

Applied Data Science Capstone Project

“Business integration trips in the Silesian region”.

Data

A description of the data and how it will be used to solve the problem.

Geometries of cities in the Silesian region were extracted from json file taken from the NYU Spatial Data Repository website¹. Available data represent the second-level administrative divisions of Poland and contain all regions (Voivodeship) and their corresponding cities in Poland. As the file format describes geometries as multipolygons, the library Shapely² was used to get the centroid location coordinates for each city. Only data containing geometries of cities belonging to the Silesian region were used to look for venues in those regions and visualise data on maps.

Venues of interest in the Silesian region were found using the Foursquare API³. Multiple searches were performed to look for all venue categories, search for selected venues and explore popular spots. All venues data were analysed and the folium library was used to visualize the cities and clustered venues in the Silesian region, whereas the choropleth map to show quantitative values of venues across geographical areas.

¹ <https://geo.nyu.edu/catalog/stanford-xh662zc5620>

² <https://pypi.org/project/Shapely/>

³ <https://developer.foursquare.com/docs/>