

Coursera's Capstone Project

IBM Applied Data Science Capstone

Opening a Hotel in Krakow, Poland

By: Tomasz Cyparski

April 2020



Introduction

Krakow was the official capital of Poland until the end of the XVI century. It has been one of the leading centers of Polish academic, economic, cultural and artistic life. It's one of Europe's most beautiful cities, its Old Town was declared the first UNESCO World Heritage Site in the world.

More than 14 million tourists have visited Krakow in 2019 according to the Malopolska Tourist Organization. From that over 3 million were foreign visitors mostly from Germany, Great Britain, Italy, France and Spain.

Based on Statista website (www.statista.com) there were approximately 14 five-star hotels in the city in 2018, an increase of 27 percent compared to the previous year.

That determine the demand for the new, high quality hotels in Krakow as the number of foreign visitors increases every year.

Business Problem

The objective of this project is to analyze and select the best location in the city of Krakow, Poland to open a new hotel. By using data science and machine learning techniques including clustering and geolocation data and using foursquare API this project will try to answer the the business question whether opening a new hotel in Krakow is still a profitable investment.

All Krakow neighborhoods will be taken into the consideration (and of course the close proximity to the center of the city would be a priority). Knowing that center part of Krakow (also known as the Old Town) is a very attractive business spot and in general very hard to find a space needed for a hotel.

Target audience

The results of this project may be the most useful for the investors, real estate agencies that are looking to do business in Krakow. With the increasing number of visitors coming to the city especially during the summer, opening a new hotel may be a good investment.

It's also worth mentioning that the number of well-known sport and cultural events that take place in Krakow is constantly growing and thus the need for a quality accommodation for guests and VIPs.

Data

In order to solve the mentioned above problem the following data will be collected:

- List of all Krakow neighborhoods
- Latitude and longitude coordinates of those neighborhoods and the city itself. They will be used to plot the map with the existing venues
- Hotels data in order to determine their locations and proximity from the city center and from each other

Data sources and methods used to extract them

The list of Krakow neighborhoods and their corresponding geographical coordinates will be extracted from Polish version of Wikipedia (https://pl.wikipedia.org/wiki/Podzia%C5%82_administracyjny_Krakowa). Web scraping technique will be used to extract the data from the website and import to Python. Similarly, all 18 Krakow neighborhoods geographic locations will be imported and processed to allow for the map creation.

Next, using Foursquare API the list of the Krakow venues will be imported and sorted and in the end those representing hotels will be used to create

clusters and based on the result the final conclusion will be made if the opening of a new hotel (and if yes, where) in Krakow is still profitable.