**Research of Minsk bakeries**

**Volha Halynskaya**

**08 July 2020**

**Introduction**

***Background:***

When I was looking topic of research I really want to use my native town at this task and definitely want to explore topic which I interested, but the main purpose was to create research using foursquare application with some restrictions and collect the data about companies.

**“What district of Minsk really needs more bakeries? “**

This topic I decided to research also because I really like morning's croissant with fresh coffee and would like to know where’s places I should visit to find it and the same time I'm planning to move and I'll know which district have enough it now.

##### ***Target audience of this research:***

It's definitely people who interested in opening bakery and looking for district to open place. In my mind this person should consider how many bakeries Minsk has now and how peoples rated it.

Opening hours in my opinion also important, because I and my colleagues and friends really like to buy fresh croissant and morning coffee when we're going to work.

##### ***Problems:***

* Information about Minsk district and it coordinates doesn’t exist in open source and I scan it manually from yandex map and create a list myself.
* API Foursquare doesn’t have a lot of users in Belarus and data is not full.

***Plan of my research:***

* Download the data – this part is concluding creating of data frame with information about districts of Minsk and it coordinates.
* Using API Foursquare I should to know id of bakeries category for searching all company this type and full their information. \*
* Join all data into one data frame including information about each company and its district.
* Deleting duplicate rows, transforming columns.
* Analyze it and appending on map.
* Using K-Means create model and fit it on data frame.
* Give a conclusion.

\* Because the free part of API service is limited by calls.

### Data

##### ***Data description of research:***

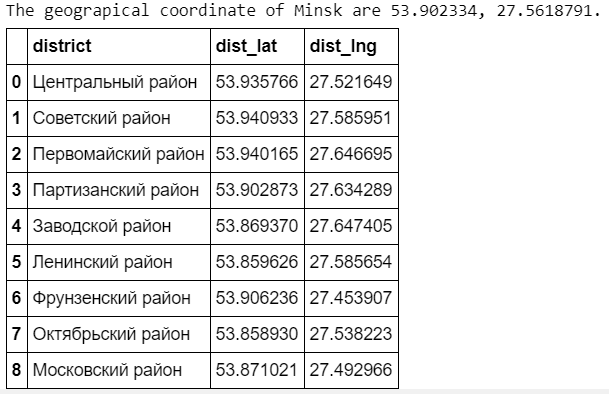
* Open source data about Minsk districts from yandex map.
* API foursquare to find all bakeries of Minsk.

I should show that in Belarus Foursquare application is not in high demand and data which I got not include all Minsk bakeries and places. The same I can consider about companies information, for example about prices.

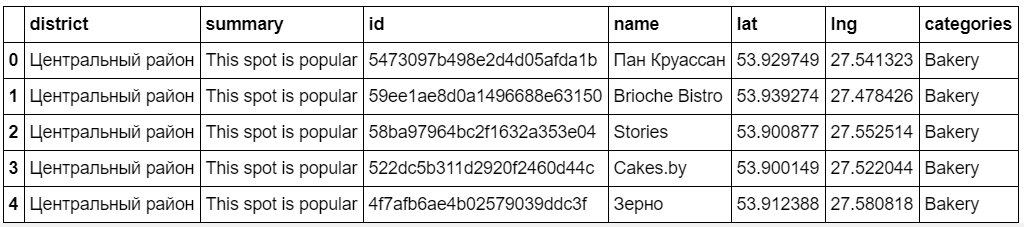
### Methodology section

Using libraries: pandas, numpy, requests, json, geodesic, translit etc. Machine learning techniques: K-means for clustering data.

Data frame Minsk districts you can see below:



Data frame bakeries of Minsk you can see below:

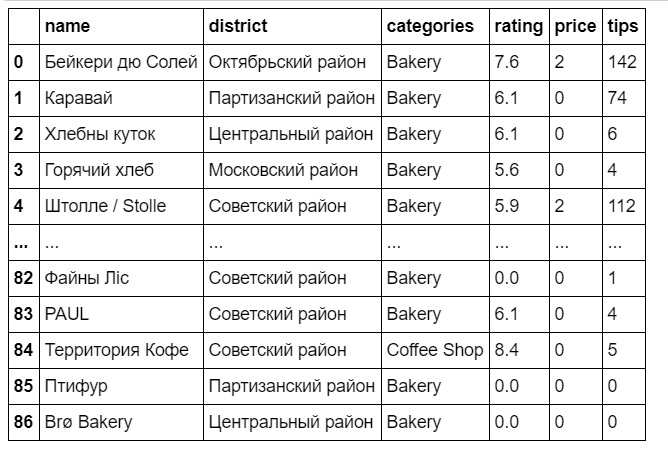


**According to plan of research:**

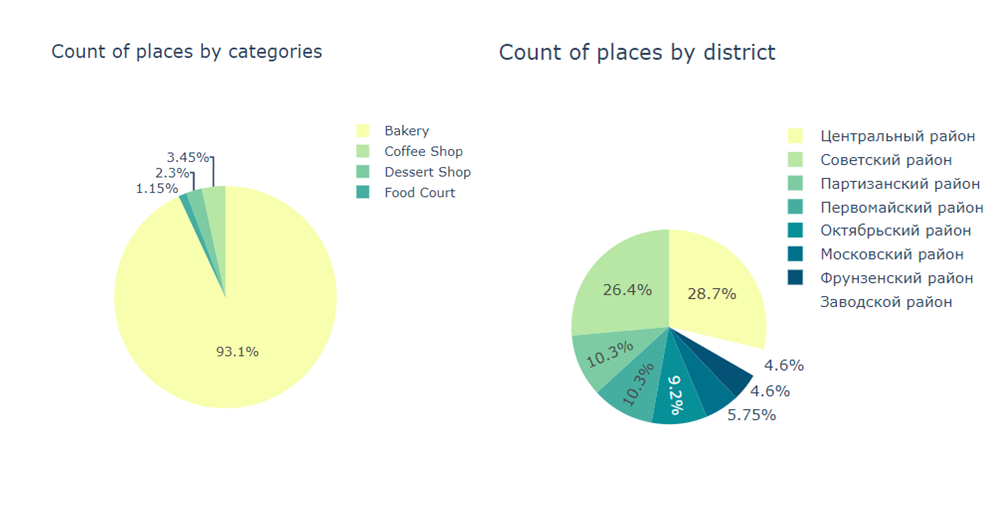
* I downloaded all names of company with category = “Bakery” from API.
* I downloaded all inside information about each company, for example rating, price etc.
* I joined 2 data frames in one using district as a key.
* I calculated the distance between central coordinate of the district and coordinate of company and delete duplicate rows.

The shape of final data frame contain (87, 21). It’s mean that my table is including 87 unique rows with 21 columns.

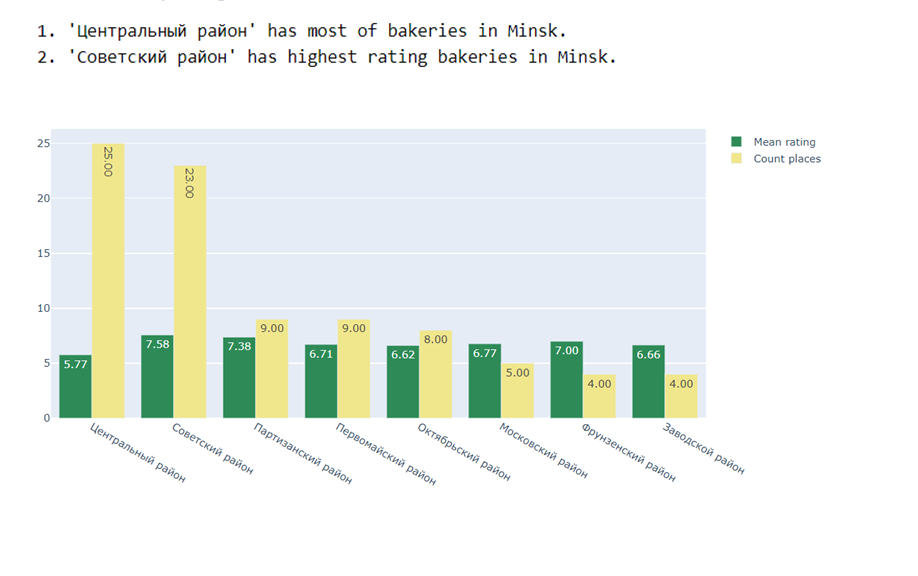
Example:



**Below you can see different plots reflective count of bakeries by each district and count of bakeries by categories from Foursquare API.**



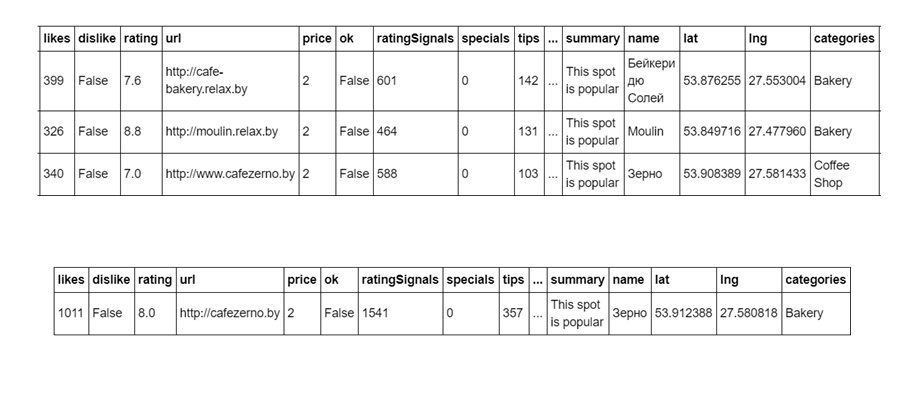
**Below you can see Mean rating and count likes of places by district:**



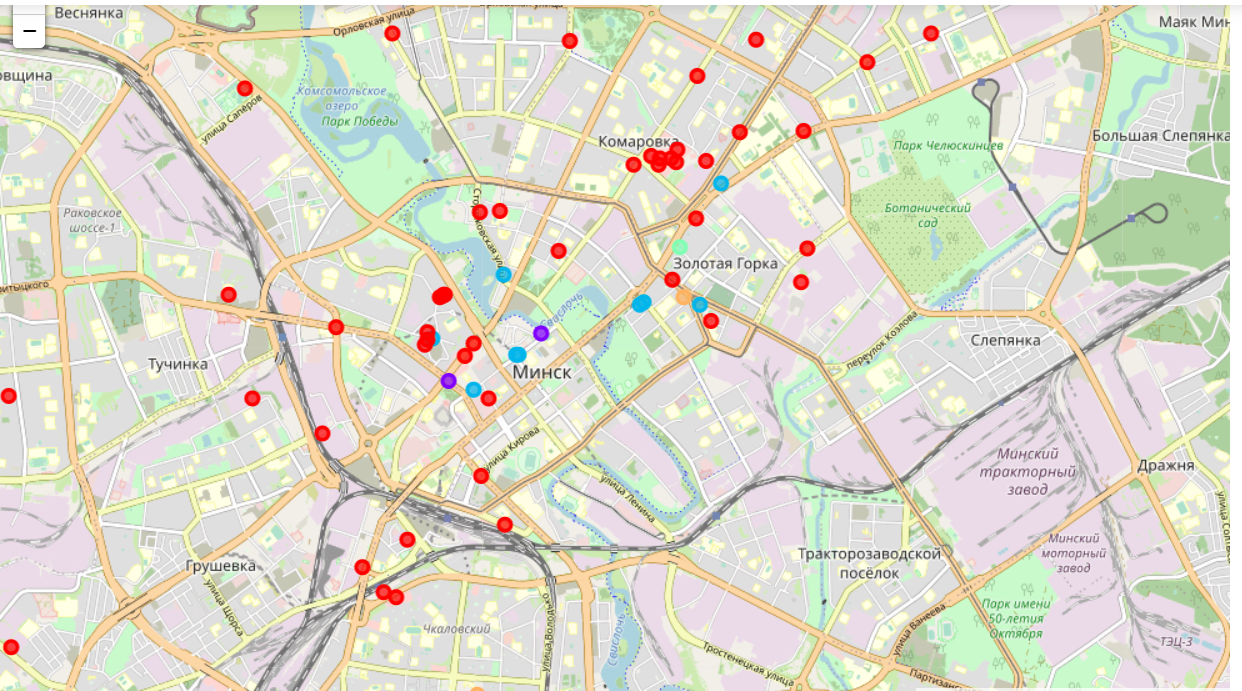
After collecting and cleaning data I used K-Means algorithm of machine learning to separate data frame by 5 clusters. Below you can see explain of clusters and example of bakeries situated on map:

Clusters 3 and 4 have the best rating, prices and likes and its orange and violet dots on map.

From 87 bakeries in Minsk just 4 of it has a really good values. Below you can see table of it:



Below you can see example of clusters situated on Minsk map:



**Results:**

Answering to the question:

**Which Minsk district really needs more bakeries?**

I should notice that most count of bakeries belong to the center of city and central district.

And it's okay because most of cultural and evening life are using there, all European cities has the same statistic.

But according to my question - bakery for me, it's firstly the morning coffee and croissant and I defiantly recommend opening bakeries not in the center of city, but in the new districts far from the center.

**Discussion:**

Really great example we can see at 'Moulin' bakery which is according to the high 4 cluster. From all 87 Minsk's bakeries just 3 from it has 4 clusters and it place is situated to new district.

**Conclusion:**

I invite everyone to Minsk to explore the best bakeries and the freshest croissants. You can visit each of these bakeries and create your own rating.