Capstone Project - The Battle of the Neighborhoods (Week 2)

Table of Contents

1. Business Problem and Background

The objective of this project is to analyze where to open a new restaurant in Bratislava (Slovakia) and recommend the type of restaurant. Bratislava has many good restaurants which are concentrated in the city center.

The data science tools will be used to find best new place where the restaurant would have a chance for good attendance, and it will be close to the City center and at the same time to recommend what type of restaurant is missing.

Target audience - Investors looking for opening a new restaurant in Bratislava, Slovakia.

2. The Data and how it will be used to solve the problem

Data / Factor to be considered is: - Number and types of existing restaurants in the neighborhood

Sources of data for the analysis: - Foursquare API as a source of data - number of restaurants, their type, location - Bratislava Neighborhoods html data from https://sk.wikipedia.org/wiki/Zoznam_ul%C3%ADc_v_Bratislave

Methods used to solve the problem: - Use Bratislava Neighborhoods data and enrich it with information from Foursquare - Create Pandas dataframe - Use K-means

3. Methodology section

3.1. Download libraries needed to perform data analysis:

The following libraries will be used: - Geopy - to locate the coordinates - Pandas - Python Data Analysis Library - Folium - to visualize data that's been manipulated in Python on an interactive leaflet map - K-means clustering

3.2. Download dataset and create Panda dataframe of Bratislava's Neighborhoods

The dataset with Bratislava's Neighborhoods is downloaded from Wikipedia (https://sk.wikipedia.org/wiki/Zoznam_ul%C3%ADc_v_Bratislave).

After data wrangling, cleansing and translating columns headers, the dataset will have four columns namely - Borough, Neighborhood, Latitude and Longitude.

Geolocation service Nominatim will be used to locate the coordinates of each Neighborhood.

	Borough	Neighborhood	Latitude	Longitude
0	Devín	Bratislava Devín	48.174586	16.982610
1	Devínska Nová Ves	Bratislava Devínska Nová Ves	48.209175	16.973894
2	Dúbravka	Bratislava Dúbravka	48.187089	17.037503
3	Jarovce	Bratislava Jarovce	48.065237	17.112942
4	Karlova Ves	Bratislava Karlova Ves	48.159240	17.052677
5	Lamač	Bratislava Lamač	48.191196	17.054279
6	Nové Mesto	Bratislava Nové Mesto	48.167162	17.138352
7	Petržalka	Bratislava Petržalka	48.110972	17.111290
8	Podunajské Biskupice	Bratislava Podunajské Biskupice	48.130042	17.207421
9	Rača	Bratislava Rača	48.211789	17.153433
10	Rusovce	Bratislava Rusovce	48.052647	17.146917
11	Ružinov	Bratislava Ružinov	48.149314	17.164505
12	Staré Mesto	Bratislava Staré Mesto	48.140589	17.112327
13	Vajnory	Bratislava Vajnory	48.204910	17.206579
14	Vinohrady	Bratislava Vinohrady	48.186776	17.133716
15	Vrakuňa	Bratislava Vrakuňa	48.145195	17.200494
16	Záhorská Bystrica	Bratislava Záhorská Bystrica	48.239288	17.038336
17	Čunovo	Bratislava Čunovo	48.029636	17.199573

3.3. Foursquare Data

The Neighborhood's Latitude and Longitude from the dataset will be used to download the data of all Food Venues from Foursquare using the Food Category ID of '4d4b7105d754a06374d81259'

To download all venues the respective categoryId will be used as per Foursquare documentation https://developer.foursquare.com/docs/build-with-foursquare/categories/

After downloading the venues in Bratislava's Neighborhoods the new databased is created with the venue's longitude, latitude and its category.

The dataset includes many places in category "Café". Those venues will be excluded in order to focus just on places with "food":

Pandas operation Groupby will be used to check the number of venues by neighborhood. Some neighborhoods have a small number of venues - opportunity for new business.

Neighborhood	Latitude	Longitude	Venue	Venue_Lat	Venue_Lon	Venue_Category
Bratislava Devín	7	7	7	7	7	7
Bratislava Devínska Nová Ves	12	12	12	12	12	12
Bratislava Dúbravka	45	45	45	45	45	45
Bratislava Jarovce	3	3	3	3	3	3
Bratislava Karlova Ves	60	60	60	60	60	60
Bratislava Lamač	45	45	45	45	45	45
Bratislava Nové Mesto	79	79	79	79	79	79
Bratislava Petržalka	86	86	86	86	86	86
Bratislava Podunajské Biskupice	23	23	23	23	23	23
Bratislava Rača	21	21	21	21	21	21
Bratislava Rusovce	9	9	9	9	9	9
Bratislava Ružinov	89	89	89	89	89	89
Bratislava Staré Mesto	75	75	75	75	75	75
Bratislava Vajnory	10	10	10	10	10	10
Bratislava Vinohrady	46	46	46	46	46	46
Bratislava Vrakuňa	26	26	26	26	26	26
Bratislava Záhorská Bystrica	14	14	14	14	14	14
Bratislava Čunovo	5	5	5	5	5	5

3.4. Visualize the data on Folium map

To visualize the Food Venues data the Folium map will be used.

3.5. Analyze Each Neighborhood

Now each Neighborhood will be analyzed using the technique "One hot encoding| where Categorical Data is transformed into Numerical Data which can be used for Machine Learning algorithms.

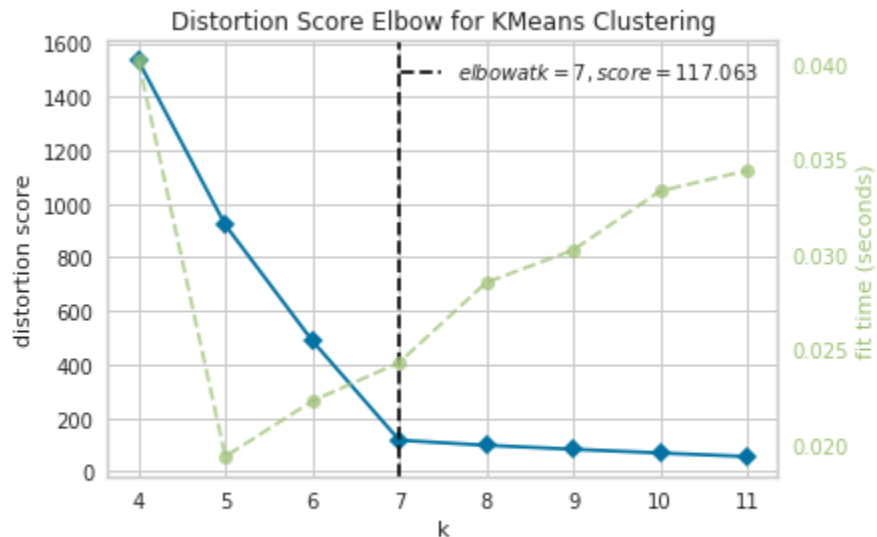|	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Bratislava Devín	Restaurant	Pizza Place	BBQ Joint	Eastern European Restaurant	Bistro	Gastropub	Food	Donut Shop	Falafel Restaurant	Fast Food Restaurant
1	Bratislava Devínska Nová Ves	Pizza Place	Restaurant	New American Restaurant	BBQ Joint	Chinese Restaurant	Soup Place	Fast Food Restaurant	Food	Vietnamese Restaurant	Eastern European Restaurant
2	Bratislava Dúbravka	Restaurant	Pizza Place	Vietnamese Restaurant	Diner	Eastern European Restaurant	BBQ Joint	Burger Joint	Fast Food Restaurant	Gastropub	Italian Restaurant
3	Bratislava Jarovce	Gastropub	Eastern European Restaurant	Pizza Place	Vietnamese Restaurant	Diner	Hungarian Restaurant	Hot Dog Joint	Greek Restaurant	French Restaurant	Food Truck
4	Bratislava Karlova Ves	Restaurant	Pizza Place	Vietnamese Restaurant	Fast Food Restaurant	Gastropub	Snack Place	Burger Joint	Italian Restaurant	Sushi Restaurant	Cafeteria
5	Bratislava Lamač	Pizza Place	Restaurant	Snack Place	Asian Restaurant	Diner	Fast Food Restaurant	BBQ Joint	Eastern European Restaurant	Gastropub	Italian Restaurant
6	Bratislava Nové Mesto	Italian Restaurant	Vietnamese Restaurant	Restaurant	Bakery	Burger Joint	Salad Place	Turkish Restaurant	Bistro	Eastern European Restaurant	Indian Restaurant

Now I create the new data frame and display the top 10 venues for each neighborhood.

	Neighborhood	Asian Restaurant	Austrian Restaurant	BBQ Joint	Bakery	Belgian Restaurant	Bistro	Breakfast Spot	Buffet	Burger Joint	Burrito Place	Cafeteria	Chinese Restaurant	Creperie	Czech Restaurant	Deli / Bodega	Diner	Donut Shop	Eastern European Restaurant	Falafel Restaurant	Fast Food Restaurant	Food	Food Court	Food Truck
0	Bratislava Devín	0.000000	0.000000	0.142857	0.000000	0.000000	0.142857	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.142857	0.000000	0.000000	0.000000	0.000000	0.000000
1	Bratislava Devínska Nová Ves	0.000000	0.000000	0.083333	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.083333	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.083333	0.083333	0.000000	0.000000
2	Bratislava Dúbravka	0.044444	0.000000	0.044444	0.000000	0.000000	0.000000	0.000000	0.000000	0.044444	0.000000	0.022222	0.022222	0.000000	0.022222	0.000000	0.066667	0.000000	0.066667	0.022222	0.044444	0.000000	0.022222	0.000000
3	Bratislava Jarovce	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.333333	0.000000	0.000000	0.000000	0.000000	0.000000
4	Bratislava Karlova Ves	0.016667	0.000000	0.000000	0.033333	0.000000	0.000000	0.000000	0.000000	0.050000	0.000000	0.033333	0.000000	0.000000	0.000000	0.016667	0.016667	0.000000	0.016667	0.016667	0.050000	0.016667	0.000000	0.000000
5	Bratislava Lamač	0.066667	0.000000	0.044444	0.000000	0.000000	0.000000	0.000000	0.000000	0.022222	0.000000	0.022222	0.022222	0.000000	0.022222	0.000000	0.066667	0.000000	0.044444	0.000000	0.066667	0.000000	0.000000	0.000000

4. Results section

4.1 Cluster Neighborhoods

I will use the Elbow Point Technique to define the optimal number of clusters. The best K value is chosen at the point in which the line has the sharpest turn. In my case, the Elbow Point is at K = 7.

: <matplotlib.axes._subplots.AxesSubplot at 0x7f08db585518>

The neighborhoods will be clustered based on their most common venues:

	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	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	Devín	Bratislava Devín	48.174586	16.982610	1	Restaurant	Pizza Place	BBQ Joint	Eastern European Restaurant	Bistro	Gastropub	Food	Donut Shop	Falafel Restaurant	Fast Food Restaurant
1	Devínska Nová Ves	Bratislava Devínska Nová Ves	48.209175	16.973894	1	Pizza Place	Restaurant	New American Restaurant	BBQ Joint	Chinese Restaurant	Soup Place	Fast Food Restaurant	Food	Vietnamese Restaurant	Eastern European Restaurant
2	Dúbravka	Bratislava Dúbravka	48.187089	17.037503	5	Restaurant	Pizza Place	Vietnamese Restaurant	Diner	Eastern European Restaurant	BBQ Joint	Burger Joint	Fast Food Restaurant	Gastropub	Italian Restaurant
3	Jarovce	Bratislava Jarovce	48.065237	17.112942	3	Gastropub	Eastern European Restaurant	Pizza Place	Vietnamese Restaurant	Diner	Hungarian Restaurant	Hot Dog Joint	Greek Restaurant	French Restaurant	Food Truck
4	Karlova Ves	Bratislava Karlova Ves	48.159240	17.052677	5	Restaurant	Pizza Place	Vietnamese Restaurant	Fast Food Restaurant	Gastropub	Snack Place	Burger Joint	Italian Restaurant	Sushi Restaurant	Cafeteria
5	Lamač	Bratislava Lamač	48.191196	17.054279	5	Pizza Place	Restaurant	Snack Place	Asian Restaurant	Diner	Fast Food Restaurant	BBQ Joint	Eastern European Restaurant	Gastropub	Italian Restaurant
6	Nové Mesto	Bratislava Nové Mesto	48.167162	17.138352	0	Italian Restaurant	Vietnamese Restaurant	Restaurant	Bakery	Burger Joint	Salad Place	Turkish Restaurant	Bistro	Eastern European Restaurant	Indian Restaurant

5. Discussion section on observations and recommendations

Cluster 0:

Cluster 0 - This analysis shows that most of restaurants are in the Neighborhoods - Stare Mesto, Nove Mesto, Ruzinov a Vinohrady. The cluster 0 is rather saturated with restaurants and there is a good diversity of all kinds of restaurants.

Recommendations for this cluster:

1. Neighborhood Ružinov - open an Italian restaurant
2. Neighborhood Nove Mesto - in Neighborhood Nove Mesto there are 2 districts - Kramare, Koliba which are high-end residential areas but there are no restaurants. One of my recommendation is to start restaurant in that arrea - it can be a luxury Italian or some general restaurant.

	Neighborhood	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
6	Bratislava Nové Mesto	Italian Restaurant	Vietnamese Restaurant	Restaurant	Bakery	Burger Joint	Salad Place	Turkish Restaurant	Bistro	Eastern European Restaurant	Indian Restaurant
11	Bratislava Ružinov	Restaurant	Gastropub	Pizza Place	Vietnamese Restaurant	Bistro	Asian Restaurant	Italian Restaurant	Burger Joint	Turkish Restaurant	Indian Restaurant
12	Bratislava Staré Mesto	Italian Restaurant	Restaurant	Burger Joint	Bistro	Vegetarian / Vegan Restaurant	Vietnamese Restaurant	Bakery	Thai Restaurant	Pizza Place	Salad Place
14	Bratislava Vinohrady	Restaurant	Italian Restaurant	Pizza Place	Vietnamese Restaurant	Bakery	Bistro	Salad Place	Burger Joint	Chinese Restaurant	Diner

Cluster 1:

Cluster 1 - The outskirts of Bratislava - small number of restaurants but there are also smaller number of inhabitants. No recommendation for this cluster.

	Neighborhood	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	Bratislava Devín	Restaurant	Pizza Place	BBQ Joint	Eastern European Restaurant	Bistro	Gastropub	Food	Donut Shop	Falafel Restaurant	Fast Food Restaurant
1	Bratislava Devínska Nová Ves	Pizza Place	Restaurant	New American Restaurant	BBQ Joint	Chinese Restaurant	Soup Place	Fast Food Restaurant	Food	Vietnamese Restaurant	Eastern European Restaurant
16	Bratislava Záhorská Bystrica	Restaurant	Pizza Place	Snack Place	BBQ Joint	Tapas Restaurant	Bakery	Gastropub	Food	Eastern European Restaurant	Falafel Restaurant

Cluster 2:

Cluster 2 - the Bratislava outskirts - small number of restaurants but both neighborhoods have quite good number inhabitants.

Recommendations for this cluster: open an Italian restaurant in any of these two neighborhoods.

	Neighborhood	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	Bratislava Podunajské Biskupice	Pizza Place	Eastern European Restaurant	Restaurant	Slovak Restaurant	Food Court	Diner	Fast Food Restaurant	Vietnamese Restaurant	Bistro	Food Truck
15	Bratislava Vrakuňa	Pizza Place	Eastern European Restaurant	Restaurant	Vietnamese Restaurant	Slovak Restaurant	Food Court	Fast Food Restaurant	Diner	Italian Restaurant	Bistro

Cluster 5:

Cluster 5 - the most of Neighborhoods are in this cluster. General restaurants and Pizza places are most common food venues. These are mostly residential Neighborhood.

Recommendations for this cluster: open an Italian restaurant in Neighborhood Petržalka.

	Neighborhood	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	Bratislava Dúbravka	Restaurant	Pizza Place	Vietnamese Restaurant	Diner	Eastern European Restaurant	BBQ Joint	Burger Joint	Fast Food Restaurant	Gastropub	Italian Restaurant
4	Bratislava Karlova Ves	Restaurant	Pizza Place	Vietnamese Restaurant	Fast Food Restaurant	Gastropub	Snack Place	Burger Joint	Italian Restaurant	Sushi Restaurant	Cafeteria
5	Bratislava Lamač	Pizza Place	Restaurant	Snack Place	Asian Restaurant	Diner	Fast Food Restaurant	BBQ Joint	Eastern European Restaurant	Gastropub	Italian Restaurant
7	Bratislava Petržalka	Pizza Place	Restaurant	Asian Restaurant	Bistro	Eastern European Restaurant	Bakery	Gastropub	Fast Food Restaurant	Vegetarian / Vegan Restaurant	Buffet
9	Bratislava Rača	Restaurant	Italian Restaurant	Pizza Place	Diner	Asian Restaurant	Slovak Restaurant	Food Court	Eastern European Restaurant	Bistro	Belgian Restaurant

6. Conclusions

Objective of this project was to analyze food venues in Bratislava and come with recommendation for a new restaurant.

This analysis used data science techniques. Final decision will require analysis also from financial, environmental, traffic and other perspectives.

Considering the above data I would recommend the following options :
1. Opening of Italian restaurant Neighborhood Nove Mesto in district Koliba

2. Opening Italian restaurant in Neighborhoods - either in Ruzinov or Petrzalka.