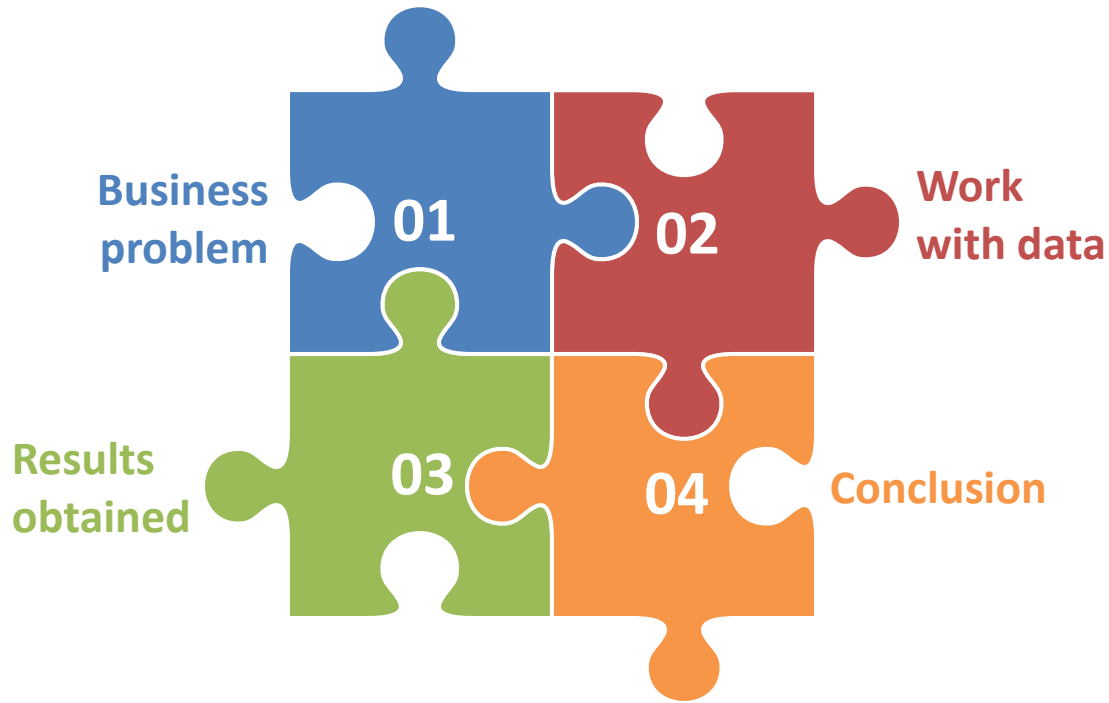


Finding the most popular venues in districts of Moscow

Viktor Klementev, klementiev.victor@gmail.com





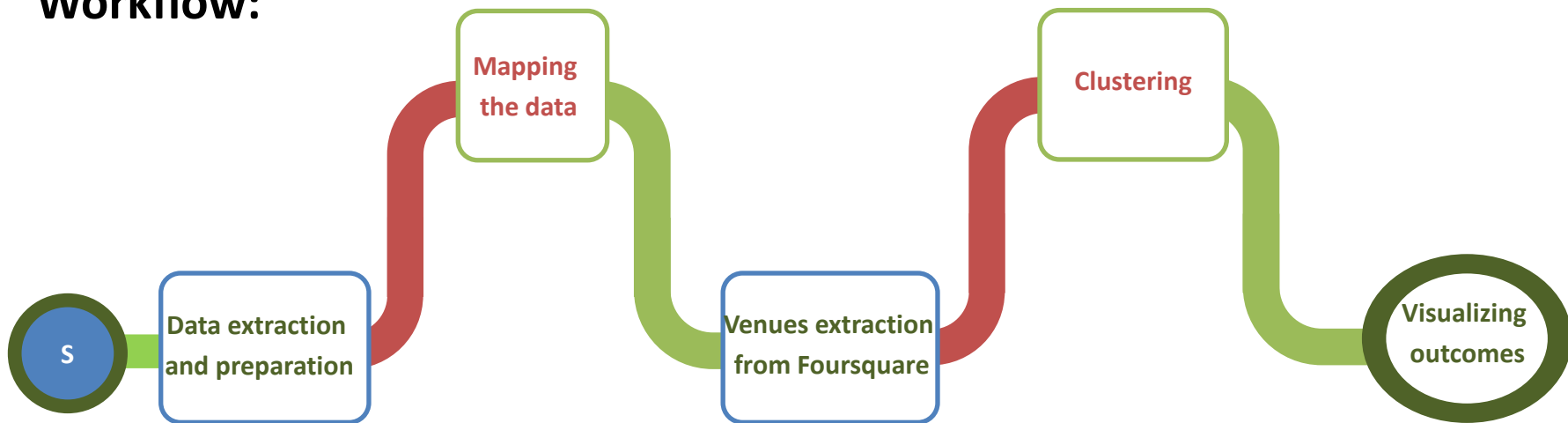
Business problem

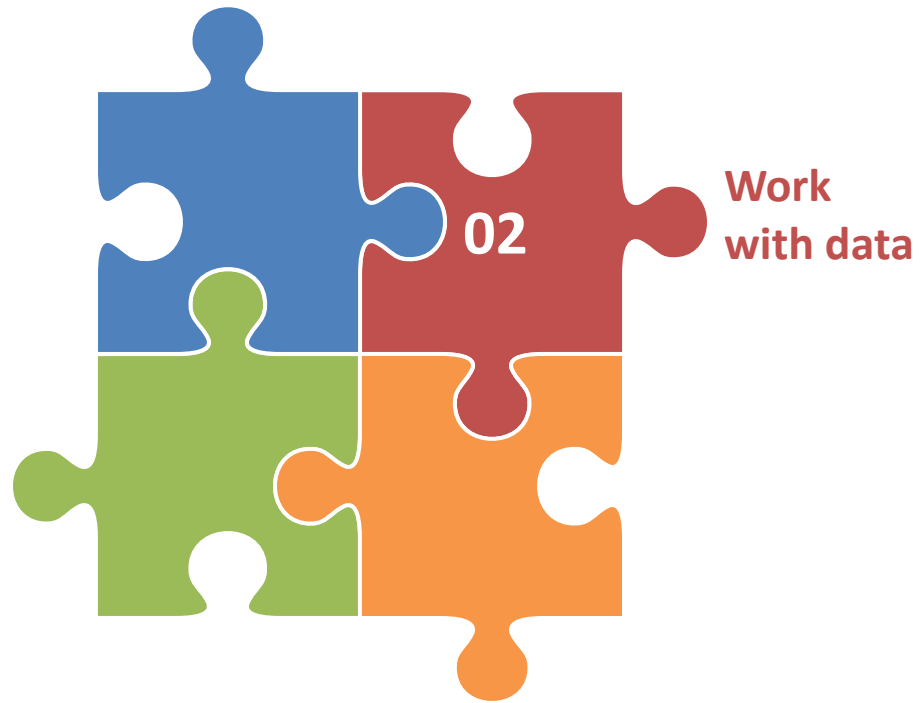
- ✓ There are 119 districts in Moscow and there should be similarities among them
- ✓ To find similarities I'm going to analyze socioeconomic indicators and venues clustering

Interested in final results parties:

- ✓ - business people who study city environments to find out new opportunities;
- urbanists and sociologists

Workflow:





Data that will be used and analyzed:

1. SOCIO ECONOMIC INDICATORS:



Population size of
every district



Real estate price in
every district



Working/residential
district feature

	District	District_en	Population	% of homeplaces	% of workplaces	Price per sq m
0	Академический	Akademichesky District	109387	41.5	58.5	205702
1	Алексеевский	Alexeyevsky District	80534	35.4	64.6	191495

2. FOURSQUARE DATA ON VENUES (2.683 VENUES, 297 UNIQUE CATEGORIES

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
2678	Yasenevo District	55.606262	37.533748	Baskin-Robbins	55.606465	37.536121	Ice Cream Shop
2679	Yasenevo District	55.606262	37.533748	Перекресток	55.607489	37.532367	Supermarket
2680	Yasenevo District	55.606262	37.533748	IL Патио	55.607347	37.532795	Italian Restaurant
2681	Yasenevo District	55.606262	37.533748	Территория	55.610085	37.535949	Nightclub
2682	Yasenevo District	55.606262	37.533748	Пряжа "Золотое Руно"	55.602885	37.529681	Arts & Crafts Store

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

There are 297 uniques venues categories.

Work with web scrapping and data cleaning :

Web scrapping of Wikipedia page(English translation is provided in the code). Before:

	0	1	2	3	4	5	6	7	8	9	10	11
0	1	Академический	Академический	ЮЗАО	5	83	109387	18762.78	2467	0	22	7
1	2	Алексеевский	Алексеевский	СВАО	5	29	80534	15223.82	1607	9	20	5

After data cleaning:

	District	Population	Population density	District square	address for Nominatim	location	point
0	Академический	109387	18762.78	5.83	Академический , ЮЗАО	(Академический, Москва, Юго-Западный администр...	(55.6897377, 37.5767712, 0.0)
1	Алексеевский	80534	15223.82	5.29	Алексеевский , СВАО	(Алексеевский, Москва, Северо-Восточный админи...	(55.8148783, 37.6506684, 0.0)

Web scrapping of real estate web site page. Before:

After data cleaning:

table1

\\n\\n,1,\\n,\\nОстоженка, \\nКропоткинская, \\t\\t\\t\\t\\t\\t\\t\\t\\tПарк культуры, 395\\xa0794,+1,3%\\n\\n,2,\\n,\\nЯкиманка, \\nНовокузнецкая, \\t\\t\\t\\t\\t\\t\\t\\t\\tПолянка, \\t\\t\\t\\t\\t\\t\\t\\t\\tТретьяковская, 376\\xa0733,+0,8%\\n\\n,3,\\n,\\nАрбат, \\nАлександровский сад, \\t\\t\\t\\t\\t\\t\\t\\t\\tАрбатская, \\t\\t\\t\\t\\t\\t\\t\\t\\tБиблиотека имени Ленина, \\t\\t\\t\\t\\t\\t\\t\\t\\tБоровицкая, \\t\\t\\t\\t\\t\\t\\t\\t\\tСмоленская, 364\\xa0093,+0,5%\\n\\n,4,\\n\\t\\t\\t\\t\\t\\t\\t\\t\\tЦентр Москвы, \\nКитай-город, \\t\\t\\t\\t\\t\\t\\t\\t\\tКузнецкий мост, \\t\\t\\t\\t\\t\\t\\t\\t\\tЛубянка, \\t\\t\\t\\t\\t\\t\\t\\t\\tОхотный ряд, \\t\\t\\t\\t\\t\\t\\t\\t\\tПлощадь Революции, \\t\\t\\t\\t\\t\\t\\t\\t\\tТеатральная, 355\\xa0644,+1,1%\\n\\n,5,\\n\\t\\t\\t\\t\\t\\t\\t\\t\\tТверской, \\nМаяковская, \\t\\t\\t\\t\\t\\t\\t\\t\\tПушкинская, \\t\\t\\t\\t\\t\\t\\t\\t\\tТверская, \\t\\t\\t\\t\\t\\t\\t\\t\\tЧеховская, 349\\xa0883,-0,8%\\n\\n,6,\\n,\\nХамовники, \\nЛужники, \\t\\t\\t\\t\\t\\t\\t\\t\\t\\tСпортивная, \\t\\t\\t\\t\\t\\t\\t\\t\\tФрунзенская, 322\\xa0677,+0,8%\\n\\n,7,\\n\\t\\t\\t\\t\\t\\t\\t\\t\\t\\tЗамоскворечье, \\nДобрынинская, \\t\\t\\t\\t\\t\\t\\t\\t\\t\\tОктябрьская, \\t\\t\\t\\t\\t\\t\\t\\t\\tПавелецкая, \\t\\t\\t\\t\\t\\t\\t\\t\\t\\tСерпуховская, 306\\xa0195,+2,1%\\n\\n,8,\\n\\t\\t\\t\\t\\t\\t\\t\\t\\tПресненский, \\nБаррикадная, \\t\\t\\t\\t\\t\\t\\t\\t\\t\\tВыставочная, \\t

	District	Price per sqm
0	Остоженка	395794
1	Якиманка	376733
2	Арбат	364093
3	Центр Москвы	355644
4	Тверской	349883

Results
obtained

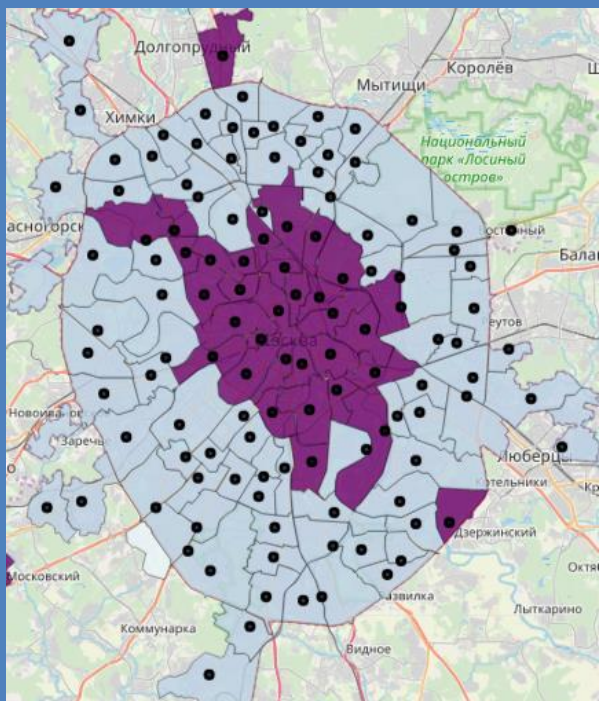


Web scrapping and data cleaning/wrangling results:

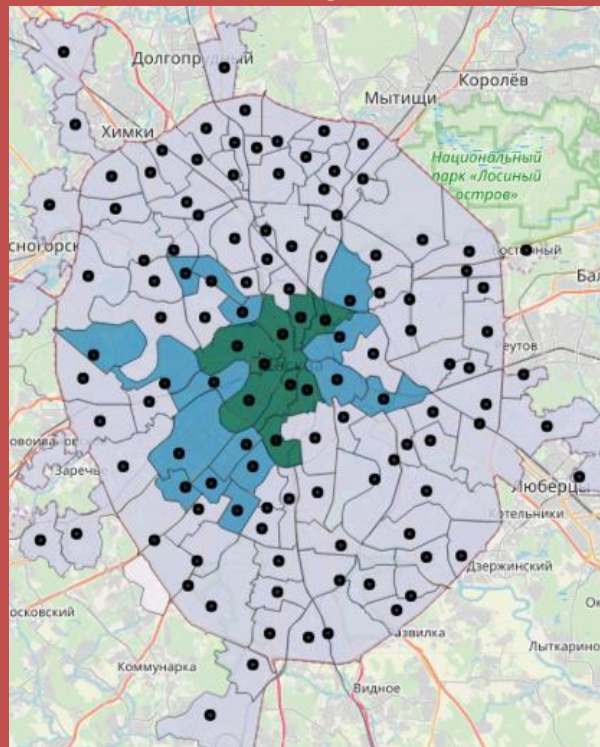
	District	District_en	Population	% of homeplaces	% of workplaces	Price per sq m	Latitude	Longitude	Work/residential district	Dummies for work/residential district	Price categories	Dummies for price
0	Академический	Akademichesky District	109387	41.5	58.5	205702	55.689738	37.576771	residential	0	average	3
1	Алексеевский	Alexeyevsky District	80534	35.4	64.6	191495	55.814878	37.650668	mostly for work	1	mass- segment	2
2	Алтуфьевский	Altufyevsky District	57596	60.5	39.5	147175	55.880255	37.581635	residential	0	mass- segment	2
3	Арбат	Arbat District	36125	12.7	87.3	362197	55.751199	37.589872	mostly for work	1	expensive	4
4	Аэропорт	Aeroport District	79486	38.5	61.5	206218	55.800402	37.533156	mostly for work	1	average	3
5	Бабушкинский	Babushkinsky District	88537	60.0	40.0	154080	55.865958	37.663894	residential	0	mass- segment	2
6	Басманный	Basmanny District	110694	17.0	83.0	224197	55.767281	37.669773	mostly for work	1	average	3
7	Беговой	Begovoy District	42781	18.8	81.2	237475	55.781917	37.566300	mostly for work	1	average	3
8	Бескудниковский	Beskudnikovsky District	79603	75.1	24.9	149875	55.863739	37.557322	residential	0	mass- segment	2

Results: chronolepth maps

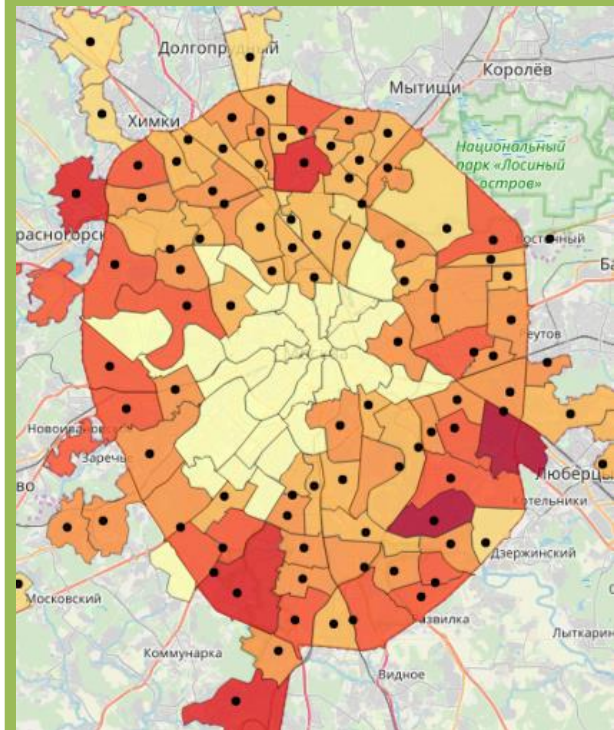
District for work /
residential district (blue)



Real estate price category (from the
center): expensive, average,
mass-segment



Division by population for
mass-segment price districts



Cluster analysis result, each district has cluster label now

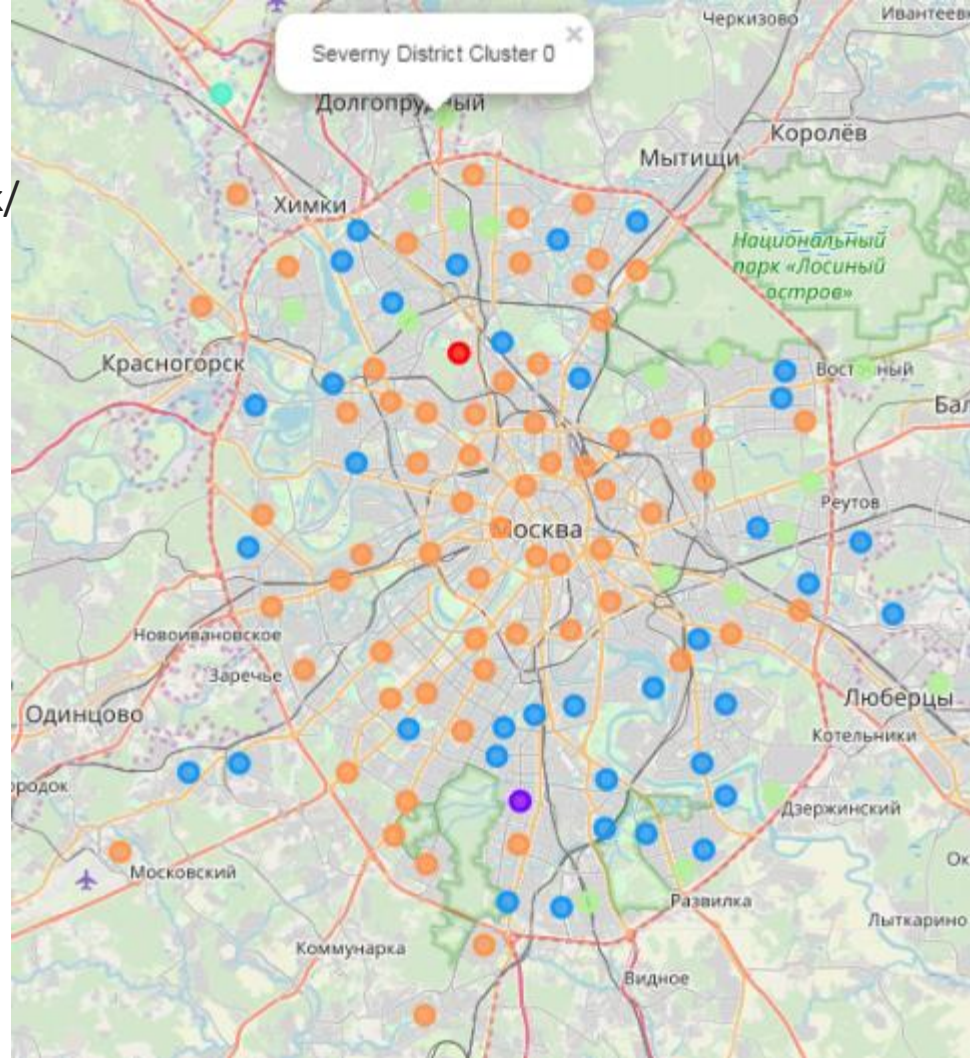
Neighborhood	Work/residential district	Price categories	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
Aeroport District	mostly for work	average	55.800402	37.533156	1	Coffee Shop	Cosmetics Shop	Café	Convenience Store	Flower Shop
Akademichesky District	residential	average	55.689738	37.576771	1	Pharmacy	Bakery	Beer Store	Pub	Park
Alexeyevsky District	mostly for work	mass-segment	55.814878	37.650668	4	Hotel	Auto Workshop	Supermarket	Mobile Phone Shop	Gym / Fitness Center
Altufyevsky District	residential	mass-segment	55.880255	37.581635	0	Supermarket	Grocery Store	Convenience Store	Building	Café
Arbat District	mostly for work	expensive	55.751199	37.589872	1	Coffee Shop	Hotel	Bar	Museum	Yoga Studio
Babushkinsky District	residential	mass-segment	55.865958	37.663894	1	Mobile Phone Shop	Cosmetics Shop	Notary	Middle Eastern Restaurant	Café
Basmanny District	mostly for work	average	55.767281	37.669773	1	Park	Hostel	Auto Workshop	Public Art	Recording Studio
Begovoy District	mostly for work	average	55.781917	37.566300	1	Dance Studio	Gym / Fitness Center	Café	Coffee Shop	Furniture / Home Store
Beskudnikovsky District	residential	mass-segment	55.863739	37.557322	4	Cosmetics Shop	Liquor Store	Moving Target	Soccer Field	Park
Bibirevo District	residential	mass-segment	55.883894	37.603577	1	Fast Food Restaurant	Clothing Store	Cosmetics Shop	Toy / Game Store	Pizza Place

Clustering results:

Division by real estate prices, population and for work/residential feature of district showed that there could be 6 groups of districts.

Clustering showed there are 3 actual clusters.
Other results of cluster analysis :

2. Some food and everyday supply venues are pre-dominantly concentrated in residential mass-segment districts.
3. The majority of work districts are in one cluster.
4. In cluster 1 supermarket is the most popular venue for residential mass-segment price districts.
5. Coffee shop is the most popular and common venue for 'mostly for work' and expensive real estate districts.



A part of cluster 1 table.

Supermarket is one of the most popular venues for residential districts

	Neighborhood	Work/residential district	Price categories	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
3	Altufyevsky District	residential	mass-segment	Supermarket	Grocery Store	Convenience Store	Building	Café
10	Biryulyovo Vostochnoye District	residential	mass-segment	Light Rail Station	Bus Stop	Supermarket	Cosmetics Shop	Outlet Store
12	Bogorodskoye District	residential	mass-segment	Supermarket	Bus Stop	Dance Studio	Pizza Place	Middle Eastern Restaurant
20	Dmitrovsky District	residential	mass-segment	Supermarket	Gym	Pizza Place	Bath House	Eastern European Restaurant
28	Ivanovskoye District	residential	mass-segment	Supermarket	Cosmetics Shop	Pizza Place	Pet Store	Fast Food Restaurant
30	Kapotnya District	mostly for work	mass-segment	Supermarket	Asian Restaurant	Deli / Bodega	Department Store	Hotel
36	Koptevo District	residential	mass-segment	Auto Workshop	Supermarket	Bus Stop	Light Rail Station	Pharmacy
54	Metrogorodok District	residential	mass-segment	Supermarket	Pharmacy	Tram Station	Sushi Restaurant	Food & Drink Shop
62	Nekrasovka District	residential	mass-segment	Supermarket	Soccer Field	Park	Concert Hall	Gym Pool
63	Nizhegorodsky District	mostly for work	average	Auto Workshop	Hotel	Bar	Supermarket	Convenience Store
65	Novogireyevo District	residential	mass-segment	Bus Stop	Supermarket	Gym	Café	Auto Workshop
70	Orekhovo-Borisovo Yuzhnoye District	residential	mass-segment	Supermarket	Gym	Dog Run	Cosmetics Shop	Sushi Restaurant
81	Ryazansky District	residential	mass-segment	Pharmacy	Supermarket	Car Wash	Sushi Restaurant	Eastern European Restaurant
87	Severnoy District	mostly for work	mass-segment	Playground	Supermarket	Pizza Place	Zoo Exhibit	Food

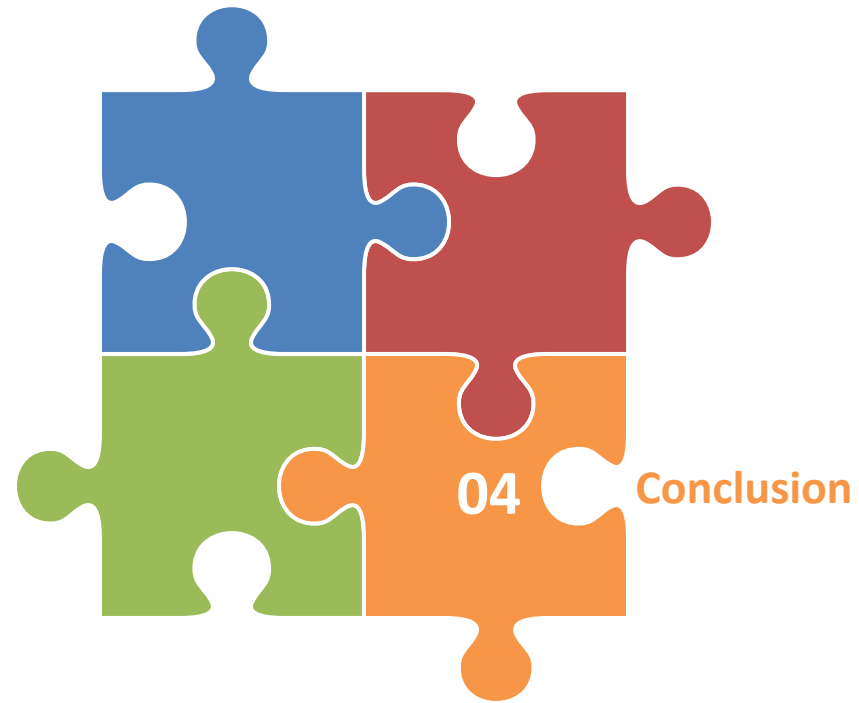
A part of cluster 2 table.

Coffee shop is popular venue for 'mostly for work' and expensive real estate districts

	Neighborhood	Work/residential district	Price categories	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Aeropot District	mostly for work	average	Coffee Shop	Cosmetics Shop	Café	Convenience Store	Flower Shop
6	Basmanny District	mostly for work	average	Park	Hostel	Auto Workshop	Public Art	Recording Studio
7	Begovoy District	mostly for work	average	Dance Studio	Gym / Fitness Center	Café	Coffee Shop	Furniture / Home Store
95	Tagansky District	mostly for work	average	Bar	Café	Coffee Shop	Health Food Store	Pharmacy
91	Sokolniki District	mostly for work	average	Toy / Game Store	Sporting Goods Shop	Hotel	Theater	Gourmet Shop
22	Dorogomilovo District	mostly for work	average	Coffee Shop	Salon / Barbershop	Wine Shop	Boutique	Smoke Shop
89	Sokol District	mostly for work	average	Bakery	Beer Bar	Electronics Store	Hobby Shop	Caucasian Restaurant
109	Yakimanka District	mostly for work	expensive	Coffee Shop	Café	Art Gallery	Restaurant	Gym / Fitness Center
100	Tverskoy District	mostly for work	expensive	Theater	Coffee Shop	Seafood Restaurant	Italian Restaurant	Yoga Studio
77	Presnensky District	mostly for work	expensive	Gym / Fitness Center	Coffee Shop	Board Shop	Café	Mobile Phone Shop
53	Meshchansky District	mostly for work	expensive	Coffee Shop	Middle Eastern Restaurant	Bakery	Caucasian Restaurant	Spa
116	Zamoskvorechye District	mostly for work	expensive	Coffee Shop	Hotel	Yoga Studio	Caucasian Restaurant	Bakery
4	Arbat District	mostly for work	expensive	Coffee Shop	Hotel	Bar	Museum	Yoga Studio
21	Donskoy District	mostly for work	expensive	Furniture / Home Store	Electronics Store	Gym / Fitness Center	Auto Workshop	Squash Court
31	Khamovniki District	mostly for work	expensive	Gym / Fitness Center	Coffee Shop	Cocktail Bar	Restaurant	Pizza Place

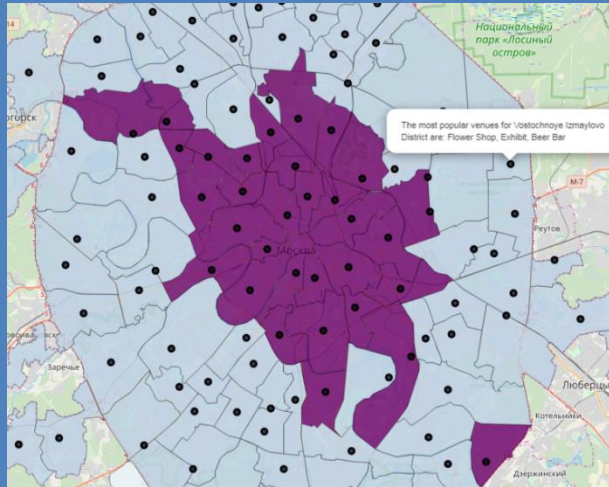
A part of cluster 5 table. Other 3 clusters has only one district

	Neighborhood	Work/residential district	Price categories	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
8	Beskudnikovsky District	residential	mass-segment	Cosmetics Shop	Liquor Store	Moving Target	Soccer Field	Park
11	Biryulyovo Zapadnoye District	residential	mass-segment	Dog Run	Soccer Field	Big Box Store	Bus Stop	Campground
13	Brateyevo District	residential	mass-segment	Supermarket	Pizza Place	Pharmacy	Concert Hall	Pedestrian Plaza
17	Chertanovo Yuzhnoye District	residential	mass-segment	Breakfast Spot	Restaurant	Pedestrian Plaza	Drugstore	Auto Workshop
26	Golovinsky District	residential	mass-segment	Waterfall	Park	Lake	Bus Stop	Middle Eastern Restaurant
27	Golyanovo District	residential	mass-segment	Park	Martial Arts Dojo	Café	Cosmetics Shop	Convenience Store
32	Khoroshyovo-Mnyovniki District	residential	mass-segment	Park	Caucasian Restaurant	Supermarket	Japanese Restaurant	Massage Studio
34	Khovrino District	residential	mass-segment	Bus Station	Campground	Supermarket	Auto Workshop	Park
37	Kosino-Ukhtomsky District	residential	mass-segment	Shopping Mall	Movie Theater	Sporting Goods Shop	Gym / Fitness Center	Big Box Store
38	Kotlovka District	residential	mass-segment	Park	Café	Bus Stop	Dance Studio	Restaurant
41	Kuntsevo District	residential	mass-segment	Bus Stop	Park	Convenience Store	Spa	Cosmetics Shop

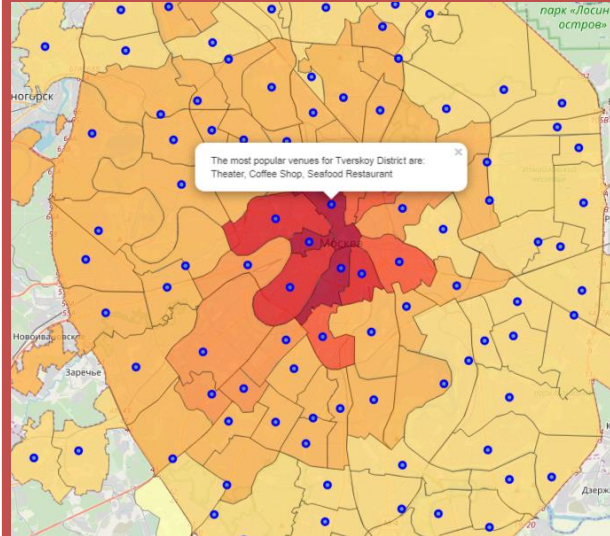


Now it's easy to identify what venues are the most popular for every district

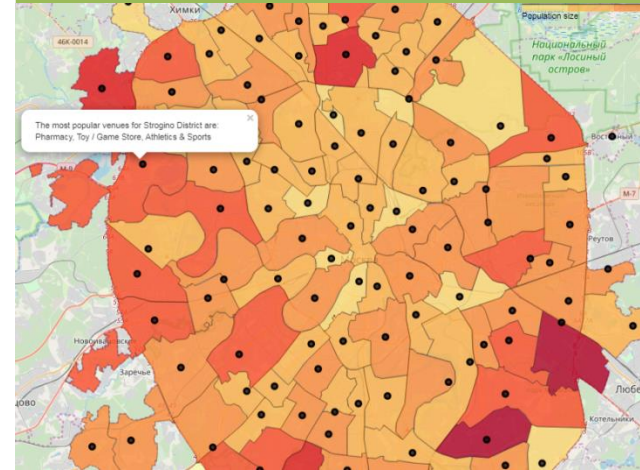
with work/residential district division



with real estate price category division



with division by population



Thank you for your attention

As there is no possibility to see folium maps in github you can look my code with folium maps visible via ibm cloud link https://eu-gb.dataplatform.cloud.ibm.com/analytics/notebooks/v2/50d4a814-4a6c-4ec7-a92e-f2099b0d599d/view?access_token=0e5b4335867c37441ff32a36a6d2a66adfd05c9c8cd7d24b61ba667d469b8e84

