Coursera Capstone Project – The Battle of Neighborhoods. Final Project

Gym in Administrative district of Prague

Introduction

- We want to open a new gym in Prague. For this business to be successful, we must find a good place. We will use Foursquare's data to determine which administrative district in Prague is opening gyms. This is necessary in order to try to find the administrative area with the least number of gyms.

Business Problem

- Prague is the capital of the Czech Republic.
  Population: 1.3 million (2018).
  The main political, economic and cultural center of the Czech Republic. A major tourist center in Europe.
  Moreover, the city has a compact size. Only 500 square kilometers.
- The dense city life leads to the emergence of bad habits: alcohol, unhealthy food, constant work sitting, laziness to train at home. And this has a negative effect on health. Nowadays, many people are forced to limit themselves to visiting sports grounds. The pandemic will not last forever (I hope), which means that people will again start visiting sports grounds and playing sports. Now is the time to think about your investment in sports. Namely, in which district of Prague you can open a new gym.
- **Target audience:** New Business who wants to open a new gym in Prague.

Data

- List of Administrative district from Wikipedia page - https://en.wikipedia.org/wiki/Districts_of_Prague

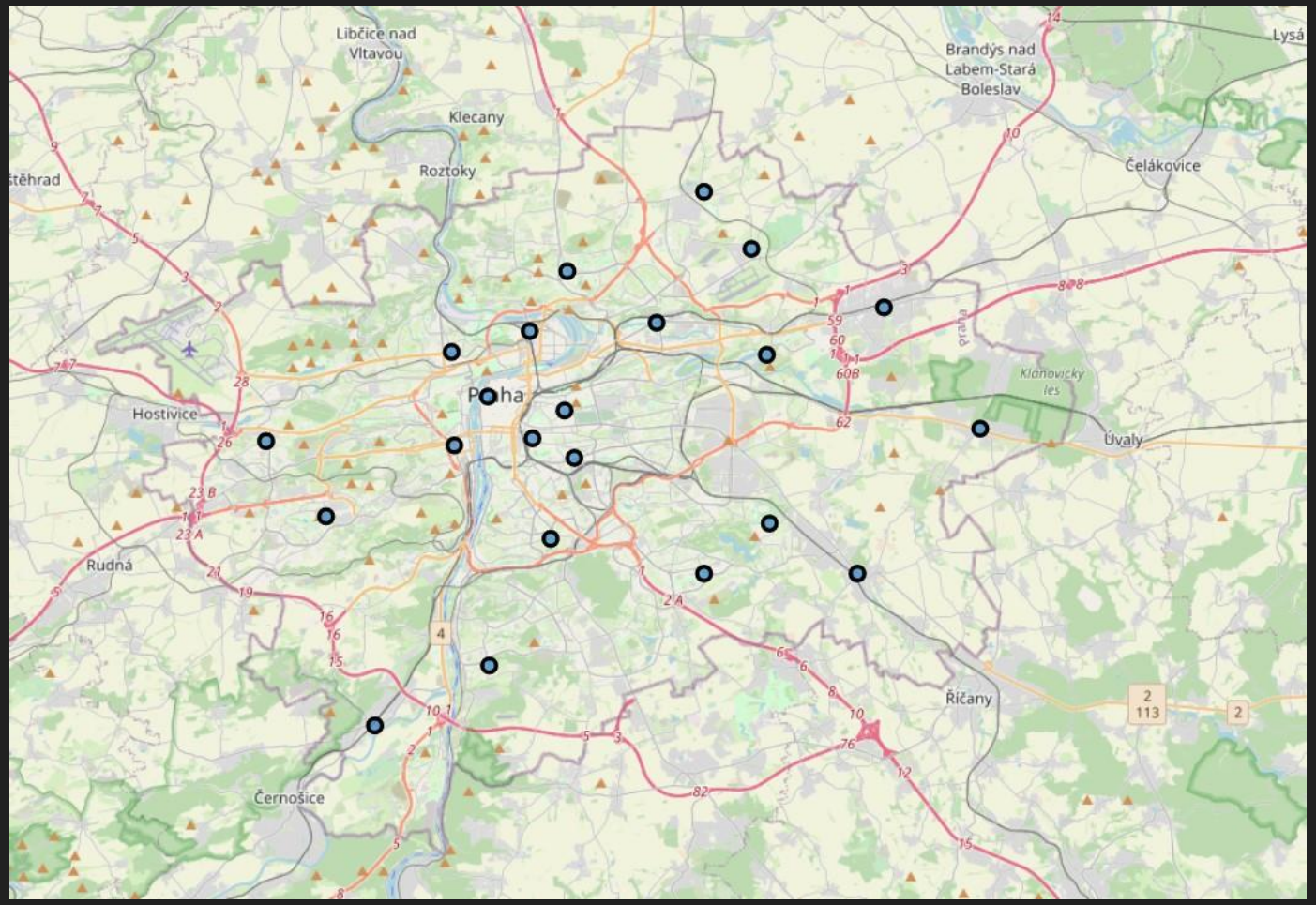
Methodology

- Parsing website with Python and special libraries – Request and Beautiful soup.
- Get the geographical coordinates in the form of latitude and longitude with Geocoder libraries.

	Administrative district	Latitude	Longitude
0	Prague 1	50.08728	14.41742
1	Prague 2	50.07394	14.43956
2	Prague 3	50.08260	14.45537
3	Prague 4	50.04232	14.44812
4	Prague 11	50.03094	14.52406
5	Prague 12	50.00201	14.41810
6	Prague 5	50.07167	14.40098
7	Prague 13	50.04925	14.33801
8	Prague 16	49.98318	14.36187
...			

Methodology

- Create a map of Prague and mark the centres of administrative districts.
- Use the Foursquare API to explore neighborhoods and segmentation.



Methodology

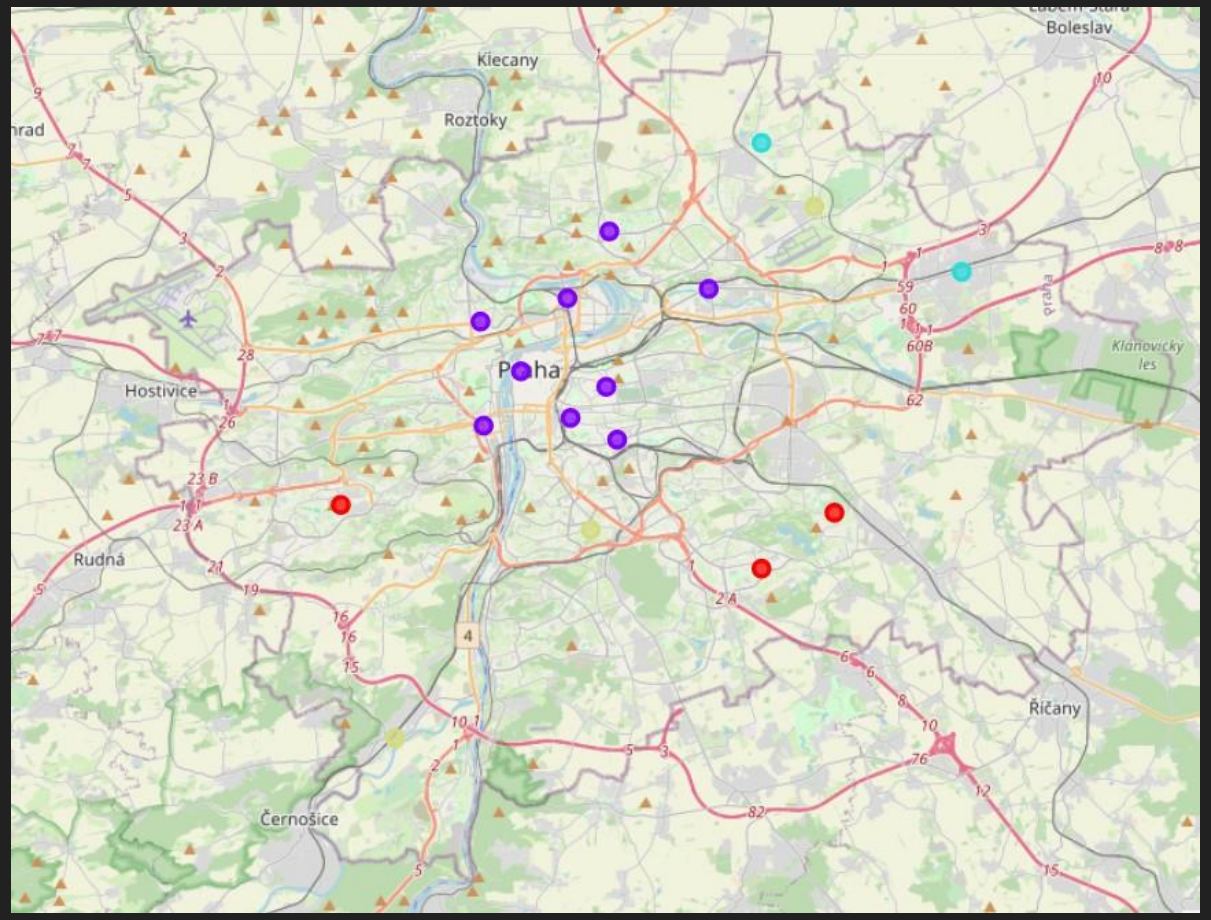
	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	GYM	GYM Latitude	GYM Longitude	Category
0	Prague 1	50.08728	14.41742	Jan Kareš Fitness	50.088716	14.423769	Gym / Fitness Center
1	Prague 1	50.08728	14.41742	Gym at Four Seasons Hotel Prague	50.087694	14.414198	Gym
2	Prague 1	50.08728	14.41742	Posilovna Právnické fakulty UK	50.091199	14.417353	College Gym
3	Prague 1	50.08728	14.41742	Sport klub U Divadla	50.085571	14.423083	Gym
4	Prague 1	50.08728	14.41742	Fitness Týn	50.088765	14.423742	Gym / Fitness Center

- Analyze each area and group the data

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Prague 1	Gym / Fitness Center	Gym	Yoga Studio	Pilates Studio	College Gym
1	Prague 10	Gym	Gym / Fitness Center	Yoga Studio	Pilates Studio	Martial Arts School
2	Prague 11	Gym	Yoga Studio	Pilates Studio	Martial Arts School	Gym Pool
3	Prague 13	Gym	Yoga Studio	Pilates Studio	Martial Arts School	Gym Pool
4	Prague 15	Gym	Yoga Studio	Pilates Studio	Martial Arts School	Gym Pool

Results

- Our research shows that despite the small number of gyms in Prague, there are areas with low or no gym density, quite close to the city center. We have 17 districts with a gym and 5 districts without a gym within a radius of 500 meters from the center of Prague with geographic coordinates 50.0864234 latitude and 14.4156772 longitude.
- The best solution would be to open a new gym in districts that don't have one at all. Prague is divided into 22 districts and as many as 5 in the city center do not have a gym. It is also a good choice to consider those districts in which gyms are in small quantities and they are among the most visited places. Since a small number of gyms are visited by a large number of clients interested in visiting gyms, the opening of a new modern gym will definitely attract clients.



Conclusion

- In this study, we modeled the distribution of gyms in Prague using data from geographic locations and the Foresquare API. Using clustering and the K-Means method, we were able to find the best location for the new gym. This can help a budding gym owner choose where to start their business. According to research, the best solution would be areas that do not have a gym.
- This work is for informational purposes only, as we do not have most of the data. For example, visit and prices at the gym. The study also used a small radius of 500 meters. This study is more suitable for the central area of Prague.

- Thanks for watching!