Analyzing and comparing different venues in Budapest and Prague

Author: Imre Kovacs, **imre.kov@gmail.com**

1. Introduction

Budapest and Prague are two quickly developing cities in East Europe. Their sizes are not too different, Budapest is the 10th largest city in the EU by population, while Prague is the 14th. Both cities are tourist hotspots, welcoming millions of visitors every year. They have similarities in their history, too, as both have been the part of the Austro-Hungarian Empire and both has been ruled by communist governments in the second half of the 20th century.

After 1990, Budapest and Prague became fast growing capitals, with more and more business opening, quickly learning how to adjust to the new capitalist era.

Budapest is the capital of Hungary, with population of 1.7 million people, living both sides of the Danube. The city has 23 districts, its overall size is 525.2 km2. Budapest's GDP is above Hungarian and EU average, 52,770 USD per capita.
In 2018, the city welcomed around 12 million international visitors.

Prague is the capital of the Czech Republic, population is 1.3 million people, at the banks for the Vltava river. The city has 22 districts, its size is 496 km2. Prague's GDP is above the national average with 62500 USD per capita.
In 2017, the city welcomed 8.4 million international visitors.

As seen from the data above, both cities are well above their own national and EU average in terms of financial availabilites. Their per capita GDP can be compared e.g. to Berlin, which has a thriving hospitality sector. The same can be said about these two cities, too, both because of their own population and the incoming tourists' needs.

2. Problem

I would like to compare the districts of the two cities, based on their venue types and categories. As it has been explained before, the two cities has a lot in common, both in their history and actual development. As per Tripadvisor, Budapest has around 1200 restaurants and around 300 cafes. Prague has almost 2000 restaurants and 600 coffee places.
I would like to compare the two cities, mainly their districts, based on their hospitality venues (restaurants, coffee shops). I would like to find similarities between the cities' districts, but also would like to find major differences, if there are any.

I will visualize the findings, both in lists and maps, trying to make it even more understandable.

The result would be really usable for a company, which currently has venues in on of the cities and would like to set up shop in similar districts of the other city. Also, as I will list the number of different venues in each cities and districts, it can be a good baseline where to open a new one.

Key objectives of analyzing the data set

1, Which are the most similar districts of Budapest and Prague, based on their venues?
2, Which are the biggest differences between districts of the two cities?
3, Which are the districts with the least and most venues?

3. Data Acquisition and sources

3.1 Data sources
I used the following data for this project:

- Districts of Budapest and Prague, with their latitudes and longitudes.
  Source: Wikipedia, I created two CSV files for them, as they are not available at one page for easy automated scrapping. Both CSV files contain all the districts of the cities with their center's coordinates.

- Restaurants and coffee shops in each district of the cities.
  Source: Foursquare API

4. Methodology

- I will load the districts geo data from the mentioned CSV files.
- Using Foursquare API, I will get all the needed venues for each district.
- I will do exploratory data analysis.
- I will create clusters of the districts based on their hospitality venues.
- The clusters will be visualized on maps using Folium and in lists, pointing out their similarities and differences

4.1 Exploratory Data Analysis
Using the free tier of the Foursquare API, I managed to get almost all of the venues of the two cities. I got 1059 venues of Prague and 894 of Budapest.
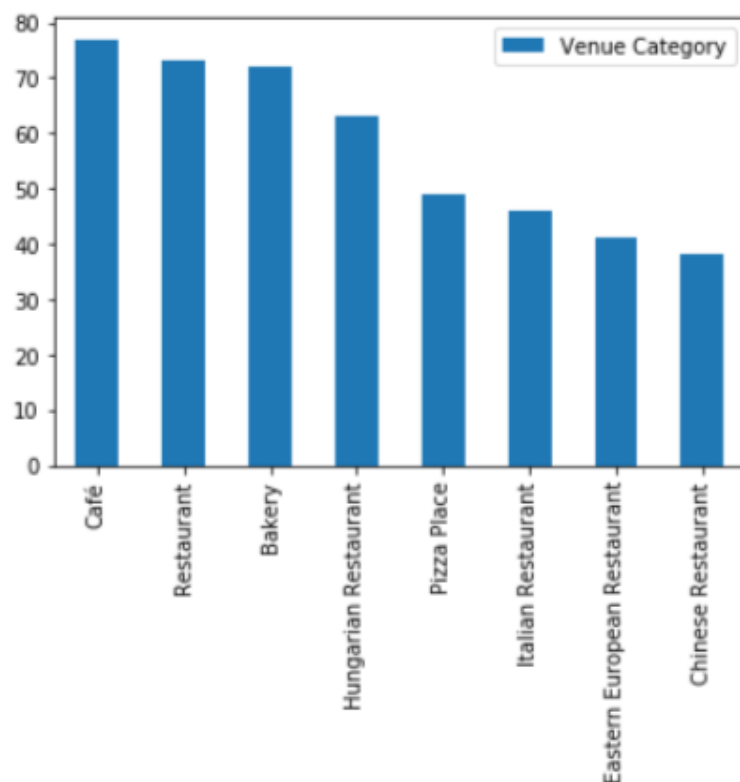
The top venue types for both cities look like this:

Type of venues in Budapest	Number of them
Café	77
Restaurant	73
Bakery	72
Hungarian Restaurant	63
Pizza Place	49
Italian Restaurant	46
Eastern European Restaurant	41
Chinese Restaurant	38
Burger Joint	29
Diner	25

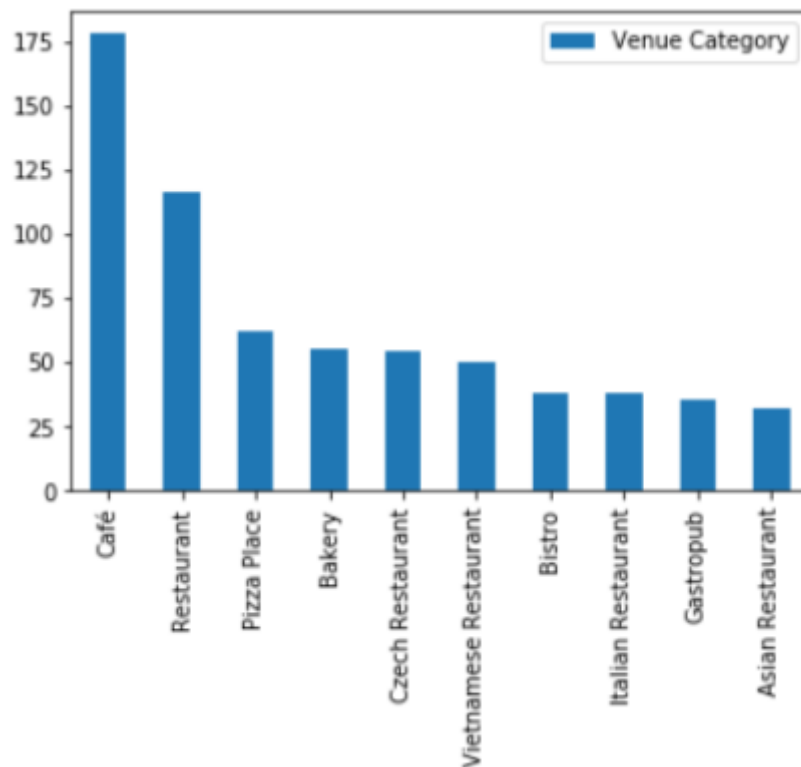| Type of venues in Prague	Number of them
Café	178
Restaurant	116
Pizza Place	62
Bakery	55
Czech Restaurant	54
Vietnamese Restaurant	50
Bistro	38
Italian Restaurant	38
Gastropub	35
Asian Restaurant	32

Some interesting results can be seen right from these two tables:

- While in both cities, the „Cafe" category is the most numerous, it is almost 2.5 times more represented in Prague. Given the same size of the cities and almost the same size of the datasets, this is an interesting difference. It should mean that Budapest still not has enough coffee places.

- Budapest has much more bakeries than Prague (72 > 55). As Hungarians are more of a „bread-eating" nation, this is not a surprise, but still, it is clearly supported by the data.

- Both cities have their local, national restaurant types in the top 5, but at neither places it is in the very top positions. Given the sizes of the cities, it can be assumed, that Prague could do more Czech restaurants.

- In Prague, there are 50 Vietnamese restaurants, while in Budapest, they are not even in the top 10. I checked the detailed dataset and in there are only 21 of those in the Hungarian capital. The reason most likely, is the bigger Vietnamese population in Prague, which is the third biggest minorities in the Czech Republic.
-

Above are the Budapest top venue types, where the type has more than 30 shops.

The same for Prague:

Given that I also created a list of the top 10 most common venues for each districts, I checked that, too for interesting findings. To save space, I only show top 5 here from some districts of Budapest.
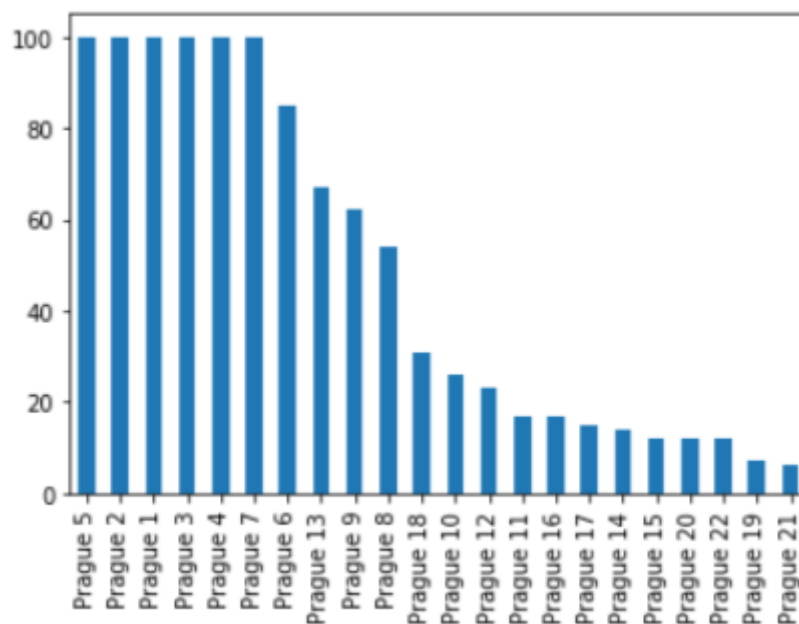
District	1st common	2nd	3rd	4th	5th
1.	Hungarian Restaurant	Café	Restaurant	Italian Restaurant	Eastern European Restaurant
4.	Fast Food Restaurant	Café	Gastropub	Bakery	Chinese Restaurant
5.	Italian Restaurant	Hungarian Restaurant	Restaurant	Eastern European Restaurant	Modern European Restaurant
15.	Pizza Place	Greek Restaurant	Bakery	Food	Donut Shop

Interesting things:

- In the first districts, the Hungarian restaurants are the most numerous of the venues. Given that this district is the Castle-district, with the most of the sightseeing and tourists, it is understandable that the local national kitchen is so well represented.

- The fifth district, the inner city, has the most of the Italian restaurants with a lot of Hungarian ones, too. The reason is twofold: while this is another tourist hotspot, which explains the Hungarian venues, it is also the main business and governmental area of the city and Hungarians tend to consider Italian food as a luxury – high quality one, so one naturally open one where there is market with money for it.
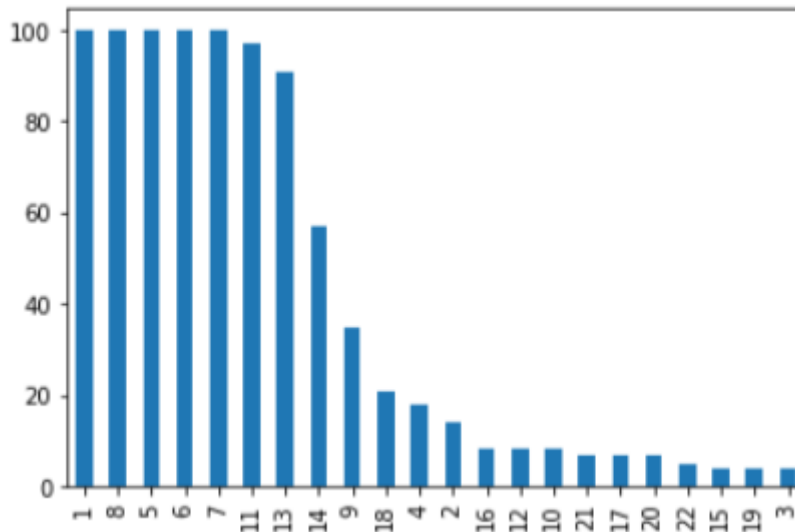
Finally, I created two bar plots to see where are the most / least venues in each cities.

Prague:

Here, the 1,2,3,4,5 districts have the most venue, which are the center / tourists ones, while the 19 and 21 have the least, which are both small districts, the furthest from the center.

Budapest:

In Budapest, districts 1,5,6,7,8 has the most venues, these are mainly the center ones, the only exception is the 8th, but that is one of the biggest, not too far from the center. 3 and 19 are with the least venues, 19 is a remote district, while the 3 is an old village, which does not have too much venues.
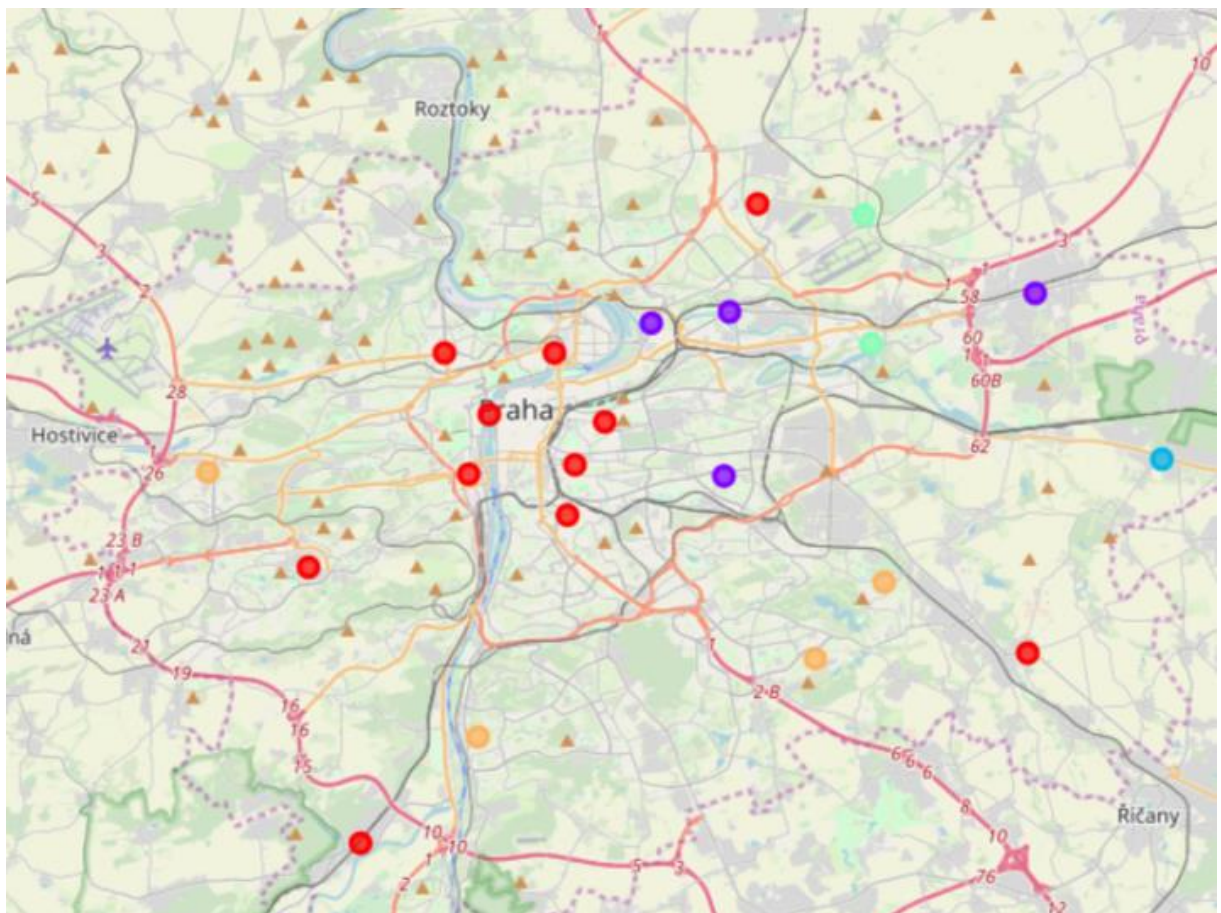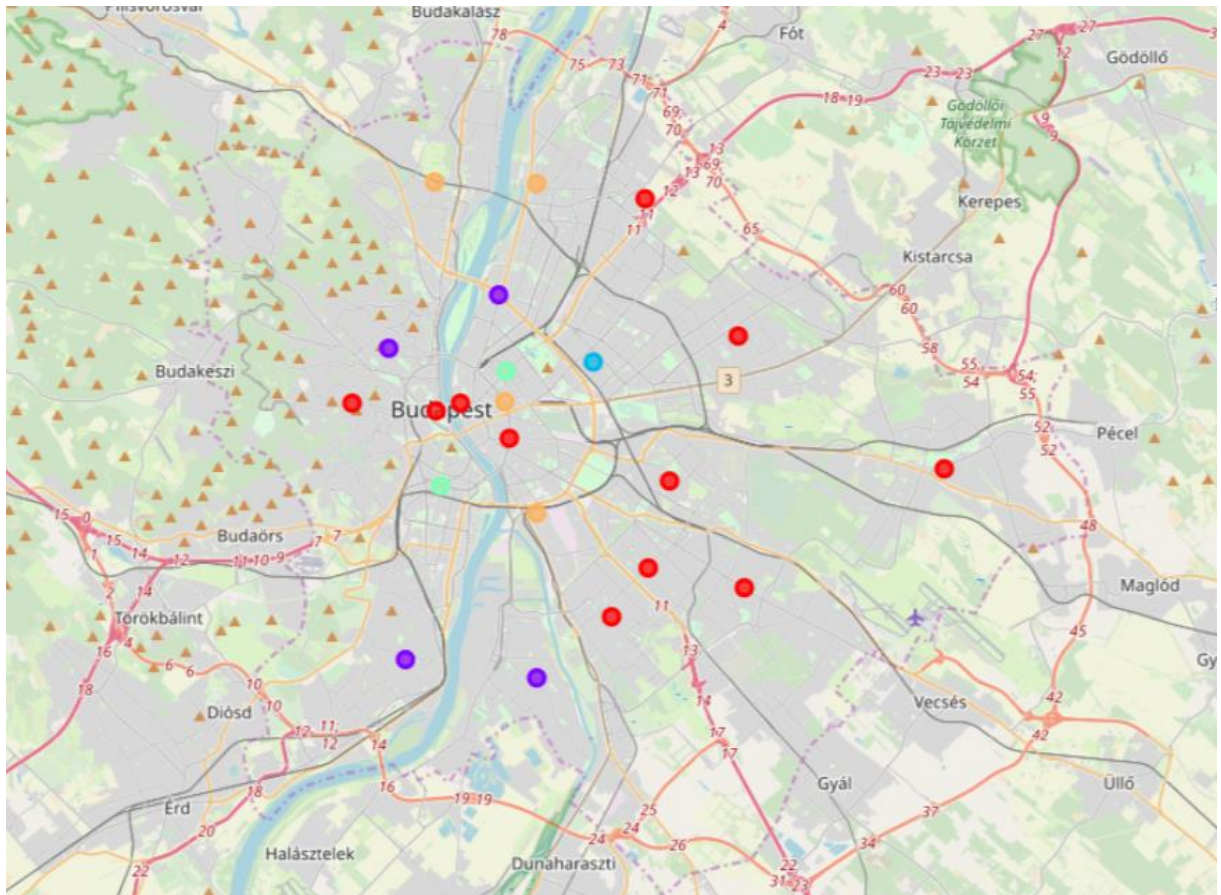
4.2 Clustering

Based on the data from Foursquare, I created the top 10 venues types on district level, using the geodata csvs and inputs for locations.

I set up a k-means clustering with 5 clusters and used the same clustering algorithm on both cities results.

The results are below.

Cluster	Number of districts	Exact districts	
		Budapest	Prague
0	11	1,5,8,10,12,15,16,17,18,19,20	1,2,3,4,5,6,7,13,16,18,22
1	4	2,13,21,22	8,9,10,20
2	1	14	21
3	2	6,11	14,19
4	4	3,4,7,9	11,12,15,17

Using Folium, I put the clusters for both cities to map:

As it can be seen, the cluster 0 type of districts are usually the center ones, even more so for Prague, where almost all of them can be found in the middle of the city. For Budapest, this cluster, while occupies the city center, it spreads to the eastern section of the city, too.

For both cities, cluster 1 types of districts are the main living ones, where the population resides and not really visited by tourists.

Cluster 2 type, while small for both cities, in Budapest, it is the district with the biggest population.

Cluster 4 seems to be the easy-going category, with cafes, fast-food restaurants, pizza places – those are the party districts in both cities.

5. Results / Discussion

Most of the results have been written above in the exploratory and clustering section, but I am summarizing those here.

1, Budapest has relatively few coffee places, compared to its size and tourist rate. Opening one in one of the center areas is still a good idea.

2, Prague could do with more local national restaurants, there are not enough Czech restaurants around, in the same sized Budapest there are much more Hungarian ones.

3, While there is numerous restaurants in the first district of Budapest, there are very few bakeries. Given that bakeries are very popular in Hungary, there should be more there to serve the tourists.

4, The result of the clustering can be seen above as well, in the table. While it is not guaranteed that just based on the clustering factors (top 10 common venues) we can make a decision, it can be a good start to explore the other city. E.g. if we have a business in one of the cluster's districts in Prague, it worth first checking the matching ones in Budapest.

6. Conclusion

I hope this paper can help entrepreneurs to some extent to understand both cities, their differences and similarities. Given their same size, background and economical power, there were many similarities between Prague and Budapest.

One way to improve this report would be to include ratings of the venues as well. Right now, because of the limitations of the free tier of the Foursquare API, it was not possible to implement.