

Capstone Project - The Battle of the Neighborhoods

Vegetarian Restaurant in Krakow

1. Introduction

Recent results show that over 3 million Poles no longer eat meat and 8.4% of Polish adults had followed a vegetarian diet or had eaten only fish but no other meat for a period of time [1]. The same report indicates that even people who eat meat want to make an effort to limit their consumption.

Warsaw, capital of the country was also ranked the 6th most vegan friendly city on the planet [2]. Thinking of this, our stakeholder wants to expand this trend to Krakow, and open a vegetarian restaurant in the city, which is one of the largest cities in Poland.

The advantages of Krakow include the fact that the city attracts tourists with its beauty, unique historic complex and cultural events.

When choosing a place to open a new venue, we should take into considerations a few key aspects, like demographics of people interested in this type of restaurant, their income, where they usually live across the city and also real state prices for our business. We also need to know the density of other restaurants and other restaurants similar to ours.

Considering that the analysis in the scope of our initial research relies on free available information and constraints imposed by the Foursquare API, used to track the nearby venues for the district in the city, we will limit our analysis to a few key points:

- Population in each district
- Number of restaurants per district and
- Number of vegetarian restaurants already in the area.

2. Data

For this preliminary analysis we will consider the list of 18 districts in the Krakow area and their coordinates, available in [3]; the population and population density in each district [4]; the number of restaurants within a certain radius from each district and the number of

vegetarian restaurants in each district. This will be limited to the free information from Foursquare API.

In the following sections we will use the Foursquare API to search the venues in each district and to sort the total number of restaurants and total number of vegetarian or vegan restaurants in each district. Using this in comparison to the population in each district, we apply an unsupervised learning technic (KMeans) to cluster the districts with regards to our feature selection in order to help us to assess the best neighborhood for the new restaurant.

2.1. Initial set of data

Krakow has 4 distinct areas with a total of 18 districts. Our first step is to get the population, area (in km^2) and density of population (number of people per km^2). This data was available in [4].

We also have available in Wikipedia [3] the latitude and longitude for each district, which we can use to produce Table 1.

We show in Figure 1 the location of the districts with the radius of each circle being proportional to the population density in each of them.

Figure 1: Map of Krakow with population density in each district highlighted.

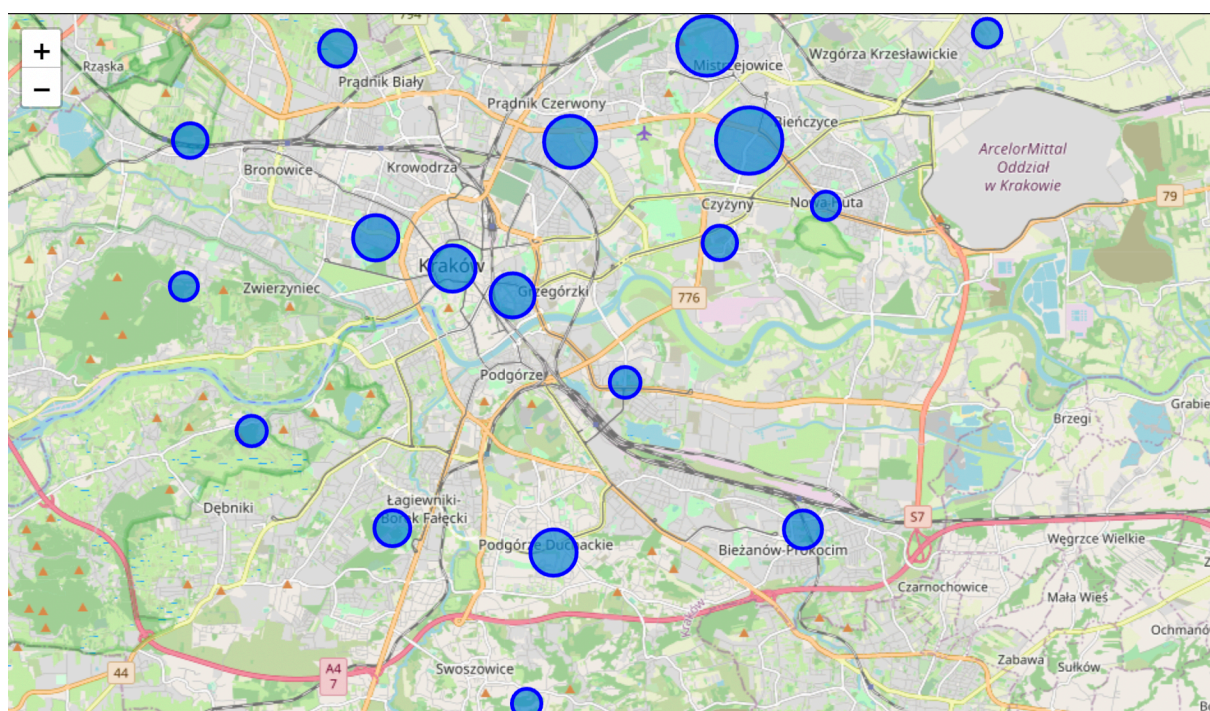


Table 1: Krakow districts, population, area (km^2), density of population (people per km^2) and coordinates.

Region	District	Population	Area	Density of population	Latitude	Longitude
Śródmieście	Stare Miasto	31 359	5 568	5 632	50.061389	19.937222
Śródmieście	Grzegórzki	29 474	5 845	5 042	50.056611	19.953278
Śródmieście	Prądnik Czerwony	46 627	6 438	7 243	50.083231	19.968844
Krowodrza	Prądnik Biały	70 647	23 419	3 017	50.099239	19.906303
Krowodrza	Krowodrza	30 223	5 619	5 379	50.066667	19.916667
Krowodrza	Bronowice	23 678	9 560	2 477	50.083333	19.866667
Krowodrza	Zwierzyniec	20 392	28 731	710	50.05825	19.864922
Podgórze	Dębnik	61 637	46 189	1 334	50.033333	19.883333
Podgórze	Łagiewniki-Borek Fałęcki	15 259	5 415	2 818	50.016472	19.921083
Podgórze	Swoszowice	27 493	25 604	1 074	49.986378	19.956869
Podgórze	Podgórze Duchackie	53 747	9 540	5 634	50.012278	19.964069
Podgórze	Bieżanów-Prokocim	62 830	18 474	3 401	50.016328	20.031161
Podgórze	Podgórze	36 885	25 667	1 437	50.041667	19.983333
Nowa Huta	Czyżyny	29 635	12 257	2 418	50.065723	20.00879
Nowa Huta	Mistrzejowice	52 011	5 590	9 304	50.099575	20.005478
Nowa Huta	Bieńczyce	41 112	3 699	11 114	50.083333	20.016667
Nowa Huta	Wzgórze Krzesławickie	20 205	23 815	848	50.101944	20.080497
Nowa Huta	Nowa Huta	51 234	65 410	783	50.072222	20.0375

2.2. Using Foursquare API

Having the latitude and longitude of each district we can use the Foursquare API (<https://developer.foursquare.com/>) to find the nearby venues in each region.

Doing some preliminary exploratory analysis, we can see the main 3 venues in each district and also notice that we have restaurants among the top 3 in many districts, including in the more densely populated ones. We present this result in Table 2.

In the next section we will merge the results from Table 1 with all the venues listed in the Foursquare API and select the features we want to use in our analysis. One important note here is that we have a limitation in the number of venues we can get for free with Foursquare, but for the current scope of our assessment (limit = 500) this is enough.

Table 2: Districts and top 3 venues based on Foursquare API data.

District	Stare Miasto	Grzegórzki	Prądnik Czerwony
1	Café	Bar	Food & Drink Shop
2	Hotel	Hotel	Bus Station
3	Italian Restaurant	Food Truck	Pizza Place
District	Prądnik Biały	Krowodrza	Bronowice
1	Coffe Shop	Café	Flower Shop
2	Grocery Store	Park	Train Station
3	Italian Restaurant	Bistro	Flea Market
District	Zwierzyniec	Dębnik	Łagiewniki-Borek Fałęcki
1	Zoo	Scenic Lookout	Shopping Mall
2	Grocery Store	Restaurant	Bus Station
3	Restaurant	Pizza Place	Gym
District	Swoszowice	Podgórze Duchackie	Bieżanów-Prokocim
1	Shoe Store	Supermarket	Convenience Store
2	Memorial Site	Convenience Store	Gym/Fitness Center
3	Liquor Store	Gas Station	Platform
District	Podgórze	Czyżyny	Mistrzejowice
1	Tram Station	Bus Stop	Tram Station
2	Bus Station	Supermarket	Park
3	Supermarket	Department Store	Sushi Restaurant
District	Bieńczyce	Wzgórza Krzesławickie	Nowa Huta
1	Supermarket	Construction & Landscaping	Park
2	Fast Food Restaurant	Diner	Bus Station
3	Tram Station	Pet Store	Café

2.3. Data cleaning and feature selection

Now we can select the venues that are interesting to our analysis. We start by sorting the number of places tagged with “Vegetarian/Vegan”, which will give us the number of vegetarian places in each district. We also select the venues with tags “Restaurant, Diner, Pizza Place” to account for all restaurants in each neighbourhood.

After some data cleaning we have the final table we will use. Values are presented in Table 3.

Table 3. Number of restaurants and of vegetarian restaurants in each district.

District	Latitude	Longitude	Area	Density of population	Total Restaurants	Total Vegetarian
Stare Miasto	50.061389	19.937222	5 568	5 632	28.0	3.0
Grzegórzki	50.056611	19.953278	5 845	5 042	28.0	2.0
Prądnik Czerwony	50.083231	19.968844	6 438	7 243	5.0	0.0
Prądnik Biały	50.099239	19.906303	23 419	3 017	5.0	0.0
Krowodrza	50.066667	19.916667	5 619	5 379	22.0	1.0
Bronowice	50.083333	19.866667	9 560	2 477	0.0	0.0
Zwierzyniec	50.05825	19.864922	28 731	710	2.0	0.0
Dębnik	50.033333	19.883333	46 189	1 334	3.0	0.0
Łagiewniki-Borek Fałęcki	50.016472	19.921083	5 415	2 818	3.0	0.0
Swoszowice	49.986378	19.956869	25 604	1 074	1.0	0.0
Podgórze Duchackie	50.012278	19.964069	9 540	5 634	1.0	0.0
Bieżanów-Prokocim	50.016328	20.031161	18 474	3 401	1.0	0.0
Podgórze	50.041667	19.983333	25 667	1 437	3.0	0.0
Czyżyny	50.065723	20.00879	12 257	2 418	5.0	0.0
Mistrzejowice	50.099575	20.005478	5 590	9 304	2.0	0.0
Bieńczyce	50.083333	20.016667	3 699	11 114	6.0	0.0
Wzgórze Krzesławickie	50.101944	20.080497	23 815	848	1.0	0.0
Nowa Huta	50.072222	20.0375	65 410	783	4.0	0.0

We will use the density of population, density of restaurants (Total Restaurants/Area) and density of vegetarian restaurants (Total Vegetarian/Area) as our features to use the clustering method KMeans in order to define a set number of clusters that will help us to assess the best place to open the new restaurant.

References:

- [1] <https://notesfrompoland.com/2020/01/14/almost-40-of-poles-claim-to-be-cutting-back-on-meat-and-8-4-have-quit-completely/>
- [2] <https://www.happycow.net/vegtopics/travel/top-vegan-friendly-cities>
- [3] https://en.wikipedia.org/wiki/Districts_of_Krak%C3%B3w
- [4] <https://www.bip.krakow.pl/zalaczniki/dokumenty/n/253681/karta>
- [5] <https://medium.com/analytics-vidhya/how-to-determine-the-optimal-k-for-k-means-708505d204eb>