The Pizza in Prague

Tomasz Biały

BUSINESS PROBLEM

Prague is capital of Czech Republic and it is one of the most visited cities in Europe by tourists. It has huge opportunities and prospects for opening a restaurant business. Our stakeholder is willing to open the pizza restaurant. In this project I am going to find the most optimal place to open a pizza restaurant in a city of Prague.

DATA COLLECTION

In order to solve the business problem described previously, I am going to collect and analyze the following data:

- Names of districts in Prague. The data is collected by parsing the *Wikipedia* page of Prague using the *Beautiful Soup* package.
- Geographic location of districts. The data is collected using the *Nominatim* library to find the geographic coordinates by name and address (geocoding).
- Name, category and geographic location of the existing food venues in Prague. The data can be retrieved with *Foursquare* - a location data provider with information about venues and events within an area of interest that can be obtained through the API.

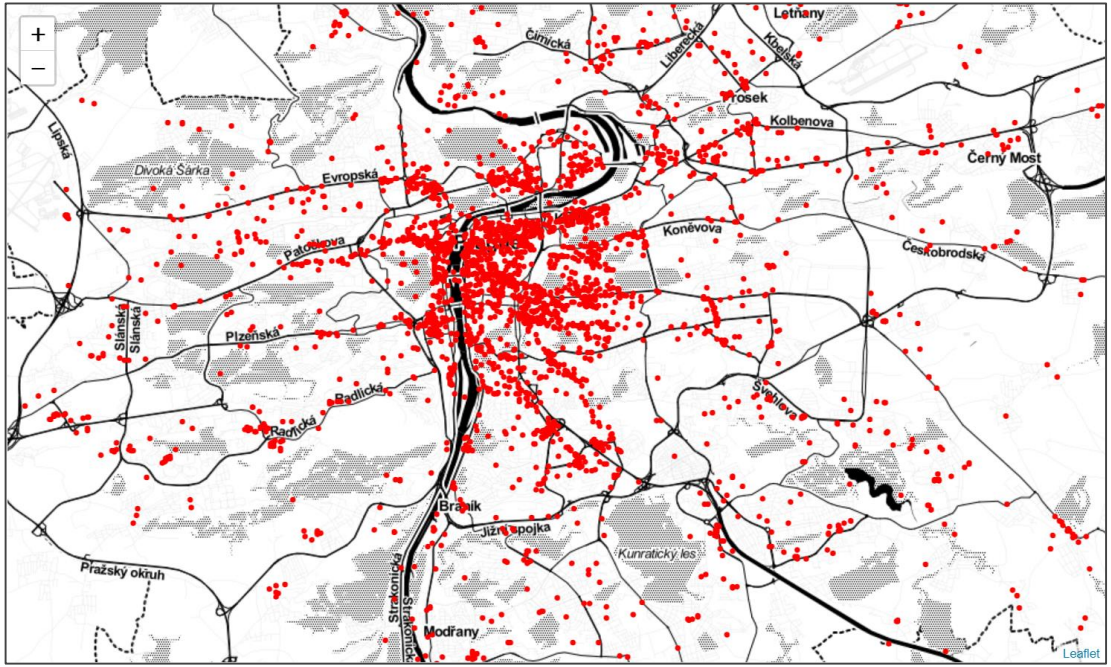
EXPLORATORY DATA ANALYSIS

The total of 3286 food venues information has been obtained and stored in dataframe during the data collection.

Venue_category	Venue_name	Venue_latitude	Venue_longitude
Vietnamese Restaurant	DuHa	50.098531	14.399085
Café	Kavárna Alibi	50.097803	14.396178
Indian Restaurant	Indian by Nature II	50.098503	14.402786
Gastropub	U Veverky	50.098991	14.402154
Pelmeni House	Bistro Váleček	50.099056	14.402783
...	...	...	...
Café	Trdlocafé	50.108025	14.581490
Café	Café Coffee Day Emporio	50.108480	14.584162
Restaurant	Sconto restaurace	50.111445	14.582952
Pizza Place	Pizzeria Gattino	50.008910	14.427877
Turkish Restaurant	Kebab House Modřany	50.008784	14.427084

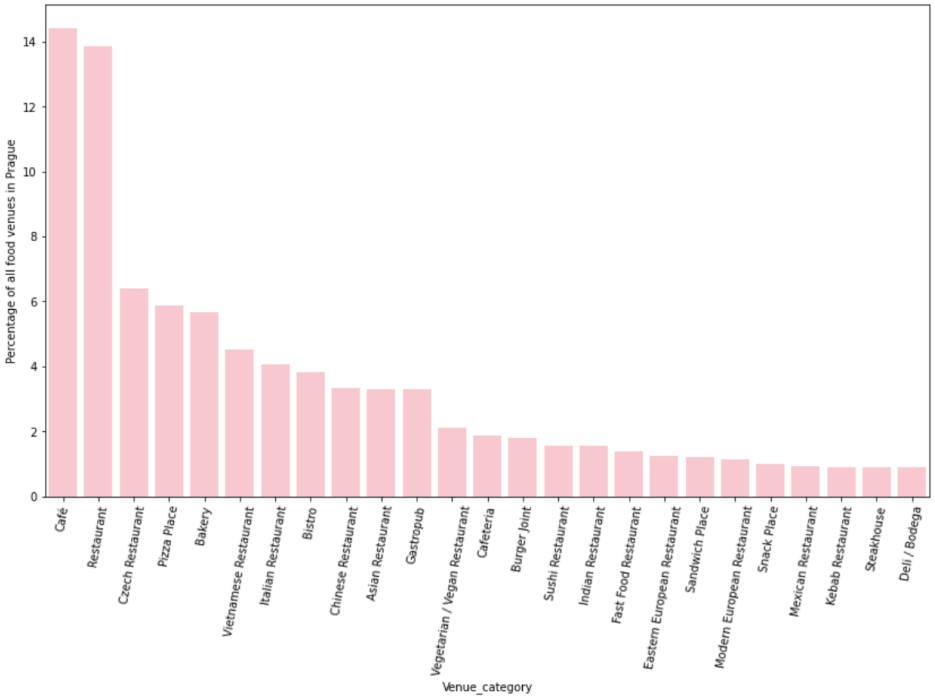
EXPLORATORY DATA ANALYSIS

It would be useful to visualize the location of the obtained venues on the map of Prague.

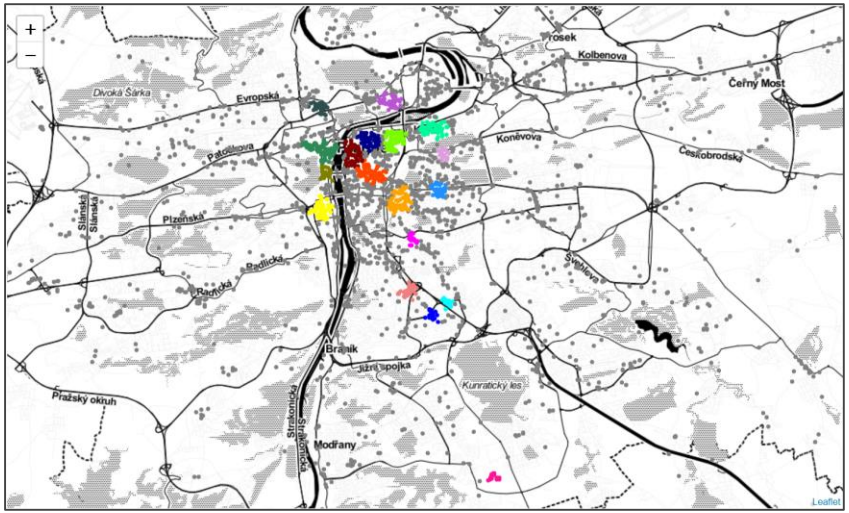


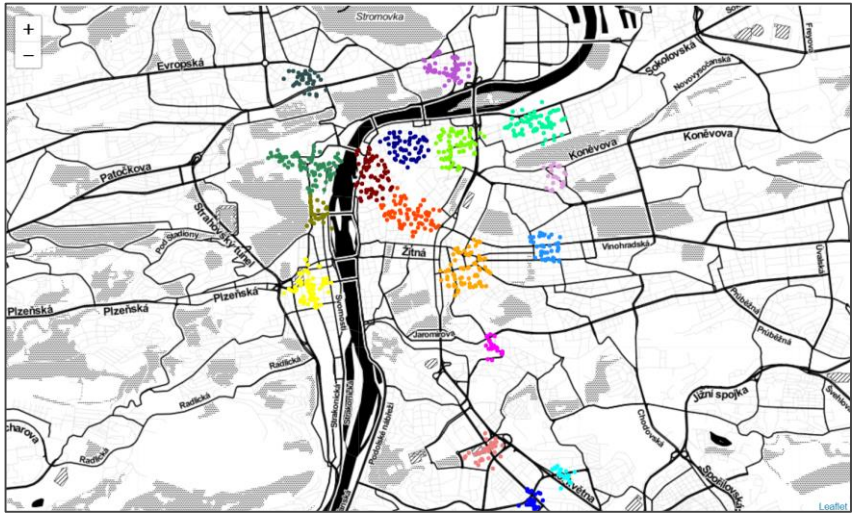
EXPLORATORY DATA ANALYSIS

Another thing that is worth checking is the popularity of specific types of the food venues.



I will cluster the locations of existing restaurants and food venues using DBSCAN clustering algorithm to find out where are the popular spots for eating out.



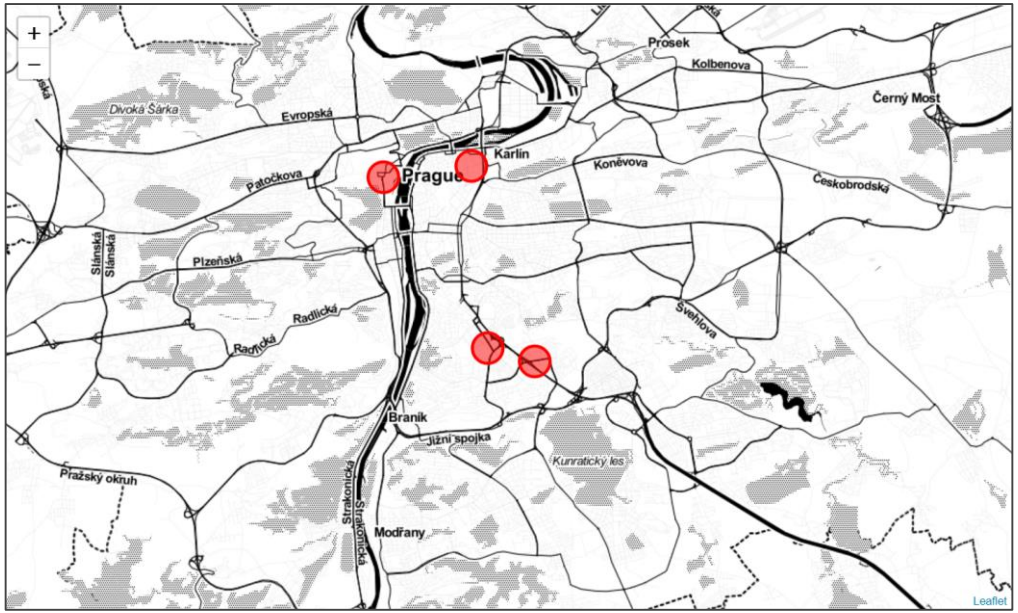


In the next step, I will check how many pizza restaurants are in each cluster in order to see how big is the competition there.

```
Cluster #1      Total places: 104      Pizza restaurants: 1.0%
Cluster #6      Total places: 84       Pizza restaurants: 4.8%
Cluster #10     Total places: 77       Pizza restaurants: 2.6%
Cluster #8      Total places: 75       Pizza restaurants: 0.0%
Cluster #5      Total places: 74       Pizza restaurants: 5.4%
Cluster #7      Total places: 72       Pizza restaurants: 4.2%
Cluster #2      Total places: 68       Pizza restaurants: 1.5%
Cluster #9      Total places: 64       Pizza restaurants: 4.7%
Cluster #4      Total places: 59       Pizza restaurants: 5.1%
Cluster #0      Total places: 42       Pizza restaurants: 2.4%
Cluster #13     Total places: 42       Pizza restaurants: 0.0%
Cluster #15     Total places: 41       Pizza restaurants: 2.4%
Cluster #3      Total places: 37       Pizza restaurants: 2.7%
Cluster #17     Total places: 36       Pizza restaurants: 2.8%
Cluster #12     Total places: 31       Pizza restaurants: 3.2%
Cluster #14     Total places: 27       Pizza restaurants: 22.2%
Cluster #11     Total places: 25       Pizza restaurants: 0.0%
Cluster #16     Total places: 25       Pizza restaurants: 4.0%
```

In order to avoid competition from businesses that serve same type of food, the recommended spots will be the ones that have the least pizza restaurants.

Finally, the recommended spots for opening a pizza restaurant are shown in the map below.



CONCLUSION

The basic data analysis was performed to identify the most optimal places to open the pizza restaurant in Prague. The clustering helped to highlight popular places where people go to eat out and categorizing the venues discarded places with the competition. Finally, the Malostranské náměstí, Náměstí Republiky, Pankrác and Michle areas were chosen as the most attractive options for setting a business.

Thank you