

Big Corporation Offshore offices placement

Rodrigo Cofré

June 06, 2020

1.Introduction

Since the beginning of March as the COVID-19 spreads, restrictions rise and the world went from commuting every day to work to a forced home office work style. The economy of several countries suddenly stopped, and this hurt small and big companies. However, one advantage big companies have (multinationals) is that they can quickly move their work force to countries where is cheaper for them to hire workers.

As an effort to make the off shores offers attractive for the locals, big corporations are trying to locate their new offices in strategic parts of town that will offer their workers proximity to venues that they conveniently could use in their way to or from the office. One popular destination for big corporation offshore activities in Europe is Poland.

This analysis will be focus in the region of Lesser Poland (Malopolska), where there are many offshore offices from big companies like HBSC or UBS. The results of this analysis could then be applied to other cities to search for possible new location for offshore activities.

2.Data acquisition and preprocessing

The data was obtained from the website **data.opendatasoft.com** making use of its API. From this we obtained information from the region of Lesser Poland about its neighborhoods, postal codes for each of them and their geographical coordinates. Using the geographical coordinates of each neighborhood and using the Foursquare API we also obtained information about nearby venues for each neighborhood at the date of 05/05/2020.

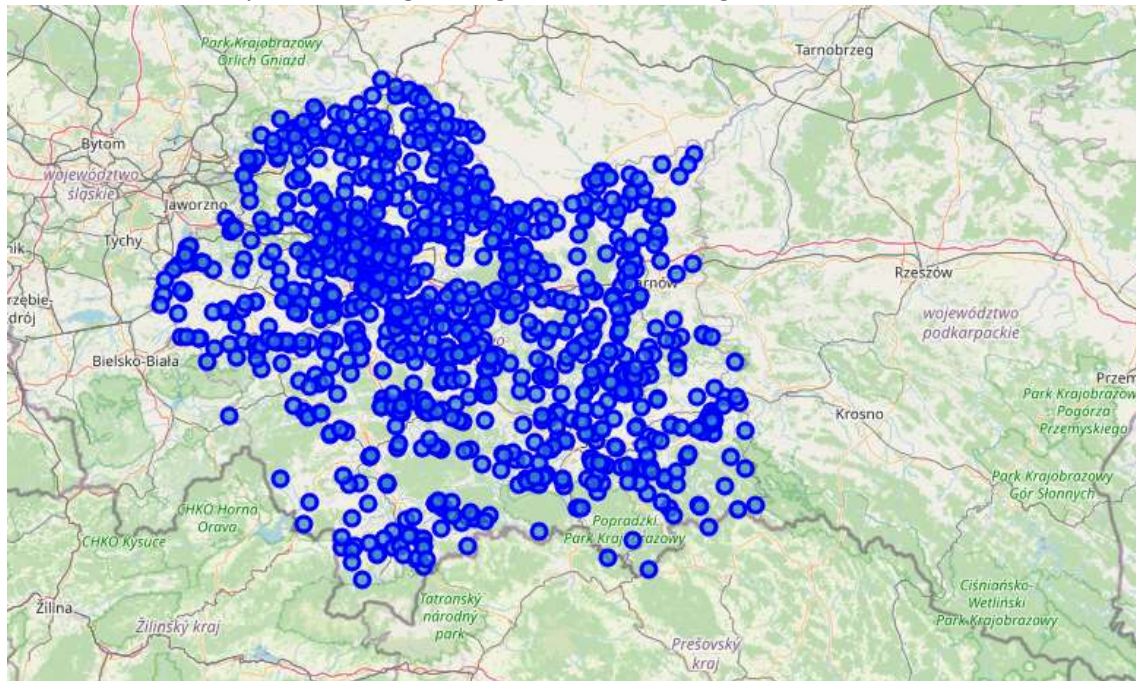
We transformed and crossed the information between these two data sets to obtain one table with the top 5 venues in each neighborhood, as is shown next.

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Andrychów,34-120	Restaurant	Plaza	Falafel Restaurant	Food	Flower Shop
1	Białka Tatrzańska,34-405	River	Falafel Restaurant	Food	Flower Shop	Flea Market
2	Białka Tatrzańska,34-530	Polish Restaurant	Restaurant	Performing Arts Venue	Diner	Eastern European Restaurant
3	Bielanka,34-723	Restaurant	Food Truck	Food & Drink Shop	Food	Flower Shop
4	Bochnia,32-700	Diner	Café	Plaza	Pizza Place	Farm

3. Methodology

3.1 Analysis

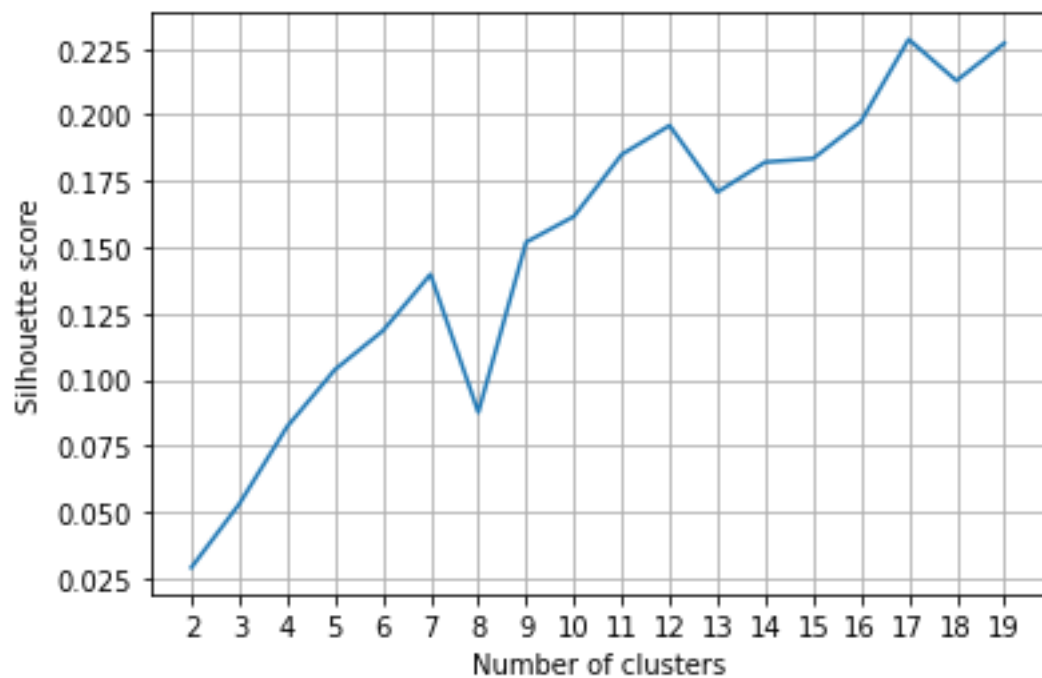
We started the analysis visualizing the region where the neighborhoods are located.



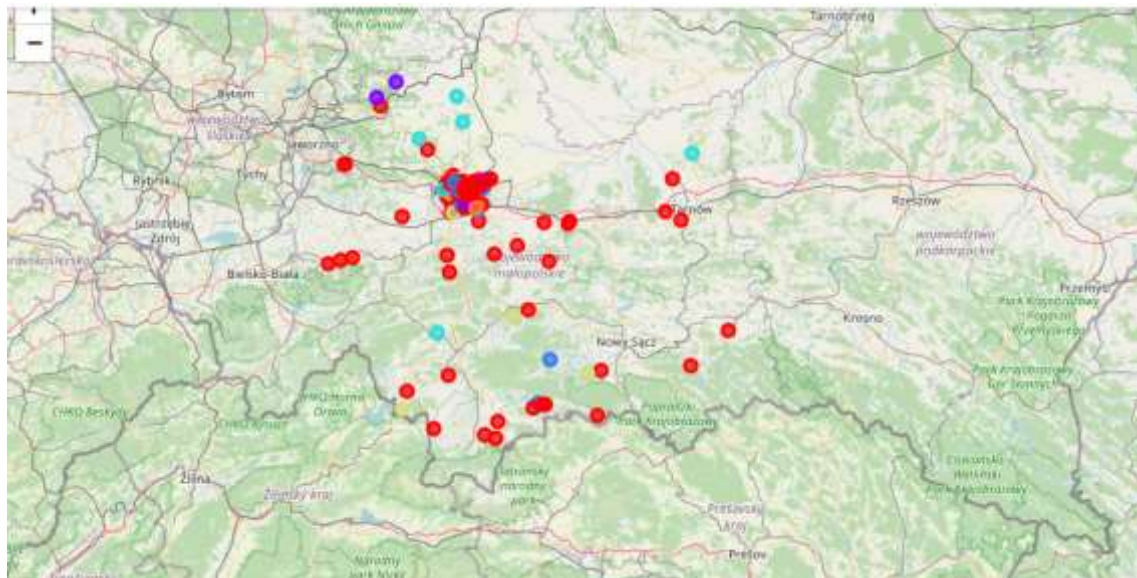
Here we can see a great number of neighborhoods, where the main concentration of them occur in densely populated areas like big cities. Later, along with the venue information from the Foursquare API for each one of these neighborhoods it was possible to obtain their top 5 venue categories as well. This allowed us narrow down the number of possible locations for a potential office, in great part since most of the neighborhoods listed do not possess any venues nearby (in 200 meters radius).

Once we had all the neighborhoods with their top 5 venues categories, we proceeded to cluster them using the frequencies of appearance of each venue in each neighborhood. First we searched for an appropriated number of cluster, for this we calculated the

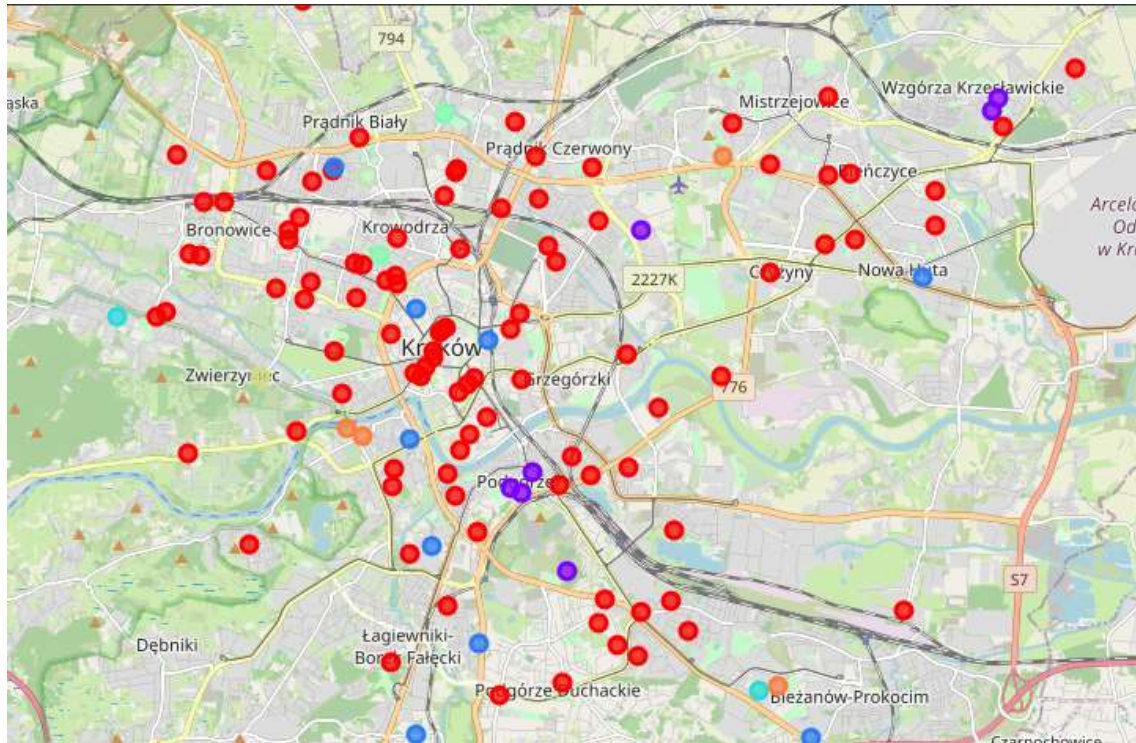
Silhouette score for a different range of number of clusters, as presented in the next plot.



The Silhouette score ranges between -1 and 1, where 1 is the best value for a given cluster. It is natural that as the number of clusters increases the score is higher. Ultimately, the clustering that characterizes the data more accurately is the clustering where each point is its own cluster. But we seek a number of clusters that allows us to characterize the points with the least number of clusters. We then choose a number of clusters equal to 7, just before the break that we can see in the plot. In the next map we see a visualization of the clusters.



Here we see how the number of points decreased drastically from the previous map shown.



Here we see a further zoom into the city of Cracow that shows the distribution of the different clusters in the city.

4. Results

Once we calculated the clusters and labeled each neighborhood. The analysis of each cluster made us to characterize them in the following for:

Cluster 1: Neighborhoods with a great variety of venues, being most of them leisure or food related.

Cluster 2: Neighborhoods close to historical places, plazas and restaurants.

Cluster 3: Neighborhoods most appealing for tourists. Near to hotels, tram and bus stations as to places to eat.

Cluster 4: Neighborhoods of mostly residential areas. They stand out in grocery stores and leisure venues.

Cluster 5: Neighborhoods without a great variety of venues.

Cluster 6: Neighborhoods without a great variety of venues.

Cluster 7: Neighborhoods closer to nature and leisure venues.

5. Discussion

From this characterization we obtained the following insights. Clusters 5,6 and 7 do not offer a great variety of venues, centering in types of venues as restaurants, parks and leisure activities. Cluster 4 represent more residential areas and we could rule out these neighborhoods as potential office locations. City regulations may be the reason behind

these areas being mostly residential. We noted that clusters 2 and 3 represent attractive neighborhoods for tourists, with mostly historical places, hotels, restaurants and means of transport nearby. Finally, cluster 1 offers a great variety of venues, being diverse between each of the neighborhoods. But most importantly, its neighborhoods are far from similar to the ones from other clusters. This last fact leads us to think that a neighborhood in this cluster may be a good candidate for an office location. With this information, a corporation wanting to locate an office in the area just needs to align what they believe will be attractive to their employees and search for a neighborhood within the cluster 1 that fits these characteristics.

6. Conclusions

The main objective of this analysis was to find possible location for an offshore office for a big corporation given information about the neighborhoods in the region and venues nearby them. Thanks to the exploratory analysis and posterior clustering was possible to characterize the neighborhoods in the region and rule out which ones were not a good fit for any kind of office.

Finally, we were left with several possible candidates, all belonging to one of the characterized clusters. Taking these possible candidates and aligning them with the corporations' priorities for their workers should narrow down the candidates to the potential office location for each corporation. Ensuring the availability of certain venues to their employees.