Venues Data Analysis

Valeria Kovergovich

October 20, 2020

1. Introduction:

1.1 Background/Description

We have a client who'd like to open a restaurant in Minsk, Belarus. We would need to analyze the city and its neighborhoods to find the best location to do so.

Minsk is the largest city in Belarus with the populationof over 2 million people a population density of 5 792 people per square kilometer. As a resident of this city, I decided to use Minsk in my project. The city is divided into 9 districts in total.

The district we're interested in is the Frunze district. It has 10 neighborhoods and there are a lot of venues located there. Therefore, we should to analyze this district we're interested in in order to figure out the best location for a new cafe. The preferred neighborhood will have a lower number of venues as compared to other neighborhood.

In order to find a solution, we'll need to create a map and cluster neighborhoods according to the venue density and analyze it.

1.2 Data Description

I used the data on districts and neighborhoods of Minsk including its coordinates (latitude and longtitude) in order to explore, analyze and cluster neighborhoods to see which one is the most appropriate for opening a new venue at this moment. I used Google Maps in order to get the coordinates of each neighborhood.

After that, I formatted the data and used it to create a map of Minsk and the Frunze district.

I used Forsquare API to get the most common venues of the Frunze District.

2. Methodology

After importing the data on the districts and neighborhoods of Minsk, I imported a data on its geospatial coordinates and merged two dataframes using Pandas which resulted in the following dataframe:

	City	Districts	Neighborhoods	Latitude	Longitude
0	Minsk	Frunze	Kamennaya Gorka	53.922072	27.428566
1	Minsk	Frunze	Zapad	53.898199	27.450072
2	Minsk	Frunze	Sukharevo	53.885941	27.433085
3	Minsk	Frunze	Kuntsevschina	53.911466	27.444215
4	Minsk	Frunze	Masyukovschina	53.925189	27.491349

In order to visualize my data, I used the folium library in Python and created a map of Minsk at first:
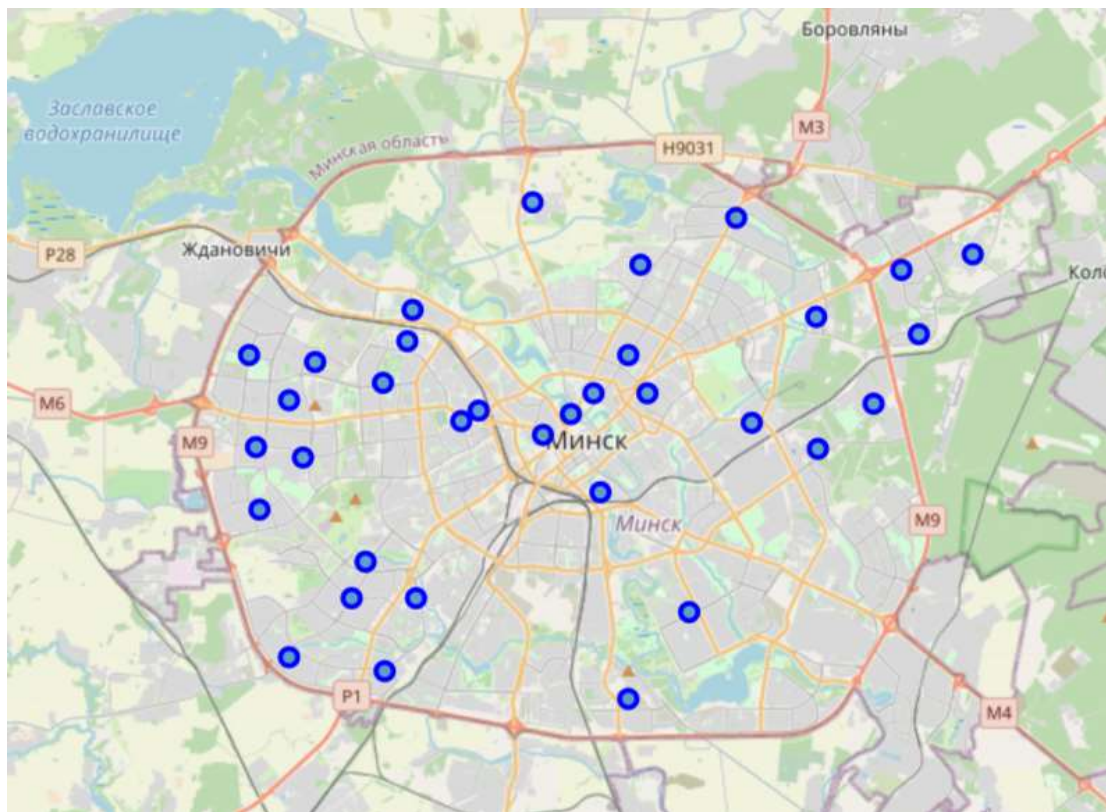
After that, I created another map of the Frunze district:
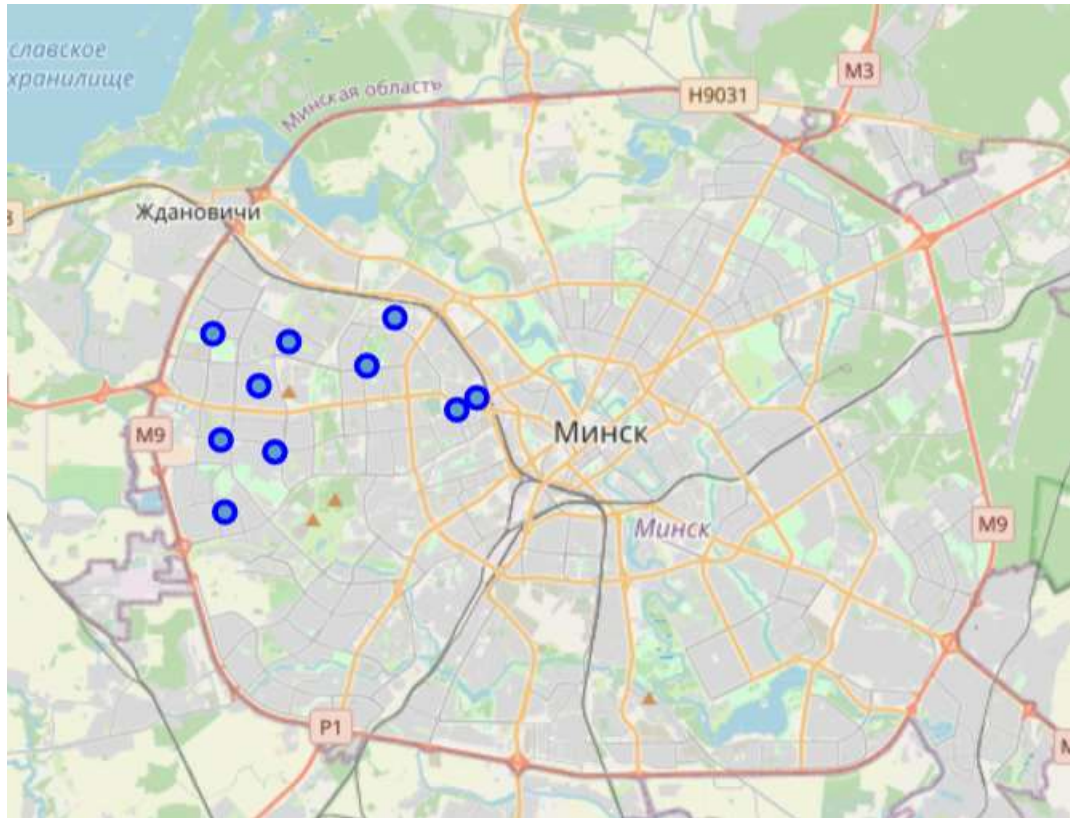
My next step was to use the Foursquare API in order to explore the neighborhoods in the Frunze district. The limit was set to 100 and the radius to 500 meters:

	Neighborhoods	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Kamennaya Gorka	53.922072	27.428566	Батутная арена Джой джамп	53.923985	27.429068	Sports Club
1	Kamennaya Gorka	53.922072	27.428566	Евроопт Hyper	53.917617	27.429462	Big Box Store
2	Kamennaya Gorka	53.922072	27.428566	Бойцовский клуб К-2 ЖД	53.923118	27.425009	Martial Arts School
3	Kamennaya Gorka	53.922072	27.428566	Стадион	53.922671	27.432494	Soccer Field
4	Zapad	53.898199	27.450072	Фрунзенский физкультурно-оздоровительный центр	53.894694	27.451526	Gym / Fitness Center

The analysis returned 51 unique values. In summary of this data 43 venues were returned by Foursquare. Here is a part of the table that includes the unique venues in the Frunze district:

|	Neighborhoods	Arcade	Athletics & Sports	Auto Garage	Bakery	Beer Bar	Beer Store	Big Box Store	Bistro	Bookstore	...	Shopping Mall	Soccer Field	Sports Club	Supermarket
0	Dombrovka	0.000000	0.0	0.000000	0.0	0.0	0.125000	0.00	0.0	0.0	_	0.000000	0.00	0.00	0.000000
1	Kalvariya	0.000000	0.1	0.000000	0.1	0.1	0.000000	0.00	0.1	0.0	_	0.000000	0.00	0.00	0.000000
2	Kamennaya Gorka	0.000000	0.0	0.000000	0.0	0.0	0.000000	0.25	0.0	0.0	_	0.000000	0.25	0.25	0.000000
3	Kaskad	0.000000	0.0	0.000000	0.1	0.1	0.000000	0.00	0.1	0.1	_	0.000000	0.00	0.00	0.000000
4	Krasnyi Bor	0.000000	0.0	0.000000	0.0	0.0	0.083333	0.00	0.0	0.0	_	0.000000	0.00	0.00	0.000000
5	Kuntsevschina	0.000000	0.0	0.333333	0.0	0.0	0.000000	0.00	0.0	0.0	_	0.000000	0.00	0.00	0.000000
6	Masyukovschina	0.333333	0.0	0.000000	0.0	0.0	0.000000	0.00	0.0	0.0	_	0.000000	0.00	0.00	0.000000
7	Sukharevo	0.000000	0.0	0.000000	0.0	0.0	0.000000	0.00	0.0	0.0	_	0.000000	0.00	0.00	0.100000
8	Tivali	0.000000	0.0	0.000000	0.0	0.0	0.083333	0.00	0.0	0.0	_	0.000000	0.00	0.00	0.000000
9	Zapad	0.000000	0.0	0.000000	0.0	0.0	0.090909	0.00	0.0	0.0	_	0.090909	0.00	0.00	0.181818

After that, I created a table including the top 10 most popular venues in each neighborhood:

	Neighborhoods	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	Dombrovka	Bus Stop	Beer Store	Pharmacy	Furniture / Home Store	Food	Jewelry Store	Kids Store	Lighting Store	Cafe	Gym / Fitness Center
1	Kalvariya	Café	Bistro	Women's Store	Fast Food Restaurant	Park	Bus Stop	Gym	Athletics & Sports	Beer Bar	Bakery
2	Kamennaya Gorka	Sports Club	Soccer Field	Big Box Store	Martial Arts School	Yoga Studio	Cafe	General Entertainment	Furniture / Home Store	Food	Fast Food Restaurant
3	Kaskad	Gym / Fitness Center	Department Store	Go Kart Track	Bakery	Beer Bar	Bistro	Bookstore	Fast Food Restaurant	Gym	Cafe
4	Krasnyi Bor	Pharmacy	Bus Line	Insurance Office	Beer Store	Mobile Phone Shop	Optical Shop	Pet Store	Lake	Gourmet Shop	Wine Shop
5	Kuntsevschina	Bus Stop	Auto Garage	Dog Run	Cafe	Go Kart Track	General Entertainment	Furniture / Home Store	Food	Fast Food Restaurant	Electronics Store
6	Masyukovschina	Bus Stop	Arcade	Athletics & Sports	Convenience Store	Go Kart Track	General Entertainment	Furniture / Home Store	Food	Fast Food Restaurant	Electronics Store
7	Sukharevo	Yoga Studio	Pizza Place	Bus Station	Bus Stop	Cosmetics Shop	Gym	Kids Store	Park	Gym / Fitness Center	Supermarket
8	Tivali	Park	Gym Pool	Convenience Store	Beer Store	Cafe	Recreation Center	General Entertainment	Tennis Stadium	Tennis Court	Auto Garage
9	Zapad	Bus Stop	Supermarket	Gym / Fitness Center	Hotel	Beer Store	Electronics Store	Shopping Mall	Martial Arts School	Convenience Store	Furniture / Home Store

Subsequently, I used clustering in order to see most common venues in different clusters.

3. Results

Here's the clustered map of the Frunze district of Minsk:
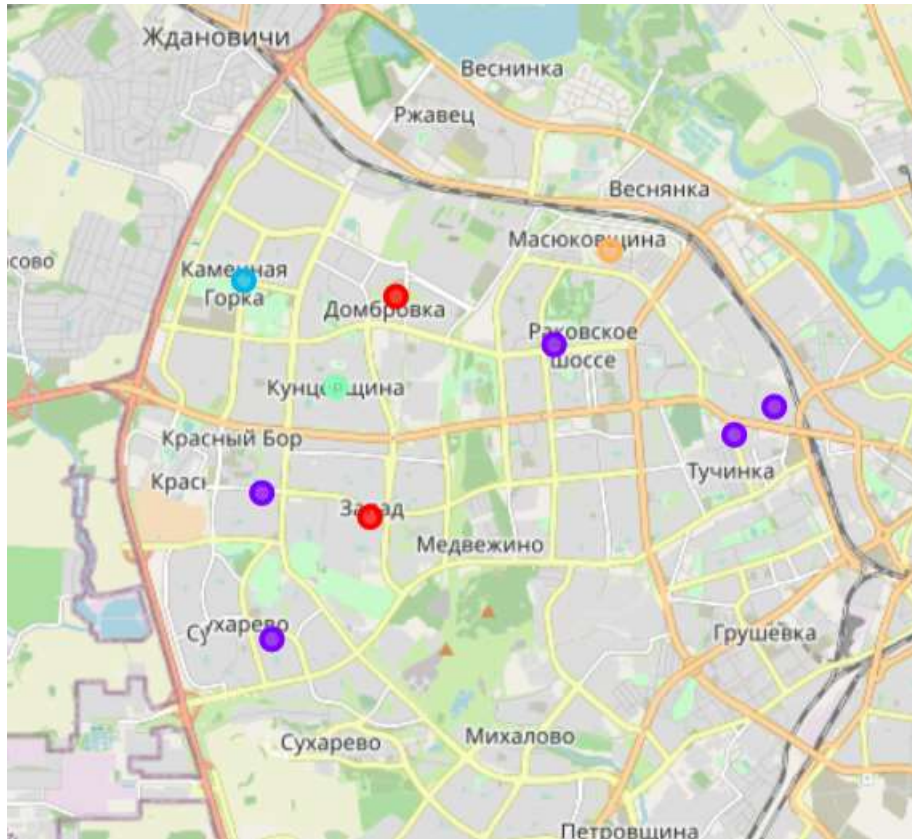
In summary section, one of my aim was also visualize the Average Housing Sale Prices for per square meter with choropleth style map. Thus, first I downloaded a json file of Second-level Administrative Divisions of the Turkey from Spatial Data Repository of NYU [2]. I cleaned the json file and pull out only city of Istanbul.

4. Discussion

Considering that Minsk is a big city with a lot of social places, it can be difficult to find a good place to open a new venue. Opening a new location without a proper preparation can be risky as there's always a strong competition in cities like Minsk.

This is the reason I performed analysis of the imported data, followed by visualization and clustering. These algorithms helped me figure out the most common venues in different neighborhoods in the Frunze district in order to see where it's better to open a new place.

5. Conclusion

In conclusion, I can say that the best neighborhood to open a new cafe would be Cluster 4 - Masyukovschina. This neighborhood doesn't have a lot of cafes as we can see from the analysis. In fact, cafes are not one of the most common places.

Opening a new place in a neighborhood that doesn't have a lot of social places will lead to an increased attention from the residents of the neighborhood thus making the business successfull.

Thank you for your attention!