COURSRA CAPSTONE PROJECT - Prague

Description of the problem

1. How much immigrants influence Prague local food industry?

Prague is a a beautiful city, based in the heart of Central Europe.

Unlike others European cosmopolitan cities, with approximately 1.2 M of inhabitants, Prague is not as large. Yet, according to local authorities number of immigrants has been growing lately, helping Prague to become one of the most beautiful multicultural cities. I myself have moved to Prague from foreign country 15 years ago and thus, I am interested to know more details, such as: from which part of world the immigrants has been coming to Prague and has been local food industry influenced by their presence? In other words, in preferred districts, are we able to identify venues with some country specifics? I recognize that to answer the latter question would require much more complex analysis, but for this capstone project, I will try to identify specific venues, restaurants that are somehow related to country of origin of the top 5 immigrant communities.

Hypothesis: Potential business opportunity could be related to food industry. Foursquare data could help me to get some general information about: 1) food industry competition/dominated cuisine within specific area and 2) potential gap in local food market that is related to foreign country specific cusine.*

*I admit this approach is highly simplified but for this exercise I hope, good enough.

2. Data source used

For this project I will be using following sources:

- Czech Statistical Bureau - open database where data related to Foreigners can be found
  https://www.czso.cz/csu/czso/otevrena-data-v-katalogu-produktu-csu?filtr=true&skupiny=29&vlastnosti=22
- http://www.geoportalpraha.cz – source offering free Prague related json data
- Wikipedia
- Foursquare – API to get the current venues in Prague

3. Methodology

All data sources used are stored on GitHub for reference.

- First step pf the analysis provides closer look at immigrant distribution in Prague. Main source of this data is Czech Statistical Office - https://www.czso.cz/. Please note that I was not able to identify any relevant data related to proportion of immigrant living in specific Prague districts, thus, I am working with overall numbers only.

Table1. – TOP 5 foreign communities in Prague

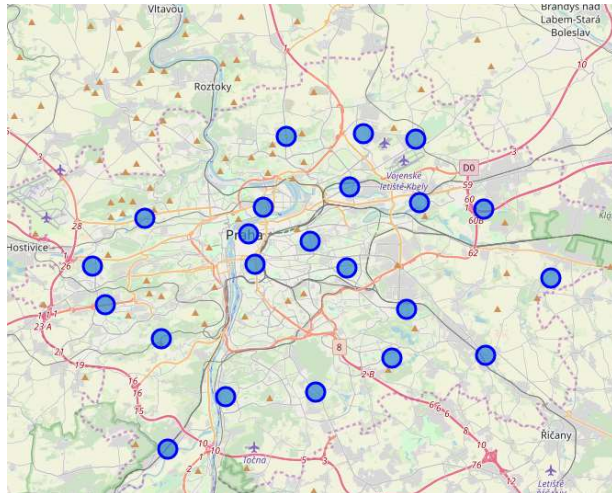|	Ctry	2016	2017	Total
0	UA	189112	194924	384036
1	SK	116272	121164	237436
2	RU	89028	91864	180892
3	VN	48848	50444	99292
4	CN	17120	19292	36412

From available data we can see that TOP 5 largest foreign communities in Prague can be split in two groups:

A) Easter Europe , B) Asia

Now, my goal it be to find out, whether we can see traces of Slavic and/or Asian cuisine in Prague food market, where the races are strongest and where are potential gaps for food industry investors.

So let's visualized Prague districts, that I am going to analyze.
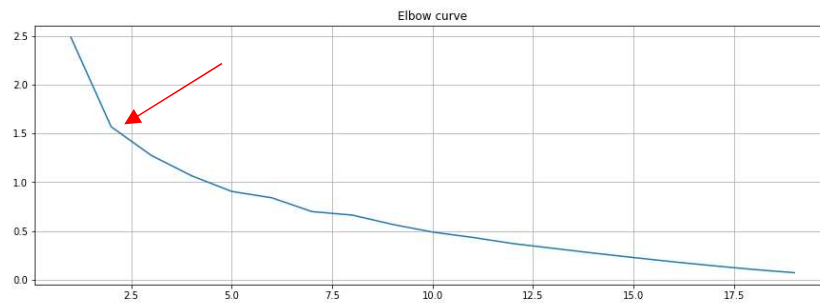
Picture1. – Prague Districts visualization

After setting up Forsquare API connection, I was able to identify There are 240 unique categories from many industries, including food industry venues. Thus, I have explored the data a little bit, and search for all venue categories that contained string "restaurant". That allowed me to get rid of venues that were beyond mt interest.

After cleansing the data, I ended up with 211 unique items. Example of the most common restaurant per Prague district is shown on table below:
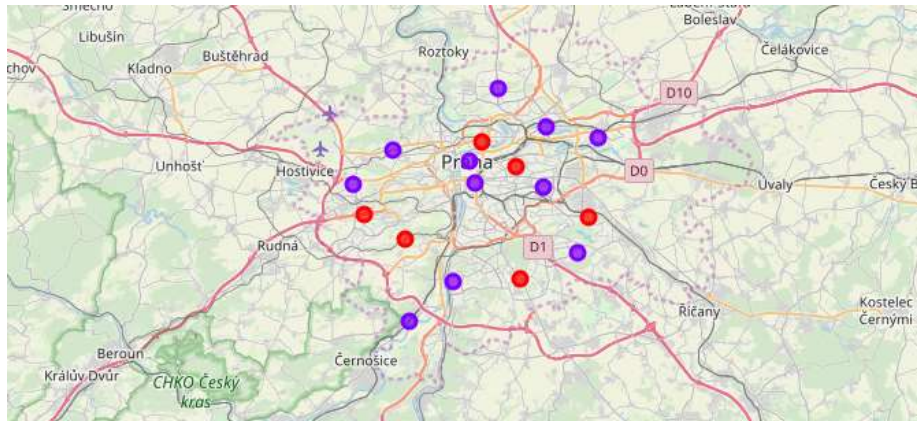
Table2. – example of most common restaurants in Prague Districts

Prague_District	1st Most Common Restaurant	2nd Most Common Restaurant	3rd Most Common Restaurant	4th Most Common Restaurant	5th Most Common Restaurant	6th Most Common Restaurant	7th Most Common Restaurant	8th Most Common Restaurant	9th Most Common Restaurant	10th Most Common Restaurant
Prague 1	Italian Restaurant	Restaurant	Vietnamese Restaurant	Brazilian Restaurant	Sushi Restaurant	Czech Restaurant	Mexican Restaurant	Asian Restaurant	French Restaurant	Vegetarian / Vegan Restaurant
Prague 10	Restaurant	Vietnamese Restaurant	Chinese Restaurant	Czech Restaurant	Mexican Restaurant	Kebab Restaurant	Sushi Restaurant	Mediterranean Restaurant	Doner Restaurant	Scandinavian Restaurant
Prague 11	Restaurant	Czech Restaurant	Turkish Restaurant	Italian Restaurant	Fast Food Restaurant	Chinese Restaurant	Doner Restaurant	Indian Restaurant	French Restaurant	Eastern European Restaurant
Prague 12	Restaurant	Italian Restaurant	Mediterranean Restaurant	Chinese Restaurant	Fast Food Restaurant	Doner Restaurant	Indian Restaurant	French Restaurant	Eastern European Restaurant	Vietnamese Restaurant
Prague 13	Restaurant	Fast Food Restaurant	Czech Restaurant	Italian Restaurant	Scandinavian Restaurant	Chinese Restaurant	Vietnamese Restaurant	Seafood Restaurant	Sushi Restaurant	Brazilian Restaurant

As a next step I wanted to crate cluster. In order to decide the right right k, I applied the elbow method and decide to set k=2:

Elbow curve

Clusters are visuzalied on the picture below. The blue color corresponds to first cluster dominated by Vietnamese cuisine. This cluester is nicely distributed across the whole city.

Conclusion

Restaurants identified within Prague District were split in two clusters.
Prague Districts grouped into first Cluster (blue) are dominated by Asian Cuisine, more specifically by Vietnamese one. Within majority of east Prague Districts, Vietnamese restaurants are the most common ones.

The Eastern European Cuisine (that embrace also Slovak, Ukraine and Russia food) is not as common at all. Few identified Easter European restaurants are concentrated in second cluster and not holding first positions at all.

When it comes to Asian cuisine, competition in Prague is quite strong. If someone wanted to open a new Asian restaurant, then Prague 8, Prague 12, Prague 14, Prague 15, Prague 18 could be areas for consideration (red cluster)
Estaern European Cuisine is quite specific, thus any Prague District could be considered.