

## **IBM Data Science Capstone Project**

**Subject: Best places to open restaurant in Warsaw**

### **Introduction/Business Problem**

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 or restaurants chains. Using insights from this project, stakeholders will be able to make better and more effective business decisions. Foursquare location data will be enhanced with statistical information to give a detailed picture of situation. The project will use analytical approach, especially clustering algorithms will be used.

### **Data**

The following data sets will be used in project:

- List of districts in Warsaw (web scraping from Wikipedia:  
[https://en.wikipedia.org/wiki/Districts\\_of\\_Warsaw](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 1,5 km radius from the center of each district), for each district the number of restaurant of each type will be counted

- Statistical information (such as number of inhabitants in each district, income, unemployment rate and other demographic data to enhance locations analysis, data will be acquired from Polish statistical office:

<https://warszawa.stat.gov.pl/publikacje-i-foldery/inne-opracowania/panorama-dzielnic-warszawy-w-2017-r-,5,19.html>)