492519,37.4366052 55.7487186,37.4363477
55.7478732,37.436691 55.7471003,37.4373777 55.7464239,37.439738 55.7452161,37.4424846 55.7435735,37.4444587 55.7426072,37.4465186 55.7418583,37.4492652
55.7411336,37.4525268 55.7406021,37.4553163 55.7406021,37.4591357 55.7412061,37.4617965 55.7418583,37.4648005 55.7428971,37.4673755 55.7440325,37.4710233
55.7461099,37.4727828 55.7472693,37.4741561 55.7485012,37.4751432 55.749878,37.4757869 55.7512305,37.476173 55.7522932,37.4765165 55.7545875,37.476731
55.7559641,37.4767739 55.7569059,37.476173 55.7617112,37.476173 55.7629909,37.476388 55.7645104,37.477087 55.7655256,37.4780555 55.7666923,37.4784216
55.7670545,37.4788659 55.7673434,37.4794184 55.7675454,37.4800361 55.7677128,37.4819188 55.7679769,37.4845108 55.76817,37.4871806 55.7681941,37.4897158
55.7679596,37.4920518 55.7673608,37.4960858 55.7661054,37.4975162 55.7656864,37.4981458 55.7655019,37.501064 55.7643671,37.5029522 55.7634013,37.5042827
55.7625321,37.5058276 55.7617836,37.5116641 55.7596587,37.5139298 55.7584219,37.5145393 55.7580891,37.5162989 55.7568334,37.5176723 55.7546358,37.5181443
55.7532109,37.5185027 55.7525846,37.5202286 55.7500222,37.5203603 55.7498517,37.5220496 55.7475834,37.5225208 55.7472106,37.5214909 55.7469001,37.5189557
55.7461855,37.5175319 55.7457204,37.5166137 55.7453674,37.5156453 55.7449183,37.514874 55.7444468,37.51413 55.7439648,37.5136712 55.7436135,37.5121459
55.7420084,37.5106486 55.7403827,37.5099303 55.7395869,37.5084012 55.7381935,37.5075455 55.7377478,37.5061072 55.73699,37.5057851 55.7368281,37.4985357
55.734604,37.4964518 55.7341356,37.4896704 55.7330801,37.4892942 55.7330146,37.4863487 55.7362459,37.4843418 55.7384702,37.4834247 55.7394783,37.481536

55.7415542,37.4801615 55.7430534,37.478433 55.7425629,37.4606938 55.7375518,37.4604358 55.7372197,37.4542183 55.7358777,37.4487873 55.7348884,37.4472514 55.7350752,37.4449168 55.7348261,37.4389451 55.7368946,37.4379707 55.737174,37.4398767 55.738506,37.4399883 55.7392307,37.4397001 55.7396765,37.4372737 55.7407898,37.4357478 55.7415528,37.4330907 55.742709,37.432147 55.7446964,37.4311171 55.7452762,37.431572 55.7456868,37.4292889 55.7475564,37.4284746 55.7477249,37.4276499 55.7482092))",Филёвский Парк,45268595,45328000,Западный,45268000,ЗАО,Муниципальный округ ... ... and so on

Collecting data

We can load data from http://gis-lab.info site. Data is archived so we have to unzip the file needed.

In [178]:

```
# Load CSV data from the resource
!wget -nc 'http://gis-lab.info/data/mos-adm/mo-csv.zip'
```

File ‘mo-csv.zip’ already there; not retrieving.

In [179]:
```python
import requests
import pandas as pd
import numpy as np
from zipfile import ZipFile

zip_file = ZipFile('mo-csv.zip')
moscow_regions = pd.read_csv(zip_file.open('mo.csv'))
moscow_regions['index']=moscow_regions.index
moscow_regions
```

Out[179]:

	WKT	NAME	OKATO	OKTMO	NAME_AO	OKATO_AO	ABBREV_AO	TYPE_MO	index
0	MULTIPOLYGON (((36.8031012 55.4408329,36.80319...	Киевский	45298555	45945000	Троицкий	45298000	Троицкий	Поселение	0
1	POLYGON ((37.4276499 55.7482092,37.4284863 55....	Филёвский Парк	45268595	45328000	Западный	45268000	ЗАО	Муниципальный округ	1
2	POLYGON ((36.8035692 55.4516224,36.8045117 55....	Новофёдоровское	45298567	45954000	Троицкий	45298000	Троицкий	Поселение	2
3	POLYGON ((36.9372397 55.2413907,36.9372604 55....	Роговское	45298575	45956000	Троицкий	45298000	Троицкий	Поселение	3
4	POLYGON ((37.4395575 55.6273129,37.4401803 55....	"Мосрентген"	45297568	45953000	Новомосковский	45297000	Новомосковский	Поселение	4
...	...	...	...	...	...	...	...	...	...
141	POLYGON ((37.7998089 55.7623198,37.7998143 55....	Ивановское	45263567	45306000	Восточный	45263000	ВАО	Муниципальный округ	141
142	POLYGON ((37.8360239 55.709776,37.8361995 55.7...	Косино-Ухтомский	45263573	45308000	Восточный	45263000	ВАО	Муниципальный округ	142
143	POLYGON ((37.8404157 55.7304867,37.8406349 55....	Новокосино	45263579	45310000	Восточный	45263000	ВАО	Муниципальный округ	143
144	POLYGON ((37.9061276 55.7062585,37.9070118 55....	Некрасовка	45290574	45391000	Юго-Восточный	45290000	ЮВАО	Муниципальный округ	144
145	MULTIPOLYGON (((37.290502 55.8019897,37.295422...	Кунцево	45268562	45320000	Западный	45268000	ЗАО	Муниципальный округ	145

146 rows × 9 columns

In [ ]:

Prepare data

Now we have to prepare our data. For further correct calculations we have to extract several lists from the gathered table.

- **List of polygonal segments which constitutes city region's borders. Extracted from moscow_regions.**
- **List of coordinates of a center for every segment. Center point is actually a geometric center of a plane figure and calculated as the arithmetic mean position of all the points. I used spherical coordinates here, and this is not entirely correct, but acceptable for our case.**
- **List of polygones' radiuses. It is calculated as the farthest point from the center, belonging to the polygon.**

In [180]:

```
import re, math

# Build a list of polylines displaying the boundaries of regions
regionPolyList = []
for index,aRow in moscow_regions.iterrows():
    strPolyList = re.findall(r"\(([\d+\.\s\,]+)\)", aRow['WKT'], re.MULTILINE)
    regionSegments = []
    for strPoly in strPolyList:
        coord_pairs = re.findall(r"[\d\.]+ [\d\.]+", strPoly, re.MULTILINE)
        polyline=[]
        for pair in coord_pairs:
            polyline.append([float(pair.split(' ')[1]), float(pair.split(' ')[0])])
        regionSegments.append(polyline)
    regionPolyList.append(regionSegments)

# Calculate coordinates of centers for regions
regionCentersList = []
for segments in regionPolyList:
    segmentCenterList = []
    for segment in segments:
        c = [0.0, 0.0]
        for point in segment:
            c[0] = c[0] + point[0]
            c[1] = c[1] + point[1]
        c[0] = c[0]/len(segment)
        c[1] = c[1]/len(segment)
        segmentCenterList.append(c)
    regionCentersList.append(segmentCenterList)

# Calculate radiuses of embracing circle for each region
radiusesList = []
for regionIndex in range(0, len(regionPolyList)):
    segments = regionPolyList[regionIndex]
    centers = regionCentersList[regionIndex]
    radiuses = []
    for segIndex in range(0, len(segments)):
        segment = segments[segIndex]
        center = centers[segIndex]
        distMax=0.0
        for point in segment:
            dist = 2*6371000*math.asin(math.sqrt(math.sin(math.radians((point[0]-center[0])/2))*math.sin(math.radians((point[0]
-center[0])/2)) + \
                                                math.cos(math.radians(center[0]))*math.cos(math.radians(point[0]))* \
                                                math.sin(math.radians((center[1]-point[1])/2))*math.sin(math.radians((center[1]
-point[1])/2))))
```

```
            if dist>=distMax:
                distMax = dist
        radiuses.append(distMax)
    radiusesList.append(radiuses)
```

Now we know coordinates of center point for every city region. We can call Foursquare API sending this central point and region's radius for a complete list of target venues.

But first lets visualize region borders on the map.

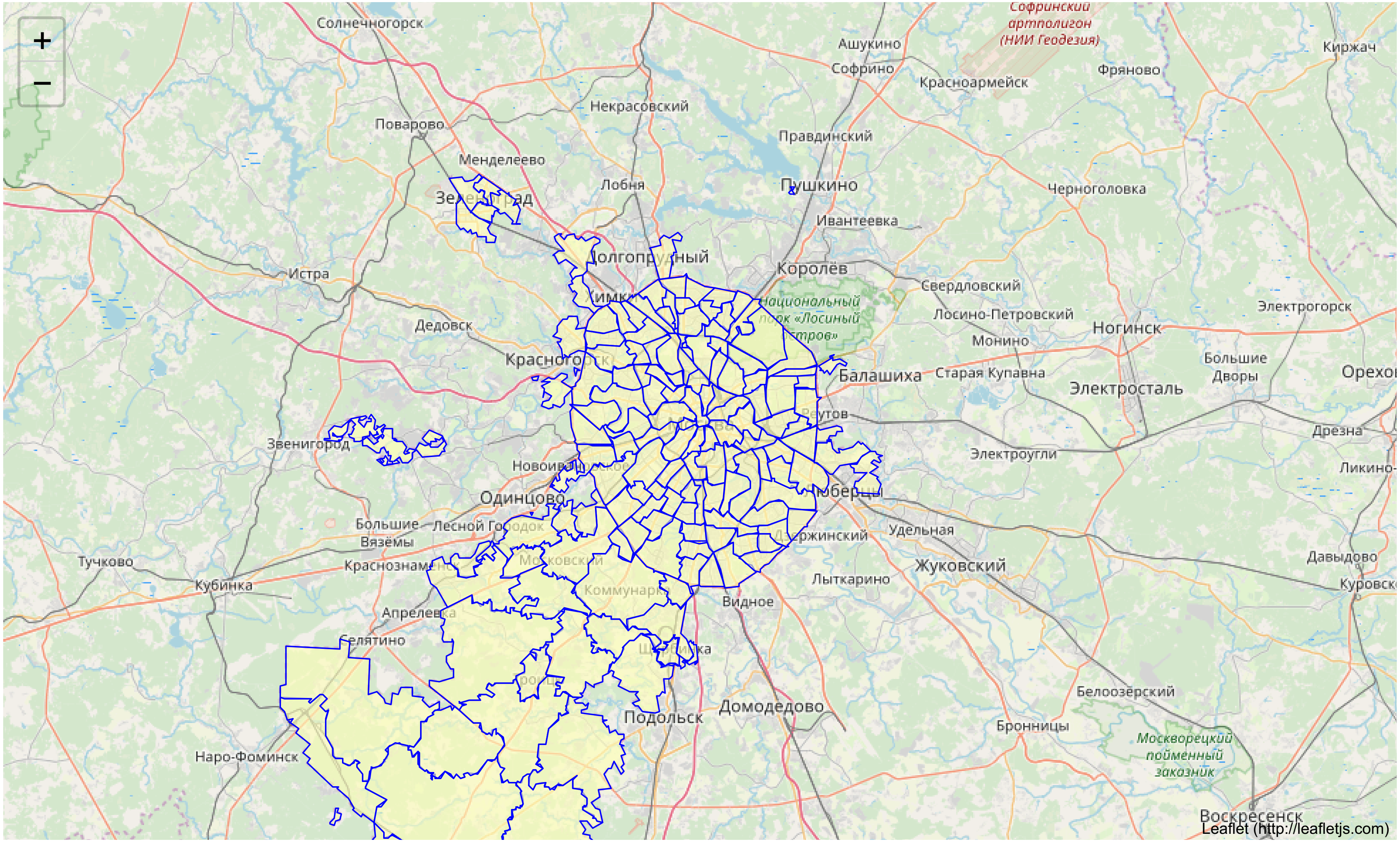
In [4]:
```
!conda install -c conda-forge folium=0.5.0 --yes
```

```
Solving environment: done


==> WARNING: A newer version of conda exists. <==
  current version: 4.5.11
  latest version: 4.7.12

Please update conda by running

    $ conda update -n base -c defaults conda



## Package Plan ##

  environment location: /home/jupyterlab/conda/envs/python

  added / updated specs:
    - folium=0.5.0


The following packages will be downloaded:

    package                    |            build
    ---------------------------|-----------------
    certifi-2019.9.11          |           py36_0         147 KB  conda-forge

The following packages will be UPDATED:

    certifi: 2019.6.16-py36_1 conda-forge --> 2019.9.11-py36_0 conda-forge


Downloading and Extracting Packages
certifi-2019.9.11    | 147 KB    | ##################################### | 100%
Preparing transaction: done
Verifying transaction: done
Executing transaction: done
```

In [182]:

```
import folium

map = folium.Map(location=[55.7532358, 37.6225412], zoom_start=8.5)

regions = folium.map.FeatureGroup()
for index in range(0, moscow_regions.shape[0]):
    regionName = moscow_regions.iloc[index, 1]
    for segmentIndex in range(0, len(regionPolyList[index])):
        segment = regionPolyList[index][segmentIndex]
        regions.add_child( folium.features.PolygonMarker(segment, color='blue', fill_color='#ffffb7') )
        #segmentCenter = regionCentersList[index][segmentIndex]
        #segmentRadius = radiusesList[index][segmentIndex]
map.add_child(regions)
map
```

Out[182]:



Even on this stage, looking at the received map, we can notice some positional patterns in the division of the city into regions. We see a pronounced central rounded region, a large appendix of the recent suburbs, protruding to the southwest, and two small enclaves in the west. In the future, it will be possible to try to compare the dependencies obtained by machine learning methods with the positioning of areas on the map.

Now we are ready to obtain venues for every city region. So lets do it.

In [183]:
```
CLIENT_ID = 'FJD2JBCQJ2XMELUIEZQXSLX01Q0QTXK1HT1WQIWTJCSPA2G1' # your Foursquare ID
CLIENT_SECRET = 'W2XREFIWN3ZANL1C4SPB5Q4G00DGT3Z3LNANDIFUYPLA0UUH' # your Foursquare Secret
VERSION = '20190923'
```

In [184]:
```
import os.path
import json
import sys

# Test whether a point belongs to a polygone
def inPolygon(x, y, segments):
    c=0
    for seg in segments:
        for i in range(1, len(seg)):
            p = seg[i]
            pp = seg[i-1]
            if (((p[1]<=y and y<pp[1]) or (pp[1]<=y and y<p[1])) and \
                (x > (pp[0] - p[0]) * (y - p[1]) / (pp[1] - p[1]) + p[0])): c = 1 - c
    return c

# Build a list of venues
def getNearbyVenues(regionIndice, names, locations, radiuses):
    LIMIT=1000
    FOLDER='res'
    symbols = (u"абвгдеёжзийклмнопрстуфхцчшщъыьэюяАБВГДЕЁЖЗИЙКЛМНОПРСТУФХЦЧШЩЪЫЬЭЮЯ",
               u"abvgdeejzijklmnoprstufhzcss_y_euaABVGDEEJZIJKLMNOPRSTUFHZCSS_Y_EUA")

    tr = {ord(a):ord(b) for a, b in zip(*symbols)}

    if not os.path.exists(FOLDER):
        os.mkdir(FOLDER)
    print('Loading\n[', end='')
    venues_list=[]
    for regIndex, name, locationsList, radiusesList in zip(regionIndice, names, locations, radiuses):
        regionNameLat = name.translate(tr)
        for location, radius, segIndex in zip(locationsList, radiusesList, range(0,len(locationsList))):
            fileName = "{}/{}_{}.json".format(FOLDER, regionNameLat, segIndex)
            # 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,
                location[0],
                location[1],
                radius,
                LIMIT)

            ioError = False
            responseCode = 0
```

```
            try:
                # first try to load local file
                if os.path.exists(fileName):
                    with open(fileName, 'r') as readfile:
                        results = json.load(readfile)
                        print('.', end='')
                        #print("Loaded locally {} - {} [size:{}]".format(regIndex, fileName, sys.getsizeof(results)))
                else:
                    # make the GET request
                    resp = requests.get(url)
                    responseCode = resp.status_code
                    if resp.status_code == 200:
                        jsonResponse = resp.json()
                        results = jsonResponse['response']['groups'][0]['items']
                        print('.', end='')
                        with open(fileName, 'w') as writefile:
                            json.dump(results, writefile)
                            print ("Save to local file {} - {}".format(regIndex, fileName))
                    else:
                        ioError = True
                        print("IOError. Response code {}".format(resp.status_code))
            except Exception as ex:
                print('Error {} calling {}'.format(ex, url))
                if os.path.exists(fileName):
                    with open(fileName, 'r') as readfile:
                        results = json.load(readfile)
                        #print ("Loaded locally due to an error {} - {}".format(regIndex, fileName))
                        print('.', end='')

            # return only relevant information for each nearby venue
            if not ioError:
                venues_list.append([(
                    int(round(regIndex)),
                    regionNameLat,
                    location[0],
                    location[1],
                    v['venue']['name'].translate(tr),
                    v['venue']['location']['lat'],
                    v['venue']['location']['lng'],
                    v['venue']['categories'][0]['name']) for v in results])

    print(']')

    nearby_venues = pd.DataFrame([item for venue_list in venues_list for item in venue_list])
    if (nearby_venues.shape[0] > 1):
        nearby_venues.columns = ['Region index',
```

```
                     'Neighbourhood',
                     'Neighbourhood Latitude',
                     'Neighbourhood Longitude',
                     'Venue',
                     'Venue Latitude',
                     'Venue Longitude',
                     'Venue Category']

    return(nearby_venues)
```

Load venues for 146 regions

In [185]:
```
venues = getNearbyVenues(moscow_regions['index'], moscow_regions['NAME'], regionCentersList, radiusesList)
```

```
Loading
[..................................................................................................
```

In [186]:
```
print(venues.shape)
venues.head()
```

```
(13493, 8)
```

Out[186]:

	Region index	Neighbourhood	Neighbourhood Latitude	Neighbourhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	0	Kievskij	55.391999	36.907018	Mogutovo	55.354818	36.890202	Campground
1	0	Kievskij	55.391999	36.907018	Park «Sosny»	55.467084	36.935202	Park
2	0	Kievskij	55.391999	36.907018	Akovlevskij park	55.467536	36.934636	Park
3	0	Kievskij	55.391999	36.907018	Derevna Aleksandrovka	55.406133	36.794478	Farm
4	0	Kievskij	55.391999	36.907018	Ponciki	55.384648	36.783685	Donut Shop

Obviously, some objects are included twice or more times, because they are situated at the intersection of circles built around neighboring centers. Therefore, we must clear out the data set.

In [187]:
```python
def removeOutcasts(row):
    if inPolygon(row[5], row[6], regionPolyList[row[0]]):
        return row
    return pd.Series()

venuesClean = venues.apply(removeOutcasts, axis=1)
venuesClean = venuesClean[venuesClean['Venue Latitude'].apply(lambda x: not math.isnan(x))]
venuesClean
```

Out[187]:

	Neighbourhood	Neighbourhood Latitude	Neighbourhood Longitude	Region index	Venue	Venue Category	Venue Latitude	Venue Longitude
5	Kievskij	55.391999	36.907018	0.0	J/d stanzia Bekasovo-1	Train Station	55.429574	36.840377
12	Kievskij	55.320082	36.911095	0.0	Machixino Style	Castle	55.322899	36.912227
13	Filevskij Park	55.748695	37.473326	1.0	PandaPark Fili	Athletics & Sports	55.750919	37.478342
14	Filevskij Park	55.748695	37.473326	1.0	PKiO «Fili»	Park	55.747953	37.484884
16	Filevskij Park	55.748695	37.473326	1.0	Filevskaa naberejnaa	Pedestrian Plaza	55.740289	37.456460
...	...	...	...	...	...	...	...	...
13471	Kunzevo	55.785919	37.330884	145.0	Ostanovka «Rublevo»	Bus Stop	55.786702	37.355136
13473	Kunzevo	55.785919	37.330884	145.0	Asna	Pharmacy	55.782227	37.358974
13482	Kunzevo	55.808512	37.376903	145.0	Biznes poselok Rublevo-Makininskij	Hostel	55.808627	37.378484
13485	Kunzevo	55.808512	37.376903	145.0	MINI Garage	Auto Workshop	55.804565	37.375400
13486	Kunzevo	55.808512	37.376903	145.0	Appart-otel_ Rublevo-Makinino	Hotel	55.810122	37.378312

4487 rows × 8 columns

Lets look at the top of venues list sorted by count.

In [188]:
```
top20v = pd.DataFrame( venuesClean.groupby(by='Venue Category').count().sort_values(by=['Venue'],ascending=False))
top20v.drop(columns = ['Neighbourhood','Neighbourhood Latitude','Neighbourhood Longitude', 'Venue','Venue Latitude','Venue Longitude'], inplace=True)
top20v.columns = ['Count']
top20v.head(7)
```

Out[188]:

	Count
Venue Category	
Park	263
Gym / Fitness Center	230
Coffee Shop	173
Supermarket	137
Cosmetics Shop	106
Pizza Place	101
Health Food Store	95

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

In [189]:
```
print('There are {} uniques categories.'.format(len(venuesClean['Venue Category'].unique())))
```

```
There are 339 uniques categories.
```

In [190]:
```
venues_onehot = pd.get_dummies(venuesClean[['Venue Category']], prefix="", prefix_sep="")
venues_onehot['Neighbourhood'] = venuesClean['Neighbourhood']
fixed_columns = [venues_onehot.columns[-1]] + list(venues_onehot.columns[:-1])
venues_onehot = venues_onehot[fixed_columns]
venues_onehot['Region index'] = venuesClean['Region index']
fixed_columns = [venues_onehot.columns[-1]] + list(venues_onehot.columns[:-1])
venues_onehot = venues_onehot[fixed_columns]
```

In [191]:
```
venues_onehot
```

Out[191]:

	Region index	Neighbourhood	ATM	Accessories Store	Adult Boutique	Airport	Airport Lounge	Airport Service	American Restaurant	Aquarium	...	Water Park	Waterfall	Waterfront	Wine Bar	Wine Shop
5	0.0	Kievskij	0	0	0	0	0	0	0	0	...	0	0	0	0	0
12	0.0	Kievskij	0	0	0	0	0	0	0	0	...	0	0	0	0	0
13	1.0	Filevskij Park	0	0	0	0	0	0	0	0	...	0	0	0	0	0
14	1.0	Filevskij Park	0	0	0	0	0	0	0	0	...	0	0	0	0	0
16	1.0	Filevskij Park	0	0	0	0	0	0	0	0	...	0	0	0	0	0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
13471	145.0	Kunzevo	0	0	0	0	0	0	0	0	...	0	0	0	0	0
13473	145.0	Kunzevo	0	0	0	0	0	0	0	0	...	0	0	0	0	0
13482	145.0	Kunzevo	0	0	0	0	0	0	0	0	...	0	0	0	0	0
13485	145.0	Kunzevo	0	0	0	0	0	0	0	0	...	0	0	0	0	0
13486	145.0	Kunzevo	0	0	0	0	0	0	0	0	...	0	0	0	0	0

4487 rows × 341 columns

In [192]:
```
venues_onehot.shape
```

Out[192]:
```
(4487, 341)
```

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

In [193]:
```
venues_grouped = venues_onehot.groupby(['Region index','Neighbourhood']).mean().reset_index()
```

In [194]:
```
venues_grouped
```

Out[194]:

	Region index	Neighbourhood	ATM	Accessories Store	Adult Boutique	Airport	Airport Lounge	Airport Service	American Restaurant	Aquarium	...	Water Park	Waterfall	Waterfront	Wine Bar	Wine Shop
0	0.0	Kievskij	0.0	0.0	0.00000	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.000000
1	1.0	Filevskij Park	0.0	0.0	0.03125	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.000000
2	2.0	Novofedorovskoe	0.0	0.0	0.00000	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.000000
3	3.0	Rogovskoe	0.0	0.0	0.00000	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.000000
4	4.0	"Mosrentgen"	0.0	0.0	0.00000	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.000000
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
141	141.0	Ivanovskoe	0.0	0.0	0.00000	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.000000
142	142.0	Kosino-Uhtomskij	0.0	0.0	0.00000	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.000000
143	143.0	Novokosino	0.0	0.0	0.00000	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.033333
144	144.0	Nekrasovka	0.0	0.0	0.00000	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.000000
145	145.0	Kunzevo	0.0	0.0	0.00000	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.016667

146 rows × 341 columns

Let's check top5 categories in each venue

In [195]:

```python
num_top_venues = 5

for hood in venues_grouped['Neighbourhood']:
    print("----"+hood+"----")
    temp = venues_grouped[venues_grouped['Neighbourhood'] == hood].T.reset_index()
    temp.columns = ['venue','freq']
    temp = temp.iloc[2:]
    temp['freq'] = temp['freq'].astype(float)
    temp = temp.round({'freq': 2})
    print(temp.sort_values('freq', ascending=False).reset_index(drop=True).head(num_top_venues))
    print('\n')
```

```
----Kievskij----
           venue  freq
0  Train Station   0.5
1         Castle   0.5
2            ATM   0.0
3         Palace   0.0
4   Perfume Shop   0.0


----Filevskij Park----
                  venue  freq
0  Gym / Fitness Center  0.12
1                  Park  0.09
2      Pedestrian Plaza  0.06
3         Cosmetics Shop  0.06
4    Athletics & Sports  0.03


----Novofedorovskoe----
                 venue  freq
0                 Park  0.67
1            Hotel Bar  0.33
2                  ATM  0.00
3            Pet Store  0.00
4  Peruvian Restaurant  0.00


----Rogovskoe----
                 venue  freq
0        Garden Center   0.5
1    Convenience Store   0.5
2                  ATM   0.0
3               Palace   0.0
4  Peruvian Restaurant   0.0


----"Mosrentgen"----
                   venue  freq
0   Gym / Fitness Center   0.5
1                   Park   0.5
2           Outlet Store   0.0
3           Perfume Shop   0.0
4  Performing Arts Venue   0.0


----Voronovskoe----
```

```
               venue  freq
0         Restaurant  0.11
1               Lake  0.11
2  Convenience Store  0.11
3             Resort  0.11
4         Steakhouse  0.06


----Mihajlovo-Arzevskoe----
               venue  freq
0  Recreation Center  0.11
1      Historic Site  0.11
2         Restaurant  0.11
3            Stables  0.11
4           Pharmacy  0.11


----Maruskinskoe----
            venue  freq
0  Scenic Lookout  0.17
1     Gas Station  0.08
2   Historic Site  0.08
3   Train Station  0.08
4   Shopping Mall  0.08


----Pervomajskoe----
                        venue  freq
0          Italian Restaurant  0.15
1                Fishing Spot  0.08
2  Construction & Landscaping  0.08
3                 Supermarket  0.08
4                         Spa  0.08


----Matuskino----
               venue  freq
0        Supermarket  0.09
1  Electronics Store  0.06
2              Plaza  0.06
3            Brewery  0.06
4       Burger Joint  0.06


----Vnukovo----
               venue  freq
```

```
0    Airport Service  0.17
1     Airport Lounge  0.13
2        Coffee Shop  0.09
3     Scenic Lookout  0.09
4  Mobile Phone Shop  0.09


----Savelki----
         venue  freq
0   Restaurant  0.09
1    Racetrack  0.09
2         Park  0.09
3   Donut Shop  0.05
4  Pizza Place  0.05


----Vnukovskoe----
               venue  freq
0             Museum  0.16
1    Bed & Breakfast  0.11
2               Park  0.05
3  Convenience Store  0.05
4        Pizza Place  0.05


----Silino----
               venue  freq
0           Pharmacy  0.12
1  Mobile Phone Shop  0.12
2       Dessert Shop  0.06
3  Electronics Store  0.06
4     History Museum  0.06


----Kokoskino----
                 venue  freq
0        Train Station  0.25
1               Forest  0.25
2                 Farm  0.25
3    Convenience Store  0.25
4  Peruvian Restaurant  0.00


----Krukovo----
         venue  freq
0         Park  0.09
```

```
1  Pizza Place  0.09
2         Café  0.03
3    Gift Shop  0.03
4          Bar  0.03


----Sukino----
                 venue  freq
0                 Park  0.10
1    Salon / Barbershop  0.06
2  Gym / Fitness Center  0.04
3                  Spa  0.04
4       Clothing Store  0.04


----Krasnopahorskoe----
             venue  freq
0             Park  0.15
1          Exhibit  0.08
2              Zoo  0.08
3  Basketball Court  0.08
4          Stables  0.08


----Nagatinskij Zaton----
                 venue  freq
0  Gym / Fitness Center  0.11
1  Other Great Outdoors  0.05
2    Athletics & Sports  0.05
3               Garden  0.05
4               Palace  0.05


----Staroe Krukovo----
                 venue  freq
0  Gym / Fitness Center  0.18
1                 Park  0.12
2         Concert Hall  0.12
3    Athletics & Sports  0.06
4       Cosmetics Shop  0.06


----Klenovskoe----
        venue  freq
0  Restaurant   0.2
1        Lake   0.2
```

```
2    Bus Stop   0.2
3   Gun Range   0.2
4  Theme Park   0.2


----Dmitrovskij----
                        venue  freq
0                 Supermarket  0.21
1                        Café  0.10
2                  Restaurant  0.07
3                        Park  0.07
4  Eastern European Restaurant  0.05


----Filimonkovskoe----
                 venue  freq
0               Resort  0.17
1   Housing Development  0.17
2          Music Venue  0.17
3  Gym / Fitness Center  0.17
4          Supermarket  0.17


----Troizk----
              venue  freq
0       Supermarket  0.08
1        Restaurant  0.08
2              Café  0.06
3              Park  0.06
4  Electronics Store  0.04


----Teplyj Stan----
              venue  freq
0  Health Food Store  0.09
1        Coffee Shop  0.06
2      Cosmetics Shop  0.06
3              Park  0.06
4       Supermarket  0.06


----Sapovskoe----
              venue  freq
0  Convenience Store  0.17
1              Café  0.17
2     Historic Site  0.17
```

```
3      Grocery Store  0.17
4               Park  0.17


----Moskovskij----
                venue  freq
0         Coffee Shop  0.13
1      Clothing Store  0.10
2  Sporting Goods Shop  0.06
3        Gourmet Shop  0.06
4       Big Box Store  0.06


----Desenovskoe----
                      venue  freq
0             Soccer Field  0.11
1                   Resort  0.06
2            Grocery Store  0.06
3  Middle Eastern Restaurant  0.06
4        Convenience Store  0.06


----Hovrino----
                 venue  freq
0  Gym / Fitness Center  0.11
1     Sushi Restaurant  0.05
2             Pharmacy  0.05
3                 Park  0.05
4          Pizza Place  0.05


----Lomonosovskij----
                venue  freq
0   Convenience Store  0.12
1         Fish Market  0.08
2   Food & Drink Shop  0.08
3  Italian Restaurant  0.08
4        Dance Studio  0.08


----Mojajskij----
                 venue  freq
0         Auto Workshop  0.17
1  Gym / Fitness Center  0.13
2       Motorcycle Shop  0.09
3                  Park  0.09
```

```
4    Salon / Barbershop  0.04


----Novo-Peredelkino----
                  venue  freq
0           Supermarket  0.14
1  Gym / Fitness Center  0.07
2          Tennis Court  0.05
3     Mobile Phone Shop  0.05
4         Grocery Store  0.05


----Strogino----
                  venue  freq
0                  Park  0.21
1    Athletics & Sports  0.06
2             Surf Spot  0.06
3  Gym / Fitness Center  0.06
4             Pet Store  0.06


----Moljaninovskij----
                         venue  freq
0                        Hotel  0.33
1  Eastern European Restaurant  0.17
2                   Restaurant  0.17
3                  Flea Market  0.17
4           Athletics & Sports  0.17


----Mitino----
                  venue  freq
0      Pedestrian Plaza  0.08
1                   Gym  0.05
2          Yoga Studio  0.05
3                  Café  0.05
4  Gym / Fitness Center  0.05


----Kurkino----
                  venue  freq
0                  Park  0.21
1                  Café  0.12
2  Gym / Fitness Center  0.08
3                Bakery  0.08
4               Brewery  0.04
```

```
----Krylatskoe----
                 venue  freq
0          Coffee Shop  0.07
1  Gym / Fitness Center  0.07
2                 Park  0.05
3           Kids Store  0.03
4      Motorcycle Shop  0.03


----Solnzevo----
                 venue  freq
0                 Park  0.11
1                 Café  0.11
2                Beach  0.06
3                  Pub  0.06
4  Caucasian Restaurant  0.06


----Sosenskoe----
                 venue  freq
0         Gourmet Shop  0.09
1       Cosmetics Shop  0.06
2  Gym / Fitness Center  0.06
3        Shopping Mall  0.06
4       Clothing Store  0.06


----Voskresenskoe----
           venue  freq
0      Rest Area  0.18
1          Hotel  0.18
2           Pool  0.09
3  Clothing Store  0.09
4         Resort  0.09


----Golovinskij----
              venue  freq
0       Coffee Shop  0.12
1  Health Food Store  0.09
2       Supermarket  0.06
3         Pet Store  0.06
4         Waterfall  0.06
```

```
----Ujnoe Tusino----
                 venue  freq
0          Coffee Shop  0.11
1          Pizza Place  0.11
2  Gym / Fitness Center  0.07
3        Cosmetics Shop  0.07
4                  Café  0.04


----Severnoe Tusino----
           venue  freq
0           Park  0.09
1  Cosmetics Shop  0.05
2      Kids Store  0.05
3           Café  0.05
4       Pharmacy  0.05


----Ceremuski----
                 venue  freq
0                  Café  0.11
1  Gym / Fitness Center  0.07
2  Caucasian Restaurant  0.04
3        Cosmetics Shop  0.04
4    Salon / Barbershop  0.04


----Pokrovskoe-Stresnevo----
                venue  freq
0                 Park  0.20
1            Surf Spot  0.20
2          Zoo Exhibit  0.07
3  Sporting Goods Shop  0.07
4                Beach  0.07


----Horosevo-Mnevniki----
                 venue  freq
0  Gym / Fitness Center  0.10
1                  Park  0.08
2          Gourmet Shop  0.05
3    Salon / Barbershop  0.05
4     Food & Drink Shop  0.05
```

```
----Ocakovo-Matveevskoe----
           venue  freq
0  Auto Workshop  0.29
1  Big Box Store  0.14
2     Bath House  0.14
3          Hotel  0.14
4           Café  0.14


----Troparevo-Nikulino----
                  venue  freq
0  Gym / Fitness Center  0.14
1                   Spa  0.07
2           Coffee Shop  0.07
3        Cosmetics Shop  0.05
4                  Park  0.05


----Levoberejnyj----
                       venue  freq
0                Coffee Shop  0.11
1                       Park  0.11
2                       Café  0.08
3         Sporting Goods Shop  0.05
4  Middle Eastern Restaurant  0.05


----Fili-Davydkovo----
                       venue  freq
0                       Park  0.09
1  Middle Eastern Restaurant  0.06
2              Shopping Mall  0.06
3             Clothing Store  0.06
4                   Boutique  0.06


----Obrucevskij----
                  venue  freq
0  Gym / Fitness Center  0.10
1      Asian Restaurant  0.05
2     Frozen Yogurt Shop  0.05
3                   Gym  0.05
4                  Park  0.05


----Razanovskoe----
```

```
                 venue  freq
0              Stables  0.15
1                 Café  0.15
2           Restaurant  0.08
3    Convenience Store  0.08
4                  Gym  0.08


----Ramenki----
         venue  freq
0         Park  0.15
1  Coffee Shop  0.10
2          Spa  0.05
3   Restaurant  0.05
4  Yoga Studio  0.05


----Vojkovskij----
                  venue  freq
0  Gym / Fitness Center  0.09
1           Coffee Shop  0.07
2                  Park  0.04
3        Cosmetics Shop  0.04
4        Clothing Store  0.04


----Sokol----
          venue  freq
0           Spa  0.06
1      Beer Bar  0.06
2          Park  0.04
3     Pet Store  0.04
4  Dance Studio  0.04


----Zapadnoe Degunino----
               venue  freq
0     Cosmetics Shop  0.33
1         Restaurant  0.17
2  Health Food Store  0.17
3      Auto Workshop  0.17
4               Pool  0.17


----Prospekt Vernadskogo----
                  venue  freq
```

```
0  Gym / Fitness Center  0.08
1          Gourmet Shop  0.08
2                  Park  0.08
3           Supermarket  0.05
4           Coffee Shop  0.05


----Ujnoe Butovo----
                venue  freq
0                Park  0.22
1  Italian Restaurant  0.05
2    Sushi Restaurant  0.05
3           Pet Store  0.05
4        Skating Rink  0.03


----Asenevo----
                  venue  freq
0                  Park  0.12
1  Gym / Fitness Center  0.12
2        Science Museum  0.08
3         Movie Theater  0.04
4            Beer Store  0.04


----Dorogomilovo----
               venue  freq
0        Coffee Shop  0.11
1               Park  0.06
2      Auto Workshop  0.04
3  Food & Drink Shop  0.04
4                Spa  0.04


----Kon_kovo----
                 venue  freq
0          Coffee Shop  0.11
1      Cosmetics Shop  0.08
2                 Park  0.08
3                 Café  0.08
4  Dumpling Restaurant  0.05


----Horosevskij----
               venue  freq
0        Coffee Shop  0.10
```

```
1     Clothing Store  0.04
2               Park  0.04
3  Health Food Store  0.04
4  Martial Arts Dojo  0.03


----Begovoj----
                  venue  freq
0  Gym / Fitness Center  0.13
1          Dance Studio  0.11
2           Coffee Shop  0.09
3           Pizza Place  0.07
4    Seafood Restaurant  0.04


----Koptevo----
                  venue  freq
0            Beer Store  0.11
1   Caucasian Restaurant  0.11
2   Gym / Fitness Center  0.11
3                Exhibit  0.06
4  Vietnamese Restaurant  0.06


----Serbinka----
                  venue  freq
0           Pizza Place  0.12
1           Supermarket  0.12
2  Gym / Fitness Center  0.12
3     Convenience Store  0.08
4  Fast Food Restaurant  0.08


----Aeroport----
               venue  freq
0        Coffee Shop  0.14
1  Health Food Store  0.09
2          Pet Store  0.09
3               Park  0.09
4     History Museum  0.05


----Presnenskij----
            venue  freq
0   Cocktail Bar  0.10
1  Jewelry Store  0.07
```

```
2    Coffee Shop  0.07
3           Park  0.07
4   Concert Hall  0.07


----Severnyj----
                 venue  freq
0       Clothing Store  0.16
1  Sporting Goods Shop  0.11
2            Nightclub  0.05
3          Coffee Shop  0.05
4        Big Box Store  0.05


----Beskudnikovskij----
                         venue  freq
0          Japanese Restaurant  0.05
1                    Bookstore  0.05
2         Caucasian Restaurant  0.05
3               Cosmetics Shop  0.05
4  Eastern European Restaurant  0.05


----Gagarinskij----
                  venue  freq
0    Italian Restaurant  0.11
1           Coffee Shop  0.09
2                  Park  0.07
3         Gourmet Shop  0.04
4  Caucasian Restaurant  0.02


----Timirazevskij----
               venue  freq
0               Park  0.08
1  Health Food Store  0.08
2      Auto Workshop  0.05
3        Yoga Studio  0.05
4     Cosmetics Shop  0.05


----Severnoe Butovo----
                  venue  freq
0                  Park  0.14
1  Gym / Fitness Center  0.10
2        Cosmetics Shop  0.10
```

```
3              Brewery  0.05
4                  Pub  0.05


----Lianozovo----
               venue  freq
0               Park  0.12
1  Salon / Barbershop  0.08
2            Stables  0.04
3        Pizza Place  0.04
4        Flower Shop  0.04


----Hamovniki----
           venue  freq
0    Yoga Studio  0.11
1           Park  0.11
2  Scenic Lookout  0.11
3      Monastery  0.05
4    Art Gallery  0.05


----Vostocnoe Degunino----
                 venue  freq
0                  Gym  0.14
1          Supermarket  0.14
2      Toy / Game Store  0.09
3  Gym / Fitness Center  0.09
4   Japanese Restaurant  0.05


----Savelovskij----
                      venue  freq
0               Coffee Shop  0.10
1      Gym / Fitness Center  0.10
2         Health Food Store  0.07
3  Middle Eastern Restaurant  0.03
4               Climbing Gym  0.03


----Akademiceskij----
                 venue  freq
0    Health Food Store  0.08
1         Dance Studio  0.06
2  Gym / Fitness Center  0.04
3       Cosmetics Shop  0.04
```

```
4             Pet Store  0.04


----Zuzino----
                  venue  freq
0  Gym / Fitness Center  0.09
1                  Park  0.09
2                Arcade  0.06
3            Restaurant  0.06
4           Supermarket  0.06


----Altuf_evskij----
               venue  freq
0        Supermarket  0.20
1      Auto Workshop  0.20
2  Convenience Store  0.05
3           Bus Stop  0.05
4      Grocery Store  0.05


----Marfino----
                  venue  freq
0     Convenience Store  0.15
1          Liquor Store  0.08
2  Caucasian Restaurant  0.08
3                Bakery  0.08
4           Flower Shop  0.08


----Certanovo Zentral_noe----
                  venue  freq
0        Clothing Store  0.13
1  Gym / Fitness Center  0.06
2           Coffee Shop  0.06
3           Supermarket  0.04
4         Cosmetics Shop  0.04


----Otradnoe----
                  venue  freq
0                  Park  0.12
1  Gym / Fitness Center  0.07
2            Kids Store  0.07
3         Cosmetics Shop  0.07
4      Toy / Game Store  0.05
```

```
----Arbat----
            venue  freq
0          Museum  0.09
1     Coffee Shop  0.07
2           Hotel  0.07
3           Plaza  0.07
4  Massage Studio  0.04


----Certanovo Ujnoe----
                       venue  freq
0                Supermarket  0.10
1                Pizza Place  0.10
2               Dance Studio  0.07
3              Auto Workshop  0.07
4  Middle Eastern Restaurant  0.03


----Butyrskij----
                venue  freq
0  Photography Studio  0.09
1         Coffee Shop  0.09
2      Science Museum  0.04
3         Supermarket  0.04
4      Clothing Store  0.04


----Tverskoj----
                venue  freq
0         Yoga Studio  0.06
1               Plaza  0.06
2               Hotel  0.04
3  Russian Restaurant  0.04
4         Coffee Shop  0.04


----Certanovo Severnoe----
                  venue  freq
0  Gym / Fitness Center  0.10
1              Gym Pool  0.08
2                  Park  0.08
3             Multiplex  0.05
4             Pet Store  0.05
```

```
----Akimanka----
                  venue  freq
0  Gym / Fitness Center  0.08
1           Art Gallery  0.08
2             Bookstore  0.05
3            Art Museum  0.05
4                  Park  0.05


----Kotlovka----
                    venue  freq
0              Beer Store  0.09
1  Furniture / Home Store  0.09
2      Photography Studio  0.09
3             Pizza Place  0.09
4      Athletics & Sports  0.05


----Ostankinskij----
              venue  freq
0              Park  0.11
1          Fountain  0.09
2    Scenic Lookout  0.07
3        Playground  0.07
4  Botanical Garden  0.04


----Donskoj----
                  venue  freq
0  Gym / Fitness Center  0.05
1            Board Shop  0.05
2          Camera Store  0.05
3                  Park  0.05
4  Caucasian Restaurant  0.02


----Bibirevo----
                  venue  freq
0                  Park  0.12
1  Gym / Fitness Center  0.08
2     Health Food Store  0.06
3           Pizza Place  0.06
4                   Bar  0.06
```

```
----Birulevo Zapadnoe----
            venue  freq
0     Supermarket  0.19
1   Big Box Store  0.12
2            Pool  0.12
3            Park  0.06
4  Cosmetics Shop  0.06


----Mar_ina Rosa----
                  venue  freq
0          Dance Studio  0.12
1               Theater  0.08
2     Health Food Store  0.08
3  Gym / Fitness Center  0.04
4     Food & Drink Shop  0.04


----Nagornyj----
               venue  freq
0           Ski Area  0.10
1       Gourmet Shop  0.10
2  Health Food Store  0.05
3     Cosmetics Shop  0.05
4      Auto Workshop  0.05


----Sviblovo----
                 venue  freq
0       Clothing Store  0.09
1       Cosmetics Shop  0.07
2        Grocery Store  0.04
3  Sporting Goods Shop  0.04
4       Lingerie Store  0.04


----Danilovskij----
                   venue  freq
0    Sporting Goods Shop  0.11
1   Gym / Fitness Center  0.11
2      Health Food Store  0.07
3          Auto Workshop  0.04
4  Vietnamese Restaurant  0.04


----Mesanskij----
```

```
                venue  freq
0               Hotel  0.08
1  Sauna / Steam Room  0.04
2                 Bar  0.04
3    Toy / Game Store  0.04
4          Hobby Shop  0.04


----Ujnoe Medvedkovo----
                venue  freq
0         Supermarket  0.12
1         Pizza Place  0.08
2   Convenience Store  0.06
3            Beer Bar  0.04
4  Athletics & Sports  0.04


----Zamoskvorec_e----
          venue  freq
0   Coffee Shop  0.10
1   Yoga Studio  0.05
2        Arcade  0.05
3         Hotel  0.05
4  Concert Hall  0.05


----Severnoe Medvedkovo----
           venue  freq
0  Auto Workshop  0.13
1       Pharmacy  0.06
2           Park  0.06
3         Bakery  0.06
4           Café  0.06


----Nagatino-Sadovniki----
                  venue  freq
0  Gym / Fitness Center  0.10
1  Caucasian Restaurant  0.08
2              Beer Bar  0.05
3                Arcade  0.05
4           Coffee Shop  0.05


----Moskvorec_e-Saburovo----
            venue  freq
```

```
0     Coffee Shop  0.17
1      Bike Trail  0.08
2           Hotel  0.08
3  Clothing Store  0.08
4    Gourmet Shop  0.08


----Zarizyno----
               venue  freq
0               Park  0.09
1         Playground  0.06
2     Cosmetics Shop  0.06
3                Gym  0.06
4  Convenience Store  0.06


----Basmannyj----
                  venue  freq
0           Coffee Shop  0.07
1          Dance Studio  0.04
2           Yoga Studio  0.04
3  Gym / Fitness Center  0.04
4               Theater  0.04


----Krasnosel_skij----
           venue  freq
0    Coffee Shop  0.17
1     Hookah Bar  0.17
2  Jewelry Store  0.08
3    Yoga Studio  0.08
4     Restaurant  0.08


----Rostokino----
                       venue  freq
0                       Park  0.14
1       Gym / Fitness Center  0.10
2                    Theater  0.07
3  Middle Eastern Restaurant  0.07
4                    Stadium  0.03


----Taganskij----
                  venue  freq
0  Gym / Fitness Center  0.05
```

```
1               Arcade  0.05
2               Bakery  0.04
3                  Bar  0.04
4          Coffee Shop  0.04


----Alekseevskij----
                 venue  freq
0       Auto Workshop  0.10
1  Convenience Store  0.07
2              Arcade  0.05
3  Mobile Phone Shop  0.05
4         Coffee Shop  0.05


----Sokol_niki----
                  venue  freq
0          Soccer Field  0.10
1                 Trail  0.04
2    Sporting Goods Shop  0.04
3            Playground  0.04
4         Auto Workshop  0.04


----Birulevo Vostocnoe----
                   venue  freq
0  Gym / Fitness Center  0.10
1                   Park  0.10
2            Supermarket  0.10
3                Theater  0.06
4   Japanese Restaurant  0.03


----Babuskinskij----
                   venue  freq
0                    Gym  0.14
1                   Park  0.14
2            Supermarket  0.10
3  Gym / Fitness Center  0.07
4               Pharmacy  0.07


----Ujnoportovyj----
                   venue  freq
0  Gym / Fitness Center  0.06
1    Photography Studio  0.06
```

```
2                Circus  0.04
3      Recording Studio  0.04
4             Nightclub  0.04


----Aroslavskij----
                   venue  freq
0                   Park  0.12
1  Vietnamese Restaurant  0.12
2          Auto Workshop  0.08
3   Gym / Fitness Center  0.08
4    Japanese Restaurant  0.04


----Pecatniki----
                  venue  freq
0                   Bar  0.17
1                  Park  0.08
2  Gym / Fitness Center  0.08
3        Clothing Store  0.08
4       Motorcycle Shop  0.08


----Bogorodskoe----
                  venue  freq
0  Gym / Fitness Center  0.14
1     Convenience Store  0.09
2               Dog Run  0.05
3                 Track  0.05
4           Coffee Shop  0.05


----Metrogorodok----
           venue  freq
0     Playground  0.17
1          Trail  0.17
2        Stables  0.17
3  National Park  0.17
4    Candy Store  0.17


----Lefortovo----
                  venue  freq
0                  Park  0.09
1  Gym / Fitness Center  0.09
2                  Café  0.06
```

```
3             Bath House  0.06
4            Coffee Shop  0.06


----Orehovo-Borisovo Severnoe----
                   venue  freq
0            Coffee Shop  0.09
1             Restaurant  0.06
2  Gym / Fitness Center  0.06
3         Clothing Store  0.04
4          Historic Site  0.04


----Losinoostrovskij----
                   venue  freq
0                   Park  0.19
1             Restaurant  0.06
2  Gym / Fitness Center  0.06
3      Mobile Phone Shop  0.06
4           Dance Studio  0.06


----Nijegorodskij----
                  venue  freq
0         Auto Workshop  0.09
1  Arts & Crafts Store  0.09
2    Martial Arts Dojo  0.09
3           Public Art  0.06
4                 Café  0.06


----Perovo----
                        venue  freq
0        Gym / Fitness Center  0.09
1              Cosmetics Shop  0.07
2                        Park  0.05
3                  Playground  0.05
4  Middle Eastern Restaurant  0.05


----Orehovo-Borisovo Ujnoe----
                   venue  freq
0            Supermarket  0.12
1  Gym / Fitness Center  0.12
2              Pet Store  0.12
3         Farmers Market  0.12
```

```
4        Cosmetics Shop  0.12


----Mar_ino----
                  venue  freq
0           Supermarket  0.11
1                  Park  0.09
2  Gym / Fitness Center  0.08
3            Beer Store  0.08
4           Pizza Place  0.04


----Vesnaki----
                  venue  freq
0  Gym / Fitness Center  0.14
1                  Park  0.14
2         Historic Site  0.14
3           Supermarket  0.07
4        Clothing Store  0.07


----Preobrajenskoe----
            venue  freq
0     Coffee Shop  0.07
1     Supermarket  0.07
2  Soccer Stadium  0.04
3          Arcade  0.04
4    Dance Studio  0.04


----Sokolinaa Gora----
                         venue  freq
0                  Coffee Shop  0.12
1         Gym / Fitness Center  0.08
2             Video Game Store  0.04
3  Eastern European Restaurant  0.04
4             Electronics Store  0.04


----Lublino----
               venue  freq
0               Park   0.3
1         Print Shop   0.1
2          Gun Range   0.1
3  Health Food Store   0.1
4             Palace   0.1
```

```
----Tekstil_siki----
                 venue  freq
0                Track  0.08
1  Caucasian Restaurant  0.08
2  Gym / Fitness Center  0.08
3      Asian Restaurant  0.04
4          Skating Rink  0.04


----Brateevo----
           venue  freq
0           Park  0.11
1  Cosmetics Shop  0.08
2     Supermarket  0.08
3           Café  0.06
4  Clothing Store  0.06


----Zablikovo----
                       venue  freq
0                Supermarket  0.08
1           Asian Restaurant  0.04
2             Cosmetics Shop  0.04
3  Middle Eastern Restaurant  0.04
4               Soccer Field  0.04


----Razanskij----
                         venue  freq
0                         Café  0.05
1  Eastern European Restaurant  0.05
2            Food & Drink Shop  0.05
3                  Pizza Place  0.05
4             Sushi Restaurant  0.05


----Izmajlovo----
                 venue  freq
0          Coffee Shop  0.07
1                Hotel  0.06
2  Sporting Goods Shop  0.04
3          Music Venue  0.04
4                 Park  0.04
```

```
----Novogireevo----
                 venue  freq
0        Cosmetics Shop  0.09
1  Gym / Fitness Center  0.06
2            Beer Store  0.06
3     Health Food Store  0.06
4                   Gym  0.06


----Kuz_minki----
              venue  freq
0              Park  0.05
1   Toy / Game Store  0.04
2        Restaurant  0.04
3     Clothing Store  0.04
4  Health Food Store  0.04


----Gol_anovo----
         venue  freq
0  Dance Studio  0.22
1     Pet Store  0.11
2        Forest  0.11
3   Gas Station  0.11
4   Candy Store  0.11


----Severnoe Izmajlovo----
                 venue  freq
0           Supermarket  0.10
1          Gourmet Shop  0.10
2             Pool Hall  0.10
3  Gym / Fitness Center  0.10
4       Supplement Shop  0.05


----Kapotna----
                      venue  freq
0        Chinese Restaurant  0.14
1  Middle Eastern Restaurant  0.14
2                 Gun Range  0.14
3                      Café  0.14
4            Cosmetics Shop  0.14
```

```
----Vyhino-Julebino----
                   venue  freq
0              Pet Store  0.06
1   Gym / Fitness Center  0.06
2    Japanese Restaurant  0.04
3            Supermarket  0.04
4         Cosmetics Shop  0.04


----Vostocnyj----
           venue  freq
0  Big Box Store  0.25
1     Restaurant  0.12
2    Candy Store  0.12
3        Stadium  0.12
4       Bus Stop  0.12


----Vostocnoe Izmajlovo----
                   venue  freq
0   Gym / Fitness Center  0.08
1            Supermarket  0.05
2            Flower Shop  0.05
3                   Park  0.05
4            Pizza Place  0.05


----Ivanovskoe----
            venue  freq
0     Supermarket  0.14
1  Cosmetics Shop  0.10
2             Gym  0.10
3            Park  0.10
4     Pizza Place  0.10


----Kosino-Uhtomskij----
                   venue  freq
0   Gym / Fitness Center  0.33
1          Shopping Mall  0.11
2            Supermarket  0.06
3    Sporting Goods Shop  0.06
4   Fast Food Restaurant  0.06


----Novokosino----
```

```
              venue  freq
0              Lake  0.07
1  Sushi Restaurant  0.07
2        Restaurant  0.03
3         Bookstore  0.03
4          Boutique  0.03


----Nekrasovka----
               venue  freq
0  Convenience Store  0.20
1        Supermarket  0.15
2           Gym Pool  0.05
3          Pet Store  0.05
4  Food & Drink Shop  0.05


----Kunzevo----
                  venue  freq
0                  Park  0.05
1           Flower Shop  0.05
2  Gym / Fitness Center  0.05
3           Supermarket  0.05
4         Auto Workshop  0.03
```

Let's put that into a *pandas* dataframe

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

In [196]:

```python
def return_most_common_venues(row, num_top_venues):
    row_categories = row.iloc[2:]
    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.

In [197]:

```
num_top_venues = 10

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

# create columns according to number of top venues
columns = ['Neighbourhood']
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['Neighbourhood'] = venues_grouped.apply(lambda x: str(x['Neighbourhood']) + ' #' + str(int(x['Regio
n index'])), axis=1)

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

neighborhoods_venues_sorted.shape
```

Out[197]:

```
(146, 11)
```

In [198]:
```
neighborhoods_venues_sorted
```

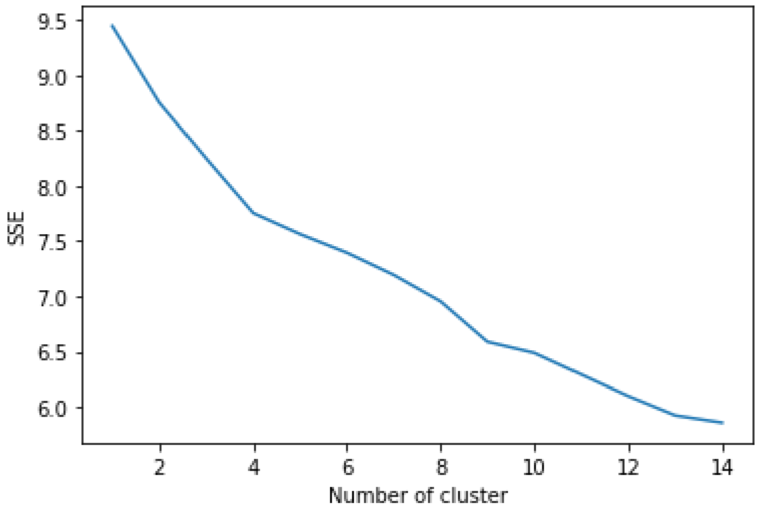
Out[198]:

	Neighbourhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Kievskij #0	Train Station	Castle	Food	Fast Food Restaurant	Film Studio	Fish Market	Fishing Spot	Fishing Store	Flea Market	Flower Shop
1	Filevskij Park #1	Gym / Fitness Center	Park	Pedestrian Plaza	Cosmetics Shop	Garden	Ski Area	Mobile Phone Shop	Gym	Go Kart Track	Coffee Shop
2	Novofedorovskoe #2	Park	Hotel Bar	Farm	Fast Food Restaurant	Film Studio	Fish Market	Fishing Spot	Fishing Store	Flea Market	Flower Shop
3	Rogovskoe #3	Garden Center	Convenience Store	Food	Farmers Market	Fast Food Restaurant	Film Studio	Fish Market	Fishing Spot	Fishing Store	Flea Market
4	"Mosrentgen" #4	Park	Gym / Fitness Center	Zoo Exhibit	Food	Film Studio	Fish Market	Fishing Spot	Fishing Store	Flea Market	Flower Shop
...	...	...	...	...	...	...	...	...	...	...	...
141	Ivanovskoe #141	Supermarket	Park	Gym	Pizza Place	Cosmetics Shop	Convenience Store	Department Store	Gym / Fitness Center	Stadium	Stables
142	Kosino-Uhtomskij #142	Gym / Fitness Center	Shopping Mall	Restaurant	Park	Sporting Goods Shop	Residential Building (Apartment / Condo)	Supermarket	Fast Food Restaurant	Big Box Store	Arcade
143	Novokosino #143	Lake	Sushi Restaurant	Health Food Store	Grocery Store	Liquor Store	Convenience Store	German Restaurant	Bookstore	Supermarket	Japanese Restaurant
144	Nekrasovka #144	Convenience Store	Supermarket	Concert Hall	Gym Pool	Auto Workshop	Pet Store	Café	Food & Drink Shop	Bath House	Lake
145	Kunzevo #145	Park	Flower Shop	Supermarket	Gym / Fitness Center	Pharmacy	Spa	Beach	Convenience Store	Auto Workshop	Food & Drink Shop

146 rows × 11 columns

Finding optimal number of clusters

In [199]:
```
from sklearn.cluster import KMeans
import matplotlib.pyplot as plt

venues_grouped_clustering = venues_grouped.drop(['Neighbourhood', 'Region index'], 1)

sse = {}
for k in range(1, 15):
    kmeans = KMeans(n_clusters=k, max_iter=1000).fit(venues_grouped_clustering)
    #venues_grouped_clustering["Clusters"] = kmeans.labels_
    sse[k] = kmeans.inertia_ # Inertia: Sum of distances of samples to their closest cluster center
#plt.figure()
plt.plot(list(sse.keys()), list(sse.values()))
plt.xlabel("Number of cluster")
plt.ylabel("SSE")
plt.show()
```



In [200]:
```
venues_grouped_clustering
```

Out[200]:

	ATM	Accessories Store	Adult Boutique	Airport	Airport Lounge	Airport Service	American Restaurant	Aquarium	Arcade	Argentinian Restaurant	...	Water Park	Waterfall	Waterfront	Wine Bar	Wine Shop	W
0	0.0	0.0	0.00000	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	...	0.0	0.0	0.0	0.0	0.000000	
1	0.0	0.0	0.03125	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	...	0.0	0.0	0.0	0.0	0.000000	
2	0.0	0.0	0.00000	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	...	0.0	0.0	0.0	0.0	0.000000	
3	0.0	0.0	0.00000	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	...	0.0	0.0	0.0	0.0	0.000000	
4	0.0	0.0	0.00000	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	...	0.0	0.0	0.0	0.0	0.000000	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	
141	0.0	0.0	0.00000	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	...	0.0	0.0	0.0	0.0	0.000000	
142	0.0	0.0	0.00000	0.0	0.0	0.0	0.0	0.0	0.055556	0.0	...	0.0	0.0	0.0	0.0	0.000000	
143	0.0	0.0	0.00000	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	...	0.0	0.0	0.0	0.0	0.033333	
144	0.0	0.0	0.00000	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	...	0.0	0.0	0.0	0.0	0.000000	
145	0.0	0.0	0.00000	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	...	0.0	0.0	0.0	0.0	0.016667	

146 rows × 339 columns

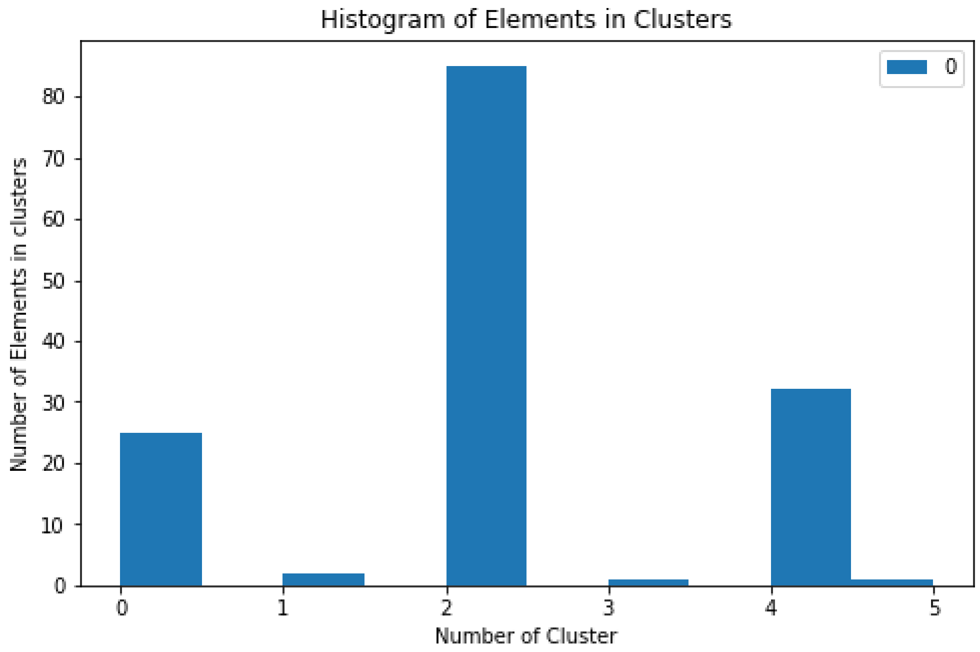
Cluster Neighborhoods

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

In [201]:
```python
# import k-means from clustering stage
from sklearn.cluster import KMeans

# set number of clusters
kclusters = 6

venues_grouped_clustering = venues_grouped.drop(['Neighbourhood', 'Region index'], 1)

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

# check cluster labels generated for each row in the dataframe
kmeans.labels_
```

Out[201]:
```
array([5, 2, 1, 3, 1, 4, 4, 4, 0, 4, 0, 0, 4, 2, 4, 2, 2, 0, 2, 2, 4, 4,
       2, 4, 2, 4, 2, 4, 2, 4, 2, 4, 0, 0, 2, 0, 2, 0, 2, 0, 2, 2, 2, 2,
       0, 2, 0, 2, 0, 0, 2, 4, 0, 2, 2, 2, 2, 0, 2, 0, 0, 2, 2, 2, 4, 0,
       2, 0, 2, 2, 2, 2, 4, 0, 2, 2, 2, 2, 4, 4, 2, 2, 2, 4, 2, 2, 2, 2,
       2, 0, 2, 2, 4, 2, 2, 2, 2, 2, 4, 2, 0, 2, 0, 4, 2, 0, 2, 2, 2, 2,
       4, 4, 2, 2, 2, 2, 2, 2, 2, 4, 2, 2, 2, 4, 2, 2, 2, 0, 2, 4, 2, 4,
       2, 2, 2, 4, 2, 2, 2, 4, 2, 4, 2, 2, 4, 2], dtype=int32)
```

In [202]:
```
pd.DataFrame(kmeans.labels_).plot(kind='hist', figsize=(8, 5))
plt.title('Histogram of Elements in Clusters')
plt.ylabel('Number of Elements in clusters')
plt.xlabel('Number of Cluster')
plt.show()
```



I choose to use six clusters. It seems it's optimal in our case. It gives us three fullfilled clusters and 3 outcasts.

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

In [203]:
```
# keep it for future analysis
dfTop10 = neighborhoods_venues_sorted
```

In [204]:

```
#neighborhoods_venues_sorted.drop(columns=['Cluster Labels'], inplace=True)
# add clustering labels
neighborhoods_venues_sorted.insert(0, 'Cluster Labels', kmeans.labels_)
neighborhoods_venues_sorted
#d_merged = moscow_regions

# merge d_grouped with initial data to add latitude/longitude for each neighborhood
#d_merged = d_merged.join(neighborhoods_venues_sorted.set_index('Neighbourhood'), on='Neighbourhood')

#d_merged.head() # check the last columns!
```

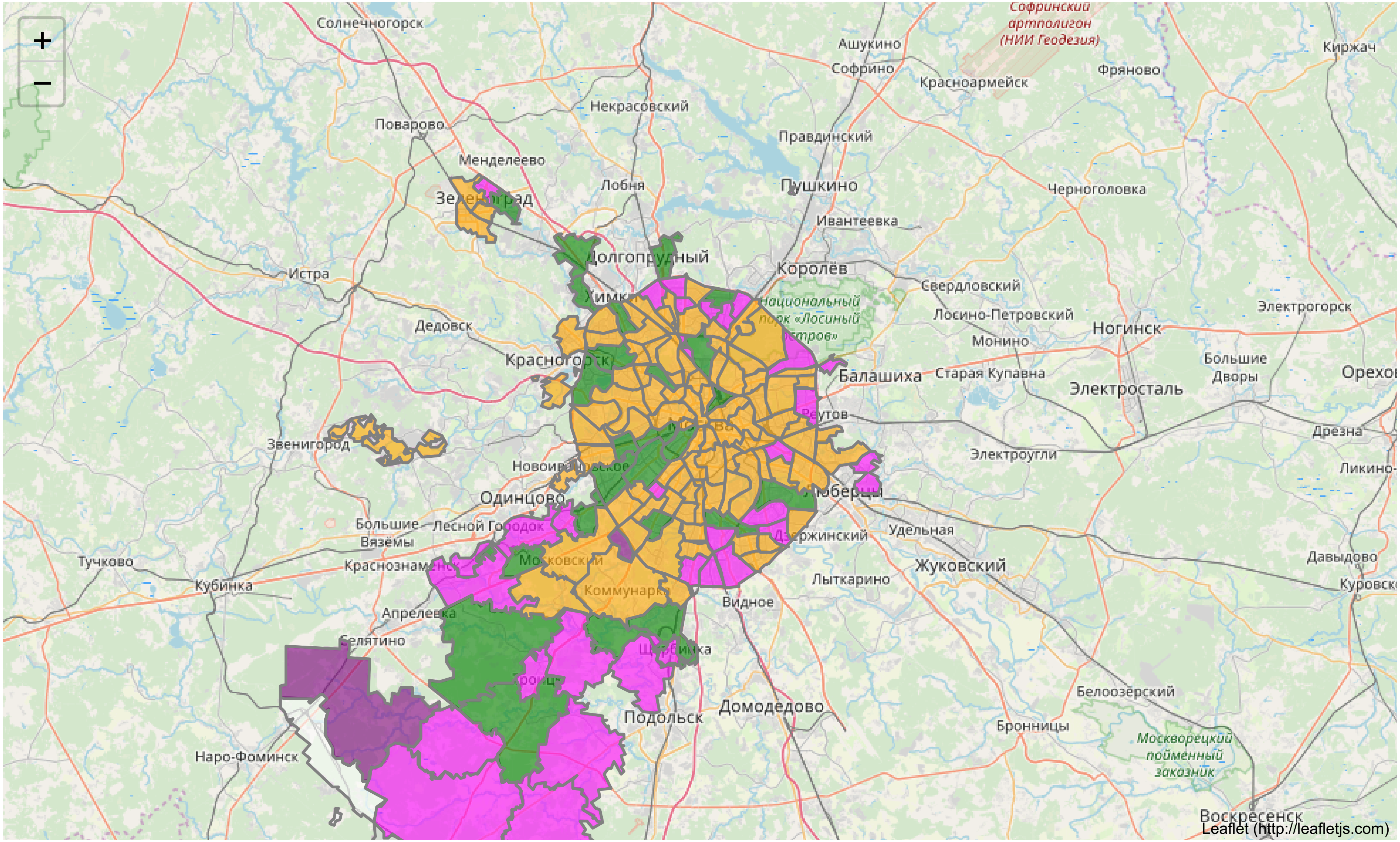
Out[204]:

	Cluster Labels	Neighbourhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	5	Kievskij #0	Train Station	Castle	Food	Fast Food Restaurant	Film Studio	Fish Market	Fishing Spot	Fishing Store	Flea Market	Flower Shop
1	2	Filevskij Park #1	Gym / Fitness Center	Park	Pedestrian Plaza	Cosmetics Shop	Garden	Ski Area	Mobile Phone Shop	Gym	Go Kart Track	Coffee Shop
2	1	Novofedorovskoe #2	Park	Hotel Bar	Farm	Fast Food Restaurant	Film Studio	Fish Market	Fishing Spot	Fishing Store	Flea Market	Flower Shop
3	3	Rogovskoe #3	Garden Center	Convenience Store	Food	Farmers Market	Fast Food Restaurant	Film Studio	Fish Market	Fishing Spot	Fishing Store	Flea Market
4	1	"Mosrentgen" #4	Park	Gym / Fitness Center	Zoo Exhibit	Food	Film Studio	Fish Market	Fishing Spot	Fishing Store	Flea Market	Flower Shop
...	...	...	...	...	...	...	...	...	...	...	...	...
141	4	Ivanovskoe #141	Supermarket	Park	Gym	Pizza Place	Cosmetics Shop	Convenience Store	Department Store	Gym / Fitness Center	Stadium	Stables
142	2	Kosino-Uhtomskij #142	Gym / Fitness Center	Shopping Mall	Restaurant	Park	Sporting Goods Shop	Residential Building (Apartment / Condo)	Supermarket	Fast Food Restaurant	Big Box Store	Arcade
143	2	Novokosino #143	Lake	Sushi Restaurant	Health Food Store	Grocery Store	Liquor Store	Convenience Store	German Restaurant	Bookstore	Supermarket	Japanese Restaurant
144	4	Nekrasovka #144	Convenience Store	Supermarket	Concert Hall	Gym Pool	Auto Workshop	Pet Store	Café	Food & Drink Shop	Bath House	Lake
145	2	Kunzevo #145	Park	Flower Shop	Supermarket	Gym / Fitness Center	Pharmacy	Spa	Beach	Convenience Store	Auto Workshop	Food & Drink Shop

146 rows × 12 columns

In [206]:
```
clusterColors=['green', 'purple','orange', 'lightgray','magenta', 'white','blue','aqua', 'violet', 'blue', ]
map = folium.Map(location=[55.7532358, 37.6225412], zoom_start=8.5)
regions = folium.map.FeatureGroup()
for index in range(0, moscow_regions.shape[0]):
    regionName = moscow_regions.iloc[index,1]
    #print(regionName)
    try:
        for segmentIndex in range(0,len(regionPolyList[index])):
            segment = regionPolyList[index][segmentIndex]
            regions.add_child( folium.features.PolygonMarker(segment, color='gray', fill_color=clusterColors[kmeans.labels_[ind
ex]], weight=2) )
            segmentCenter = regionCentersList[index][segmentIndex]
            segmentRadius = radiusesList[index][segmentIndex]

    #            regions.add_child(
    #                folium.features.Circle(location=segmentCenter,radius=segmentRadius,fill=False,color='green')
    #            )

    #            regions.add_child(
    #                folium.features.CircleMarker(
    #                    segmentCenter,
    #                    radius=8,
    #                    color='yellow',
    #                    fill=True,
    #                    fill_color='red',
    #                    fill_opacity=0.6,
    #                    popup=str(index) + "." + regionName
    #                )
    #            )
    except:
        print('Error on index {}'.format(regionName) )



map.add_child(regions)


# display map
map
```

Out[206]:



Results

The map gives us illustration of clusters distribution. We have three big groups and three odd ones. Let's try to analyze their content. What are the common features of these clusters? Let's see and give them applicable names.

Cluster 0

In [120]:

```
dfTop10[dfTop10['Cluster Labels']==0]
```

Out[120]:

	Cluster Labels	Neighbourhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
8	0	Pervomajskoe #8	Italian Restaurant	Park	Tennis Court	Stables	Supermarket	Spa	Hotel	Construction & Landscaping	Fishing Spot	Gymnastics Gym
10	0	Vnukovo #10	Airport Service	Airport Lounge	Mobile Phone Shop	Scenic Lookout	Coffee Shop	Bar	Hotel	Performing Arts Venue	Gym Pool	Shopping Mall
11	0	Savelki #11	Park	Racetrack	Restaurant	Stables	Beach	Theater	Gourmet Shop	Sports Club	Soccer Field	Café
17	0	Krasnopahorskoe #17	Park	Basketball Court	Hotel	Eastern European Restaurant	Seafood Restaurant	Snack Place	Exhibit	Outdoors & Recreation	Burger Joint	Stables
32	0	Strogino #32	Park	Beach	Athletics & Sports	Surf Spot	Pet Store	Gym / Fitness Center	Pool	Trail	Massage Studio	Sauna / Steam Room
33	0	Moljaninovskij #33	Hotel	Restaurant	Eastern European Restaurant	Athletics & Sports	Flea Market	Food	Fast Food Restaurant	Film Studio	Fish Market	Fishing Spot
35	0	Kurkino #35	Park	Café	Bakery	Gym / Fitness Center	Pet Store	Gastropub	Butcher	Golf Course	Caucasian Restaurant	Fruit & Vegetable Store
37	0	Solnzevo #37	Park	Café	Pub	Art Gallery	Fast Food Restaurant	Bakery	Beach	Train Station	Coffee Shop	Car Wash
39	0	Voskresenskoe #39	Rest Area	Hotel	Movie Theater	Italian Restaurant	Beach	Resort	Restaurant	Clothing Store	Pool	Zoo Exhibit
44	0	Pokrovskoe-Stresnevo #44	Surf Spot	Park	Zoo Exhibit	Soccer Stadium	Canal Lock	Beach	Gym / Fitness Center	Flower Shop	Furniture / Home Store	Restaurant
46	0	Ocakovo-Matveevskoe #46	Auto Workshop	Café	Hotel	Bath House	Park	Big Box Store	Fishing Spot	Fishing Store	Flea Market	Flower Shop
48	0	Levoberejnyj #48	Coffee Shop	Park	Café	Middle Eastern Restaurant	Sporting Goods Shop	Multiplex	History Museum	Shoe Repair	Blini House	Trail
49	0	Fili-Davydkovo #49	Park	Boutique	Clothing Store	Middle Eastern Restaurant	Shopping Mall	Sushi Restaurant	Hobby Shop	Liquor Store	Multiplex	Gourmet Shop
52	0	Ramenki #52	Park	Coffee Shop	Yoga Studio	Spa	Restaurant	Gourmet Shop	Caucasian Restaurant	Liquor Store	Gym / Fitness Center	Salon / Barbershop

	Cluster Labels	Neighbourhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
57	0	Ujnoe Butovo #57	Park	Italian Restaurant	Pet Store	Sushi Restaurant	Japanese Restaurant	Flower Shop	Gym	Beer Bar	Beer Store	Gastropub
59	0	Dorogomilovo #59	Coffee Shop	Park	Food & Drink Shop	Spa	Fountain	Auto Workshop	Record Shop	Modern European Restaurant	Theater	Clothing Store
60	0	Kon_kovo #60	Coffee Shop	Cosmetics Shop	Café	Park	Pool	Yoga Studio	Eastern European Restaurant	Dumpling Restaurant	Beer Store	Middle Eastern Restaurant
65	0	Aeroport #65	Coffee Shop	Health Food Store	Park	Pet Store	Hotel	Auto Workshop	Hockey Arena	Liquor Store	Palace	Convenience Store
67	0	Severnyj #67	Clothing Store	Sporting Goods Shop	Zoo Exhibit	Kids Store	Restaurant	Pizza Place	Coffee Shop	Park	Forest	Nightclub
73	0	Hamovniki #73	Yoga Studio	Scenic Lookout	Park	Health & Beauty Service	Art Gallery	Movie Theater	Monastery	Seafood Restaurant	Gymnastics Gym	Massage Studio
89	0	Ostankinskij #89	Park	Fountain	Playground	Scenic Lookout	Garden	Russian Restaurant	Science Museum	Concert Hall	Botanical Garden	Stables
100	0	Severnoe Medvedkovo #100	Auto Workshop	Café	Bakery	Pharmacy	Park	Japanese Restaurant	Gaming Cafe	Multiplex	Bookstore	Flower Shop
102	0	Moskvorec_e-Saburovo #102	Coffee Shop	Laser Tag	Hotel	Park	Clothing Store	Gourmet Shop	Salon / Barbershop	Tennis Court	Tea Room	Pedestrian Plaza
105	0	Krasnosel_skij #105	Hookah Bar	Coffee Shop	Flower Shop	Hotel	Yoga Studio	Jewelry Store	Bookstore	Restaurant	Pub	Hotel Bar
127	0	Lublino #127	Park	Dance Studio	Print Shop	History Museum	Palace	Gun Range	Chinese Restaurant	Health Food Store	Food & Drink Shop	Food

Cluster 1

In [121]:
```
dfTop10[dfTop10['Cluster Labels']==1]
```

Out[121]:

	Cluster Labels	Neighbourhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
2	1	Novofedorovskoe #2	Park	Hotel Bar	Farm	Fast Food Restaurant	Film Studio	Fish Market	Fishing Spot	Fishing Store	Flea Market	Flower Shop
4	1	"Mosrentgen" #4	Park	Gym / Fitness Center	Zoo Exhibit	Food	Film Studio	Fish Market	Fishing Spot	Fishing Store	Flea Market	Flower Shop

Cluster 2

```
In [76]: dfTop10[dfTop10['Cluster Labels']==2]
```

Out[76]:

	Cluster Labels	Neighbourhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
1	2	Filevskij Park #1	Gym / Fitness Center	Park	Pedestrian Plaza	Cosmetics Shop	Garden	Ski Area	Mobile Phone Shop	Gym	Go Kart Track	Coffee Shop
13	2	Silino #13	Pharmacy	Mobile Phone Shop	Toy / Game Store	Cosmetics Shop	Market	Electronics Store	Pizza Place	History Museum	Forest	Blini House
15	2	Krukovo #15	Park	Pizza Place	Fast Food Restaurant	Bar	Coffee Shop	Paper / Office Supplies Store	Paintball Field	Beer Bar	Gym / Fitness Center	Gym
16	2	Sukino #16	Park	Salon / Barbershop	Food & Drink Shop	Gym / Fitness Center	Cosmetics Shop	Convenience Store	Restaurant	Spa	Clothing Store	ATM
18	2	Nagatinskij Zaton #18	Gym / Fitness Center	Eastern European Restaurant	Sculpture Garden	Caucasian Restaurant	Park	Scenic Lookout	Coffee Shop	Market	History Museum	Palace
...	...	...	...	...	...	...	...	...	...	...	...	...
138	2	Vyhino-Julebino #138	Gym / Fitness Center	Pet Store	Japanese Restaurant	Cosmetics Shop	Park	Coffee Shop	Fast Food Restaurant	Supermarket	Sushi Restaurant	Grocery Store
140	2	Vostocnoe Izmajlovo #140	Gym / Fitness Center	Park	Pizza Place	Flower Shop	Supermarket	Farmers Market	Paper / Office Supplies Store	Exhibit	Middle Eastern Restaurant	Theater
142	2	Kosino-Uhtomskij #142	Gym / Fitness Center	Shopping Mall	Restaurant	Park	Sporting Goods Shop	Residential Building (Apartment / Condo)	Supermarket	Fast Food Restaurant	Big Box Store	Arcade
143	2	Novokosino #143	Lake	Sushi Restaurant	Health Food Store	Grocery Store	Liquor Store	Convenience Store	German Restaurant	Bookstore	Supermarket	Japanese Restaurant
145	2	Kunzevo #145	Park	Flower Shop	Supermarket	Gym / Fitness Center	Pharmacy	Spa	Beach	Convenience Store	Auto Workshop	Food & Drink Shop

85 rows × 12 columns

Cluster 3

In [77]:
```
dfTop10[dfTop10['Cluster Labels']==3]
```

Out[77]:

	Cluster Labels	Neighbourhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
3	3	Rogovskoe #3	Garden Center	Convenience Store	Food	Farmers Market	Fast Food Restaurant	Film Studio	Fish Market	Fishing Spot	Fishing Store	Flea Market

In [ ]:

In [78]:

```
dfTop10[dfTop10['Cluster Labels']==4]
```

Out[78]:

	Cluster Labels	Neighbourhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th M Comm Ver
5	4	Voronovskoe #5	Lake	Resort	Convenience Store	Restaurant	Pool	Steakhouse	Farmers Market	Ski Trail	River	P
6	4	Mihajlovo-Arzevskoe #6	Stables	Recreation Center	Restaurant	Historic Site	Bistro	Pharmacy	Food & Drink Shop	Farm	Eastern European Restaurant	Fish St
7	4	Maruskinskoe #7	Scenic Lookout	Train Station	Racetrack	Park	Supermarket	Resort	Shopping Mall	Restaurant	Convenience Store	Gas Stat
9	4	Matuskino #9	Supermarket	Brewery	Burger Joint	Electronics Store	Plaza	Café	ATM	Fish Market	Toy / Game Store	Fast Fc Restaur
12	4	Vnukovskoe #12	Museum	Bed & Breakfast	Convenience Store	Historic Site	Badminton Court	Coffee Shop	Big Box Store	Exhibit	Café	For
14	4	Kokoskino #14	Train Station	Convenience Store	Forest	Farm	Flower Shop	Farmers Market	Fast Food Restaurant	Film Studio	Fish Market	Fish S
20	4	Klenovskoe #20	Lake	Gun Range	Restaurant	Theme Park	Bus Stop	Film Studio	Fish Market	Fishing Spot	Fishing Store	F Mar
21	4	Dmitrovskij #21	Supermarket	Café	Park	Restaurant	Pizza Place	Eastern European Restaurant	Soccer Field	Auto Workshop	Fast Food Restaurant	G
23	4	Troizk #23	Restaurant	Supermarket	Café	Park	Big Box Store	Trail	Bakery	Gym	Rest Area	Electron St
25	4	Sapovskoe #25	Historic Site	Convenience Store	Park	Café	Gym	Grocery Store	Forest	Food Court	Food & Drink Shop	Fount
27	4	Desenovskoe #27	Soccer Field	Park	Convenience Store	Supermarket	Bus Stop	Butcher	Café	Beer Store	Grocery Store	Farm Mar
29	4	Lomonosovskij #29	Convenience Store	Italian Restaurant	Dance Studio	Fish Market	Food & Drink Shop	Cosmetics Shop	Caucasian Restaurant	Health Food Store	Pharmacy	Gift Sh
31	4	Novo-Peredelkino #31	Supermarket	Gym / Fitness Center	Convenience Store	Mobile Phone Shop	Coffee Shop	Pizza Place	Fast Food Restaurant	Tennis Court	Grocery Store	Foo Drink Sh
51	4	Razanovskoe #51	Stables	Café	Restaurant	Gym / Fitness Center	Gym	Lake	Fast Food Restaurant	Convenience Store	Park	Airp
64	4	Serbinka #64	Gym / Fitness Center	Pizza Place	Supermarket	Convenience Store	Fast Food Restaurant	Climbing Gym	Hobby Shop	Hockey Arena	Flower Shop	BBQ Jo

	Cluster Labels	Neighbourhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Mo Comm Ver
72	4	Lianozovo #72	Park	Salon / Barbershop	Furniture / Home Store	Gas Station	Middle Eastern Restaurant	Gym	Gym / Fitness Center	Sporting Goods Shop	Fast Food Restaurant	Big B St
78	4	Altuf_evskij #78	Auto Workshop	Supermarket	Café	Convenience Store	Gym / Fitness Center	Grocery Store	Big Box Store	Pharmacy	Eastern European Restaurant	Salo Barbersh
79	4	Marfino #79	Convenience Store	Gas Station	Hotel Bar	Flower Shop	Grocery Store	Caucasian Restaurant	Middle Eastern Restaurant	Pet Store	Liquor Store	Bak
83	4	Certanovo Ujnoe #83	Supermarket	Pizza Place	Auto Workshop	Dance Studio	Middle Eastern Restaurant	Shoe Store	Restaurant	Pet Store	Pedestrian Plaza	P
92	4	Birulevo Zapadnoe #92	Supermarket	Big Box Store	Pool	Train Station	Park	Flea Market	Café	Hockey Rink	Flower Shop	Food Drink Sh
98	4	Ujnoe Medvedkovo #98	Supermarket	Pizza Place	Convenience Store	Sushi Restaurant	Gym	Athletics & Sports	Beer Bar	Coffee Shop	Fast Food Restaurant	Ten Co
103	4	Zarizyno #103	Park	Convenience Store	Butcher	Playground	Gym	Cosmetics Shop	Health Food Store	Pub	BBQ Joint	I
110	4	Birulevo Vostocnoe #110	Gym / Fitness Center	Park	Supermarket	Theater	Japanese Restaurant	Soccer Field	Bar	Beer Bar	Beer Store	Fount
111	4	Babuskinskij #111	Gym	Park	Supermarket	Pharmacy	Gym / Fitness Center	Japanese Restaurant	Food & Drink Shop	Multiplex	Café	Fast Fo Restaur
119	4	Losinoostrovskij #119	Park	Dance Studio	Supermarket	Gym / Fitness Center	Pizza Place	Restaurant	Convenience Store	Mobile Phone Shop	Theme Park Ride / Attraction	Caucas Restaur
123	4	Mar_ino #123	Supermarket	Park	Beer Store	Gym / Fitness Center	Pharmacy	Eastern European Restaurant	Convenience Store	Pizza Place	Caucasian Restaurant	Ro
129	4	Brateevo #129	Park	Supermarket	Cosmetics Shop	Fast Food Restaurant	Pizza Place	Clothing Store	Café	Tennis Court	Gym / Fitness Center	Pedestr Pla
131	4	Razanskij #131	Pizza Place	Sushi Restaurant	Park	Eastern European Restaurant	Café	Food & Drink Shop	Flower Shop	Theater	Beer Garden	Beer St
135	4	Gol_anovo #135	Dance Studio	Pizza Place	Gas Station	Park	Candy Store	Pet Store	Forest	Café	Fish Market	Fish S

	Cluster Labels	Neighbourhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th M Comm Ver
139	4	Vostocnyj #139	Big Box Store	Bakery	Restaurant	Bus Stop	Candy Store	Stadium	Reservoir	Dive Bar	Diner	Fish S
141	4	Ivanovskoe #141	Supermarket	Park	Gym	Pizza Place	Cosmetics Shop	Convenience Store	Department Store	Gym / Fitness Center	Stadium	Stab
144	4	Nekrasovka #144	Convenience Store	Supermarket	Concert Hall	Gym Pool	Auto Workshop	Pet Store	Café	Food & Drink Shop	Bath House	La

Cluster 5

In [79]:
```
dfTop10[dfTop10['Cluster Labels']==5]
```

Out[79]:

	Cluster Labels	Neighbourhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	5	Kievskij #0	Train Station	Castle	Food	Fast Food Restaurant	Film Studio	Fish Market	Fishing Spot	Fishing Store	Flea Market	Flower Shop

Discussion

Clusters 1,3 and 5 are clear exceptions. The properties of their elements differ sharply from others, since the size of these clusters is extremely small. It makes sense to exclude them from consideration. Clusters 0,2,4 have comparable sizes.

Given the features of filling objects with clusters, the following typical names can be given.

- **Cluster 0. *Green area of parks and cafes*. This cluster is dominated by park areas, cafes and coffee shops. Other objects to a lesser extent characterize this cluster.**
- **Cluster 2. *Shops and cafes*. Park zones also prevail in this cluster, but there is a significant preponderance regarding various types of stores. Other objects to a lesser extent characterize this cluster.**
- **Cluster 4. *Shops and supermarkets*. As for cluster 2, the predominance of shopping facilities is becoming overwhelming. Other objects to a lesser extent characterize this cluster.**

Conclusion

- We were able to get the division of the regions of the city of Moscow into several clusters.
- The accuracy of this method is somewhat degraded due to the lack of the number of objects returned by the Foursquare API. But this does not prevent a general idea of the properties of the obtained groupings.
- Also, some simplifications of mathematical methods introduce an error when calculating coordinates on a spherical surface, but this does not violate general principles.