Coursera Capstone Project

IBM Applied Data Science Capstone

Part of IBM Data Science Processional Certificate

Property opportunities in Olsztyn, Poland

photo copyright: https://bit.ly/2Swn1vt

Problem Statement

I would like this project to help me in my personal dilemma. Therefore, I will neither be focusing on a business plan as such, nor will I try to sell this idea to a client (other than myself). Because of that I believe I should start with a short introduction. Hi, I'm Piotr. I live and work in UK, but I was born and raised in Poland. Olsztyn is a town located in North East Poland. Its surrounded by forest, 1000+ lakes (so they say) and It's a place where I was born couple of decades ago. My family is still here, and I do visit them as often as I can.

Some time ago I started to explore the idea of purchasing a house in Olsztyn. The problem is that the town changed so much since I left, that I wouldn't have the slightest idea which areas should I consider. Although my family is still here, they have no interest in property market and I would need to rely on my own research.

This is where this assignment comes into play: I plan to leverage Foursquare location data to help me make aforementioned decision and re-discover the area. There are multiple criteria that can be taken into consideration here: access to medical services, proximity to shopping centers, entertainment venues, roads, schools, lakes, forests, prices of land, building materials etc. In order to simplify this task, I will focus on the selected data available through Foursquare and will list it in the next paragraph. The goal is to find a location that's in a proximity to desired services and can potentially increase in value over time. Naturally, this analysis will be also beneficial for those who don't plan to settle in Olsztyn but instead are looking for investment opportunity.

Data

This paragraph will include a list of necessary data sources and the ways I plan to utilize them in order to solve my problem.

1. List of neighborhoods in Olsztyn; taken from Wikipedia.
2. Coordinates (Latitude and Longitude) of those neighborhoods that will be used to plot a map of the area; using geocoding library
3. List of POI's and venues, particularly related to (sourced using Foursquare API):
    o arts & entertainment
    o college & university
    o medical services

In order to get the list of neighborhoods in the area, I will use BeautifulSoup library. As you may know, it's a Python package that is used for extracting data from HTML documents by scraping them and creating a parse tree. After obtaining that list I will use geocoder library to allocate latitude and longitude coordinates accordingly to aforementioned areas and later visualize a map using folium. The last step would be to utilize Foursquare API and fetch data related to entertainment, schools and medical services (in order to simplify task at hand).