

Problem Background

Warsaw is the capital of Poland, fast growing business centre in Central Europe. Disposable income in the City is much higher than in other Polish regions and now it is even above European Union average. It attracts a lot of investors, new business district is emerging and more and more houses are being built in residential areas.

This generates a lot of opportunities for new businesses as long as many challenges from the city authorities to keep the capital growth sustainable.

Problem Description

Poles spend relatively low on away from home consumption. There is a lower number of restaurants or bars than in neighbouring countries, and that generates an opportunity for businesses. An investor wants to invest in new venue in Warsaw. He needs to decide that type of outlet he will open and in which neighbourhood.

He asked me to advise them about the above business questions:

- How many outlets they should open
- Where they should locate them

Data collection

I collected data Warsaw districts - their characteristics and business potential. Also will explore Foursquare API data in order to understand the competitive environment.

Crucial will be selecting decision criteria and finally decide on what should be the main factor of choosing location and number of outlets.

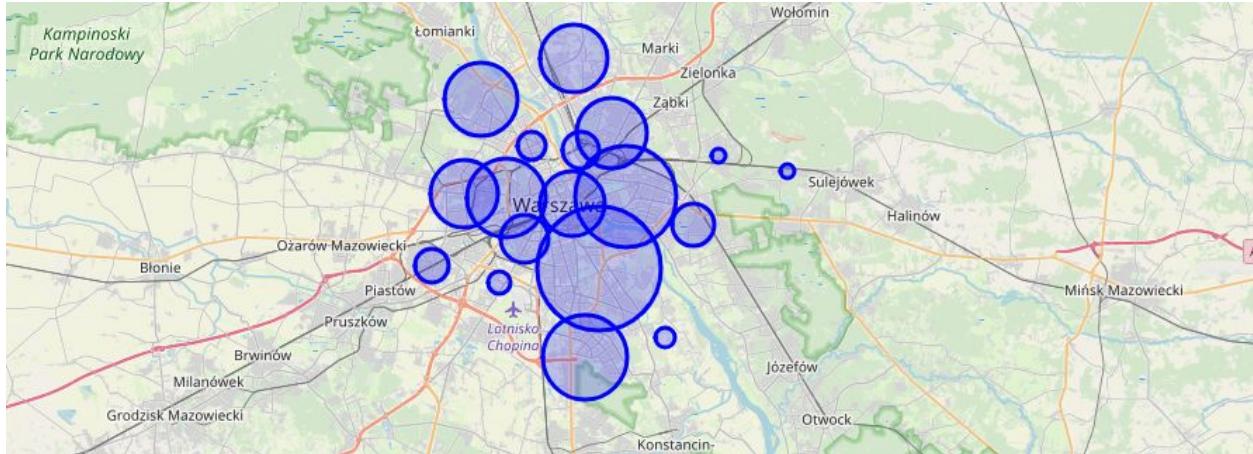
Success Criteria

1. Choosing any potential locations based on specified criteria. The criteria are:
 - a. Small / large number of competitors
 - b. And all other data that will be available after exploration phase
- 2.

The data

The data used come from:

1. Wikipedia definition of 18 neighbourhoods
2. Foursquare API data - this dataset helped me understand the competitive environment and find the best location for venues. It will also tell me what type of outlets are in each of the neighbourhoods. The question then is if we want to use the potential of the district or bring something new to competitive environment
3. Neighbourhoods with their potential (total number of inhabitants) are shown on Map



And below you can find statisticsa about neighbourhoods:

Borough	Inhabitants	Density	Area	
0	Mokotów	217577	6143	35.42
1	Praga-Południe	178726	7986	22.38
2	Ursynów	150273	3432	43.79
3	Wola	140111	7275	19.26
4	Bielany	132026	4082	32.34
5	Targówek	123941	5117	24.22
6	Bemowo	122210	4898	24.95
7	Śródmieście	117005	7515	15.57
8	Białołęka	119374	1634	73.04
9	Ochota	83081	8547	9.72
10	Wawer	75991	991	79.70

11	Praga-Północ	64904	5683	11.42
12	Ursus	59261	6331	9.36
13	Żoliborz	51441	6073	8.47
14	Włochy	41929	1465	28.63
15	Wilanów	35511	1021	36.73
16	Wesoła	25106	1094	22.94
17	Rembertów	24148	1251	19.30

Data Analysis

When analyzing data I came to 2 potential solutions:

- Have the best outlet type in every district
- Have only one type and choose which districts to enter

1st approach was solved by selecting the most popular (frequent) venue type in Warsaw and comparing it with frequency in each of the districts. The ratio shown the potential of investment:

----Bemowo----

```
venue freq freqtot ratio
0 Park 0.02 0.04 1.97
```

----Białołęka----

```
venue freq freqtot ratio
0 Supermarket 0.08 0.04 0.44
```

----Bielany----

```
venue freq freqtot ratio
0 Park 0.02 0.04 1.76
```

----Mokotów----

```
venue freq freqtot ratio
0 Supermarket 0.01 0.04 3.56
```

----Ochota----

	venue	freq	freqtot	ratio
0	Restaurant	0.01	0.02	2.24

----Praga-Południe----

	venue	freq	freqtot	ratio
0	Hotel	0.01	0.02	2.33

----Praga-Północ----

	venue	freq	freqtot	ratio
0	Supermarket	0.01	0.04	3.63

----Rembertów----

	venue	freq	freqtot	ratio
0	Café	0.12	0.07	0.54

----Targówek----

	venue	freq	freqtot	ratio
0	Gym / Fitness Center	0.03	0.03	0.95

----Ursus----

	venue	freq	freqtot	ratio
0	Coffee Shop	0.02	0.03	1.3

----Ursynów----

	venue	freq	freqtot	ratio
0	Café	0.02	0.07	3.1

----Wawer----

	venue	freq	freqtot	ratio
0	Café	0.06	0.07	1.09

----Wesoła----

	venue	freq	freqtot	ratio
0	Pizza Place	0.17	0.03	0.16

----Wilanów----

	venue	freq	freqtot	ratio
0	Supermarket	0.05	0.04	0.69

----Wola----

	venue	freq	freqtot	ratio
0	Burger Joint	0.01	0.01	1.4

----Włochy----

	venue	freq	freqtot	ratio
0	Café	0.04	0.07	1.84

----Śródmieście----

	venue	freq	freqtot	ratio
0	Gym / Fitness Center	0.01	0.03	2.89

----Żoliborz----

	venue	freq	freqtot	ratio
0	Hotel	0.01	0.02	2.33

2nd approach helped us to understand in which type of venue we need to invest in general and which neighbourhoods are missing it.

The most popular venue in Warsaw is: Café

And the list of neighbourhoods that do not have Cafe in the list of top venues is below:

	Borough	Inhabitants	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
2	Ursynów	150273	Convenience Store	Chinese Restaurant	Supermarket
5	Targówek	123941	Supermarket	Plaza	Market

8	Białołęka	119374	Flower Shop	Asian Restaurant	Outdoors & Recreation
10	Wawer	75991	Hotel	Bus Station	Bakery
12	Ursus	59261	Supermarket	Italian Restaurant	Hotel
14	Włochy	41929	Supermarket	Hotel	Turkish Restaurant
15	Wilanów	35511	Restaurant	Golf Course	Gym / Fitness Center
16	Wesoła	25106	Train Station	Forest	Neighborhood

Recommendations:

Let's assume we want to start with 3 venues, there are 2 potential options for investment:

1. **Investing in Supermarket in Mokotów and Praga Północ with additional investment in Cafe in Ursynów**
or
2. **Investing in Cafe in Ursynów, Targówek and Białołęka**