**Capstone Project – Restaurants in Kraków**

**Introduction**

Kraków is the second biggest city of Poland and the most visited one. Over 700.000 people live in Kraków and more than 14 million tourists visit Kraków every year. It is one of the most beautiful cities in whole Europe and it is an old city, almost the whole city survived the second world war and the historic buildings are in really good condition. The city of Kraków is divided into 18 administrative district, each with a degree of autonomy within the municipal government.

A close up of a map

Description automatically generated

The oldest neighborhoods of Kraków were incorporated into the city before the late 18th century. They include the Old Town, the Wawel, which is the site of the Royal Castle and the Cathedral, Kazimierz, the Jewish neighborhood and the ancient town of Kleparz, all of them are part of the district I.

As you can see, Kraków is not a huge city, but receive many tourists every year. So, if you want to open a restaurant in Kraków, where would be the best location to do it? When we think of it by the investor, we expect from them to prefer the district with the higher number of tourists and city residents, so the restaurant will have higher amount of customers, but the investors also prefer the districts where there is a lower real estate cost and the type of business is less intense.

Although the center of Kraków have many restaurants and the real estate cost is more expensive, it is always full of both tourists and city residents, so it is a perfect place for a restaurant. Therefore, the aim of the project is to find a good location to open a restaurant in the center of Kraków based on the location of the rest of the restaurants. We can create a map to know where the restaurants are located in this district and choose a good location for the new restaurant. Let’s alto make sure that the **audience** is explicitly defined to be the **local restaurant** **entrepreneurs** in Kraków, and they should care about this problem because the location of the new restaurant has a significant impact on the expected returns.

**Data**

The data used to solve this problem is geolocation data collected from Foursquare, which has a radius of 2000 and a limit of 100, so we select 100 restaurants from the center of Kraków. Adequate explanation and discussion of the data is the following. We receive the data as json with all of the restaurants, and many information, as you can see in the following picture.

A screenshot of text

Description automatically generated

We transform the data into a single dataframe, containing the name of the restaurant, the category of their food, the address, the postal code and the location defined by lat, which stands for latitude, and lng, which stands for longitude. Although the decision is based on the location, so we only need the lat and lng categories, any restaurant which does not include all of the categories, will be deleted.

A screenshot of a cell phone

Description automatically generated

Finally, we have a dataframe with 63 rows which contains all of the categories, and we can start our analysis.