

Restaurants in Warsaw

Capstone Project

Introduction

- The objective of this problem is to analyse restaurant locations in districts of Warsaw (the capital city of Poland).
- This analysis will allow to select the best places to open a new restaurant.
- The main audience of this project are entrepreneurs willing to launch a new restaurant.

Datasets

- List of districts in Warsaw (web scraping from Wikipedia: https://en.wikipedia.org/wiki/Districts_of_Warsaw)
- Geographical coordinates of districts (using geopy package)
- Foursquare data (extracted by Foursquare API using geographical coordinates, data will include information about restaurants in 3 km radius from the center of each district), for each district the number of restaurant of each type will be counted

Methodology

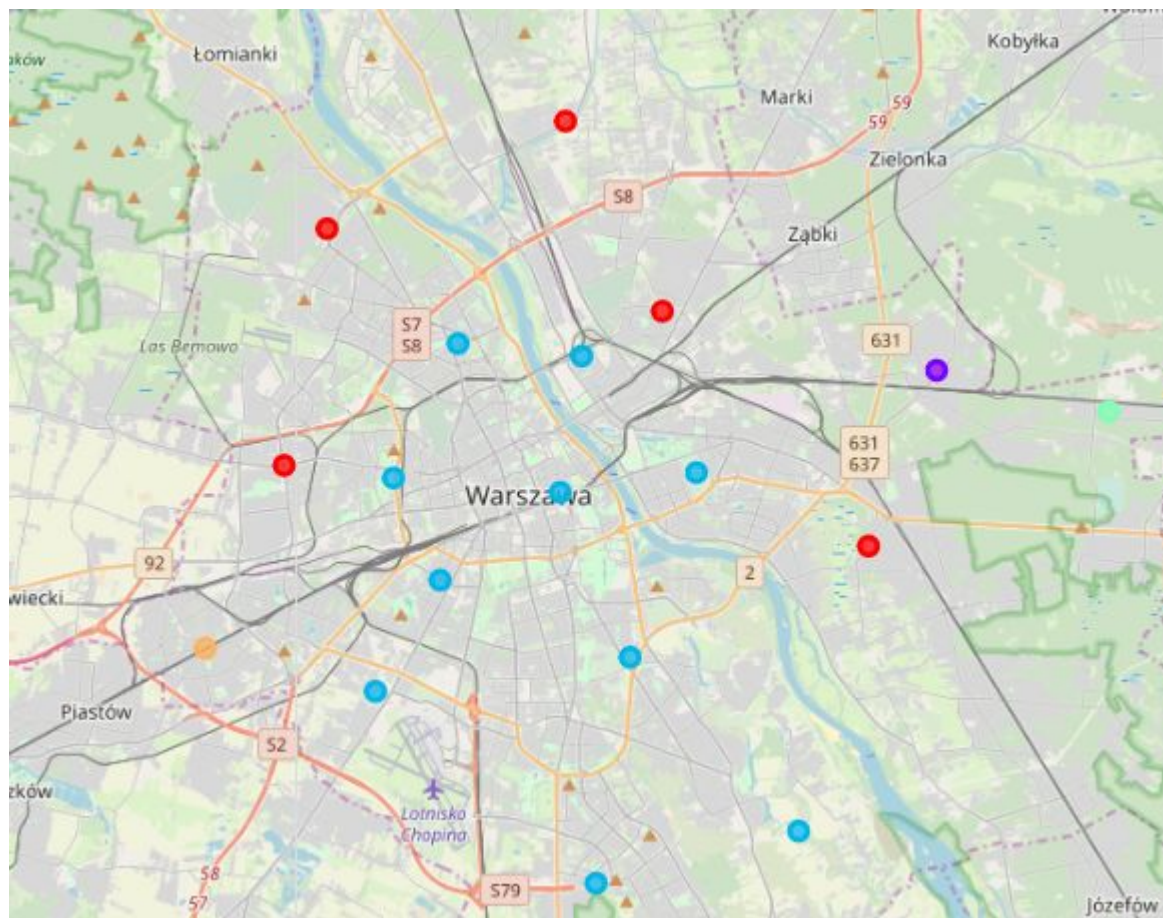
- Wikipedia web scraping
- Adding geographical coordinates
- Getting restaurants data from Foursquare API
- Grouping districts by frequency of each restaurant type
- Clustering
- Clusters visualisation and presentation

Results

5 clusters:

Two biggest clusters:

- Central and peripheral districts.
- 3 clusters consisted of 1 districts that are far from the center of the city



Summary

- Most of the restaurants in Warsaw are concentrated in blue cluster that consists of the most populous districts in the city.
- It is recommended not to open the most popular restaurant category in each districts because there are too saturated. In each district, the best way is to open restaurant from the second half of the most popular venues, because those restaurants are popular in each district but the market for them is not saturated.