# 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.

Map of districts of the City of Kraków
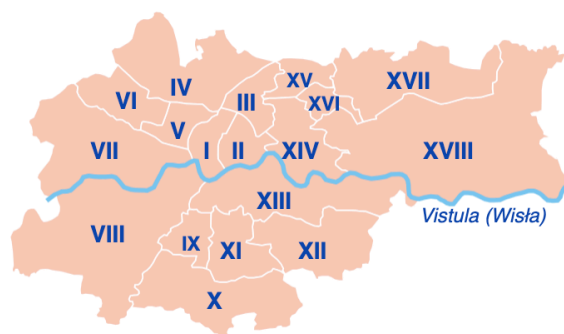
Interactive map. For more information, click on district number.

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 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.

```
Out[8]: {'reasons': {'count': 0,
          'items': [{'summary': 'This spot is popular',
            'type': 'general',
            'reasonName': 'globalInteractionReason'}]},
        'venue': {'id': '597cc55b6dcf041c42573b04',
         'name': 'Pizzatopia',
         'location': {'address': 'Szewska 22',
          'lat': 50.062943238240265,
          'lng': 19.933833687142826,
          'labeledLatLngs': [{'label': 'display',
            'lat': 50.062943238240265,
            'lng': 19.933833687142826}],
          'postalCode': '31-009',
          'cc': 'PL',
          'city': 'Kraków',
          'state': 'Lesser Poland',
          'country': 'Polska',
          'formattedAddress': ['Szewska 22', '31-009 Kraków', 'Polska']},
         'categories': [{'id': '4bf58dd8d48988d1ca941735',
           'name': 'Pizza Place',
           'pluralName': 'Pizza Places',
           'shortName': 'Pizza',
           'icon': {'prefix': 'https://ss3.4sqi.net/img/categories_v2/food/pizza_',
            'suffix': '.png'},
           'primary': True}],
         'photos': {'count': 0, 'groups': []}},
        'referralId': 'e-3-597cc55b6dcf041c42573b04-0'}
```

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.

	uid	name	food	address	postalcode	lat	lng
0	597cc55b6dcf041c42573b04	Pizzatopia	Pizza	Szewska 22	31-009	50.062943	19.933834
1	53e678c5498e35f10c8bcbfa	Bianca Pasta & Wine	Italian	Plac Mariacki 2	31-042	50.061899	19.939741
2	56377311498ea3a0befc336c	Bonerowska Gourmet Steak & Fish	Steakhouse	Rynek Główny 42	31-013	50.062246	19.937987
3	4e6f50cb483bd41a3db4dfa4	MoaBurger	Burgers	Mikołajska 3	31-027	50.061306	19.941310
4	4b813ff5f964a520259c30e3	Café Camelot	Café	Świętego Tomasza 17	31-022	50.063098	19.939010
5	4facdb7fe5e83b277dae66de	Charlotte - Chleb i Wino	Café	Plac Szczepański 2	31-011	50.063455	19.935069
6	4c00dc18ceabb7138aef8752	Bunkier Café	Café	Plac Szczepański 3a	31-011	50.063577	19.934080
7	4be1a1a885670f474eae2bd7	Morskie Oko	Polish	Plac Szczepański 8	31-011	50.063814	19.935762
8	592005654940bc3f72327973	La Petite France	French	Szpitalna 20	31-024	50.062825	19.941311
9	53a1380f498e903f74523f65	Massolit Bakery & Cafe	Bakery	Smoleńsk 17	31-108	50.058975	19.930351

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