

Analyzing and comparing different venues in Budapest and Prague

Imre Kovács



Both cities are booming tourist hotspots

- **Prague:** 8.4 million visitors a year.
- **Budapest:** 12 million visitors a year.

However, differences in their hospitality business exist and should be analysed.

E.g.: which are the most similar districts of Budapest and Prague, based on their venues?

or

Which are the districts with the least and most venues?

Data Acquisition and sources

- Districts of Budapest and Prague, with their latitudes and longitudes.
 - **Source:** Wikipedia
- Restaurants and coffee shops in each district of the cities. Managed to get 1059 venues for Prague and 894 for Budapest.
 - **Source:** Foursquare
- Disclaimer: given that I have access only to free-tier data, results can be limited.

Methodology used

- Loaded the districts geo data from the mentioned CSV files.
- Using Foursquare API, got all the needed venues for each district.
- Exploratory data analysis has been done.
- I created clusters of the districts based on their hospitality venues.
- The clusters were visualized on maps using Folium and in lists, pointing out their similarities and differences.

Exploratory Data Analysis results

| 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 |

Exploratory Data Analysis results

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.



Exploratory Data Analysis results

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.



Exploratory Data Analysis results

While there are 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.



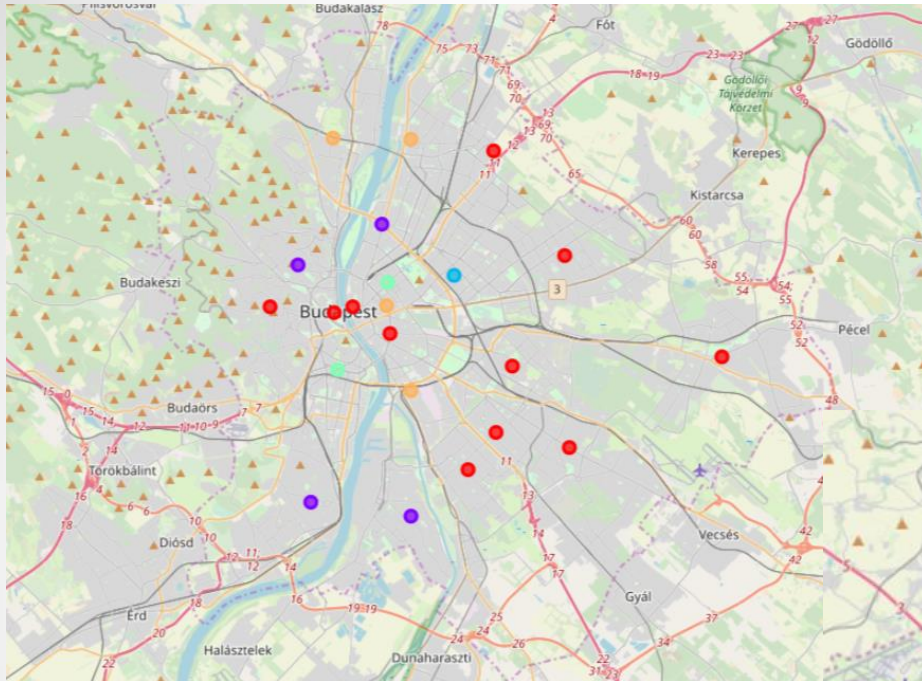
Clustering

- Based on the data from Foursquare, I created the top 10 venues types on district level.
- I set up a k-means clustering with 5 clusters and used the same clustering algorithm on both cities results.

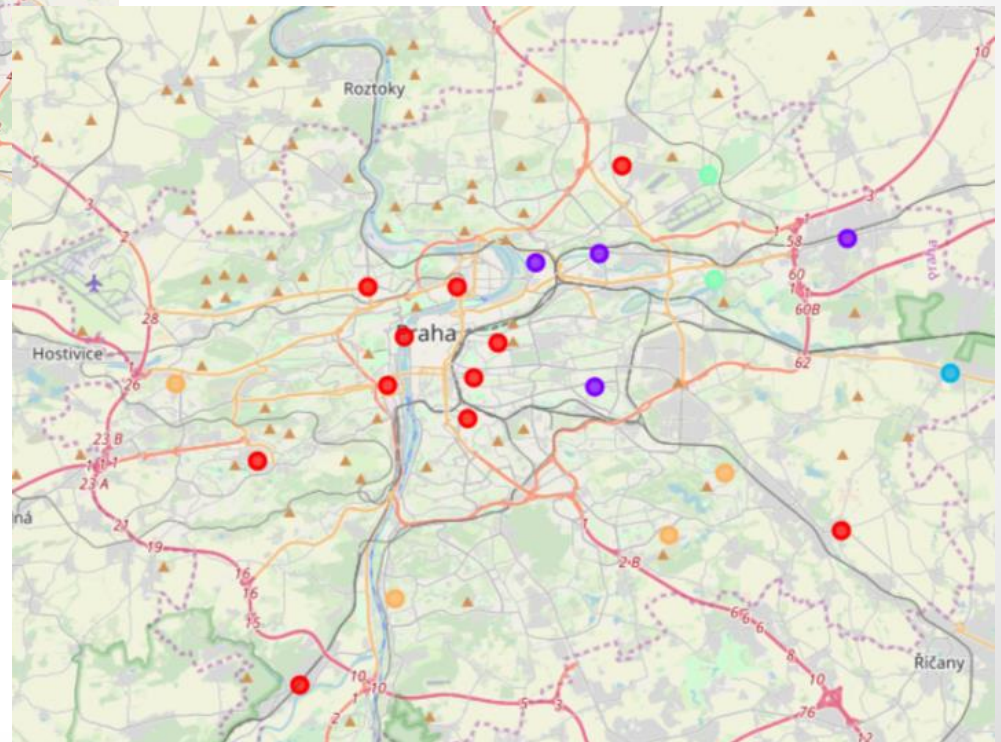
| Cluster | Number of districts | Exact districts | |
|---------|---------------------|-------------------------------|---------------------------|
| | | Budapest | Prague |
| 0 | 11 | 1,5,8,10,12,15,16,17,18,19,20 | 1,2,3,4,5,6,7,13,16,18,22 |
| 1 | 4 | 2,13,21,22 | 8,9,10,20 |
| 2 | 1 | 14 | 21 |
|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

Clustering on map

- Budapest



Prague -



Clustering on map

Different clusters have been identified:

- **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.

Conclusion and future directions

- 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.
- Also, including local GDP data of the given districts would help to properly balance the financial side of such calculations.