The best Oslo Borough in times of COVID-19

María Ángeles Jiménez Sigstad

December 2, 2020

1. Introduction

In Oslo, there are many restaurants currently dealing with a low demand due to COVID-19. A new restaurant owner wants to buy a new location to place nearby the main epicentre to increase their brand. Since bars and restaurants are not able to sell alcohol under the current restrictions, it is necessary to go back to past results known about which venues people in Oslo prefer and which Boroughs are more populated. The latter would aid into analysing which would be the best suitable place to open a new restaurant. So, the main objective is finding the best Oslo Borough to place a new restaurant. For this, we will use freely available data about the population in Oslo per Borough and venue information in the surrounding areas acquired from the Foursquare API. Then, we will find where if we can find other Borough with the same characteristics using a Clustering algorithm as a mean to give other options for the buyer in case that it finds it less attractive than other Boroughs or to provide a bigger range of solutions to the problem.

2. Data collection

All data was collected using Python (version 3.7, 2018, Python Software Foundation). Information about Boroughs of Oslo was acquired from (Wikipedia, 2020). Then, to find the latitude and longitude of the Boroughs the geocoder library (Carriere, 2018). Nearby venues around the Boroughs were found using the Foursquare API (Foursquare Developers, 2020).

3. Methodology

Everything that is described here was done using Python (version 3.7, 2018, Python Software Foundation) and a Jupyter Notebook within Watson Studio in the IBM Cloud.

- Data cleaning

The number of residents were normalised to a 0-1 scale for better visualization. Also, venue categories obtained from the Foursquare API for each Borough were used to create a binary feature depending on their presence of absence per Borough.

	Borough	Residents
0	Grünerløkka	1.000000
1	Frogner	0.909149
2	GamleOslo	0.891923
3	Nordstrand	0.712985
4	NordreAker	0.709183
5	Østensjø	0.665370
6	VestreAker	0.646676
7	Alna	0.636421
8	Sagene	0.500691
9	SøndreNordstrand	0.327198
10	St.Hanshaugen	0.323712
11	Ullern	0.198439
12	Bjerke	0.164622
13	Stovner	0.161568
14	Grorud	0.000000

Fig.1. Normalized (0-1) number of residents per Borough where the highest populated is 1 and the lowest is 0.

- Borough information

A table (Wikipedia, 2020) containing the number of residents per Borough was acquired using the BeautifulSoup library (Richardson, 2020). The normalized (0-1 range) number of residents was used as a measure of population per Borough.

- <u>Nearby venues per Borough</u>

To get information about the venues, we got the nearby venues per Borough using the Foursquare API (Foursquare Developers, 2020). Then, we got the frequency of specific types of venues per Borough to see which were the ten most important venues for every Borough.

	Borough	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Alna	Electronics Store	Furniture / Home Store	Grocery Store	Supermarket	Department Store	Bus Station	Bookstore	Kids Store	Flower Shop	Sporting Goods Shop
1	Bjerke	Supermarket	Pizza Place	Soccer Field	Grocery Store	Asian Restaurant	Racetrack	Gym / Fitness Center	Chinese Restaurant	Bus Station	Shopping Mall
2	Frogner	Gourmet Shop	Italian Restaurant	Scandinavian Restaurant	Furniture / Home Store	Wine Shop	French Restaurant	Indian Restaurant	Movie Theater	Park	Cosmetics Shop
3	GamleOslo	Coffee Shop	Record Shop	Park	Café	Plaza	Cocktail Bar	Italian Restaurant	Pizza Place	Restaurant	Moving Target
4	Grorud	Convenience Store	Grocery Store	Wine Shop	Pizza Place	Bus Station	Supermarket	Metro Station	Soccer Field	Eastern European Restaurant	Cosmetics Shop
5	Grünerløkka	Coffee Shop	Bar	Park	Italian Restaurant	Burger Joint	Café	Cocktail Bar	Brewery	Pub	Asian Restaurant

Fig. 2. The ten most important venues per Borough.

- <u>Clustering</u>

To further assess which were the most important Boroughs for setting up a new restaurant, we trained an unsupervised Clustering Machine Learning tool, called KMeans, using the scikit-learn library (Pedregosa *et al.*, 2011; Buitinck *et al.*, 2013). The KMeans algorithm divided our data into five different clusters. Afterwards, the resulting cluster labels (0-4) were added to the ten most important venues per Borough to check if the Boroughs had similar features and if the clusters solved our question.

	Borough	Residents	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue
0	Grünerløkka	1.000000	59.92433	10.76119	1	Coffee Shop	Bar	Park	Italian Restaurant
1	Frogner	0.909149	59.91886	10.70239	1	Gourmet Shop	Italian Restaurant	Scandinavian Restaurant	Furniture / Home Store
2	GamleOslo	0.891923	59.90624	10.77147	1	Coffee Shop	Record Shop	Park	Café

Fig. 3. The top three most populated Boroughs were found in Cluster 1.

- <u>Data visualization</u>

Boroughs and clusters were visualized into the map of Oslo using the Folium library and tables here are outputs obtained from the script created by Jiménez Sigstad (2020).

4. <u>Results and Discussion</u>

Using the first table explained above, we could assess that Grünerløkka, Frogner and Gamle Oslo are the top three epicentres (Fig.1). Interestingly, the three of them were found in the same cluster (Fig.3, Fig.4: Cluster 1-purple) and their most common venues (Fig.2) were gourmet-related (Coffee or gourmet shot) ones. So, we could infer that the three might be suitable places for a new restaurant. But, Grünerløkka had the highest number of residents and the second most common venues was a bar (Fig.1B). So, a safe option might be Grünerløkka. It is important to mention that Frogner's second most common venue is an Italian restaurant. Hence, the selection of the location of the new restaurant depends on the need of the new restaurant owner.

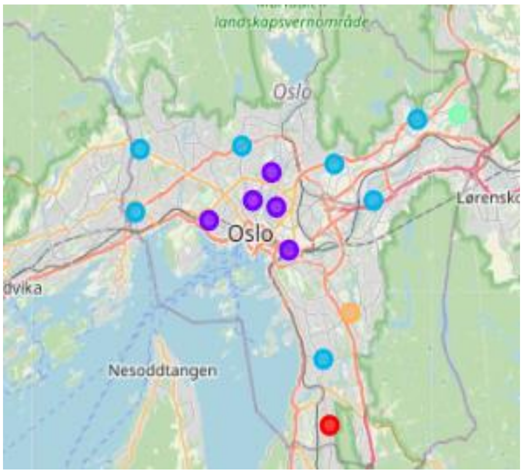

Fig. 4. Map showing the found clusters (Purple = Cluster).

There were a few limitations. Due to the emergency of the buyer, we could not assess the efficiency of our cluster selection. Also, other features, such as number of offices in the nearby areas or sound contamination may influence the decision of the residents into where they prefer to go to a restaurant. So, these features should be added in the future to improve the algorithm.

In conclusion, our results can be used for the buyer to select the best Borough to place their restaurant. The same methodology can be applied for coffee shops and bars. Thus, we believe that his could be automatised and improved to reach a bigger audience or businesses.

5. References

Buitinck, L. *et al.* (2013) 'API design for machine learning software: experiences from the scikit-learn project', *European Conference on Machine Learning and Principles and Practices of Knowledge Discovery in Databases*, pp. 1–15.

Carriere, D. (2018) 'geocoder 1.38.1'. Available at: https://pypi.org/project/geocoder/.

Foursquare Developers (2020) 'Foursquare API'. Available at: https://developer.foursquare.com/docs/api-reference/venues/details/.

Jiménez Sigstad, M. Á. (2020) 'Capstone W2.ipynb'. Available at: https://github.com/mariaangelesjs/Coursera_Capstone/blob/master/Capstone W2.ipynb.

Pedregosa, F. *et al.* (2011) 'Scikit-learn: Machine Learning in Python', *Journal of Machine Learning Research*, 12, pp. 2825–2830.

Richardson, L. (2020) 'beautifulsoup4 4.9.3'. Available at: https://pypi.org/project/beautifulsoup4/.

Wikipedia (2020) *List of Boroughs of Oslo*. Available at: https://en.wikipedia.org/wiki/List_of_boroughs_of_Oslo.