1. **Introduction**

This is the capstone project of Michiel Mathijssen in order to obtain the IBM credential.

1. **What are the business drivers of a restaurant:**

Economic and market outlook

Benchmarking of performance vs peers

Competitors’ strategies and in particular their approach to discounting

Site by site analysis including impact of :

i) own and competitor openings;

ii) owned vs franchised sites

iii) “vintage analysis”

Potential for further sites

Competitors’ roll out strategies

Property optimisation

‘=> LOCATION in function of your competitors plays an important role

1. **Business problem and target audience**

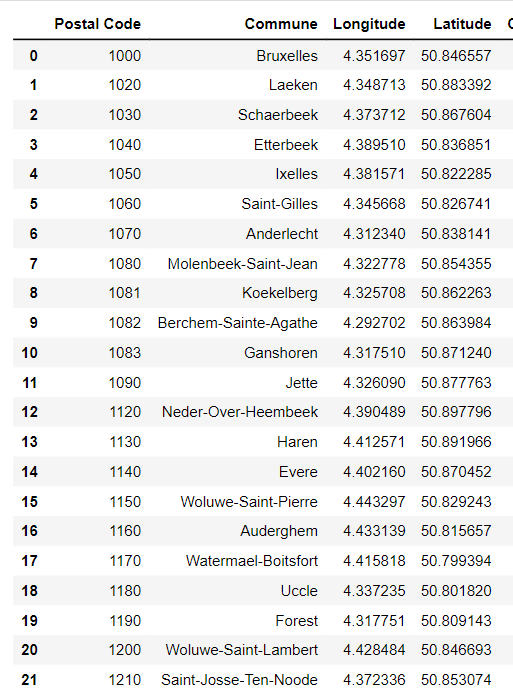
There are entrepreneurs who want to open a press shop or a supermarket store in brussels that mainly focus on the people who travel. I.e. they want to sell high margin products that are readily available. The target audience would notice this store because they are passing by or because they need something quickly that is nearby (e.g. you’re cooking and you ran out of eggs). In this case the consumers would be less price sensitive and would just buy it on-the-go.

1. **Description of the data**

The data used consisted of the all 22 communes in brussels together with there geolocation.

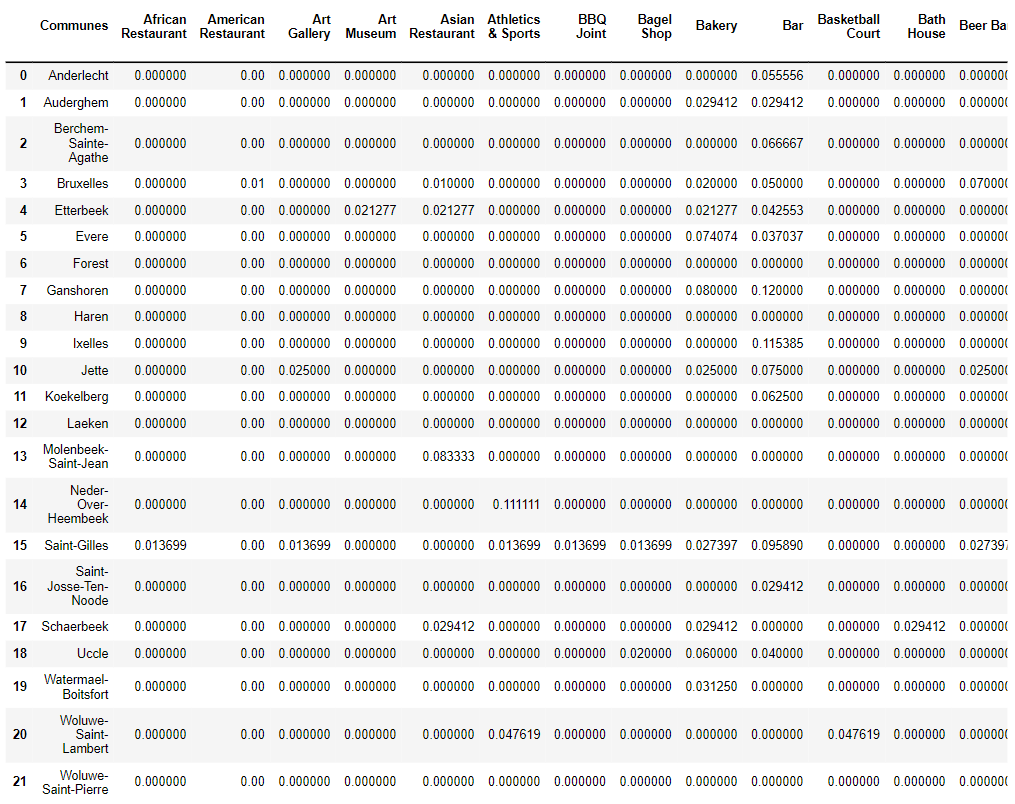
Geopy api used to obtain the latitude and longitude of the respective districts.

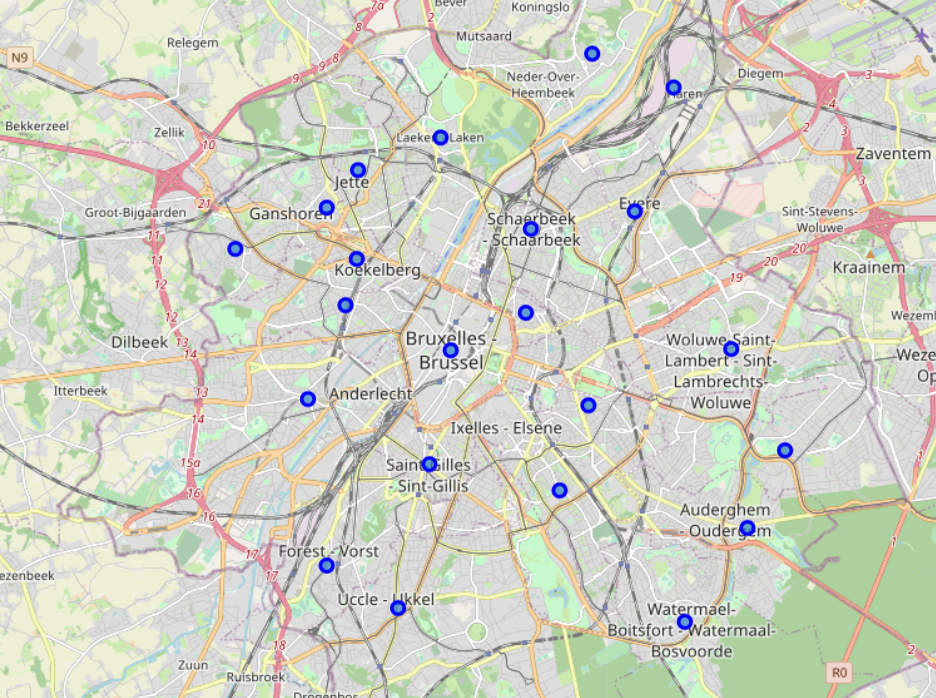
Removed any missing values presented in the dataset. The final dataset consisted of 4 columns being 1) postal code 2) name of the commune 3) longitude & 4) latitude.



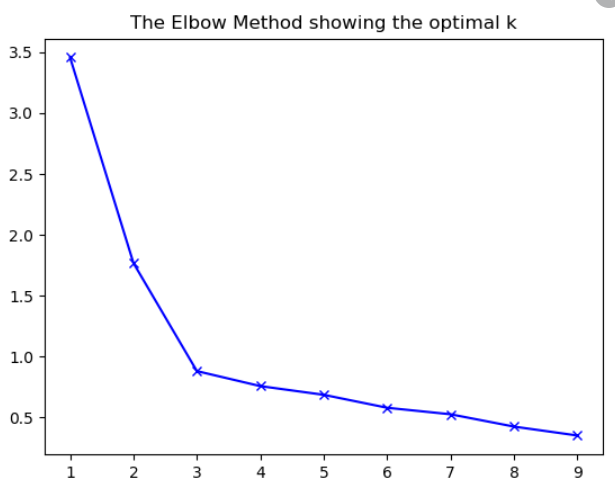
1. **Methodology**

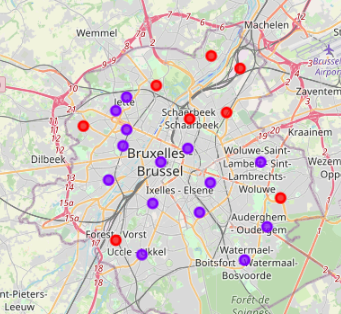
Next obtained all the venues with a limited of 100 and a radius of 500 for the venues. Then we performed one-hot encoding to obtain the most common venues in each commune. We used the obtained venues in the dataset to find the best locations for a press shop or convenience store in brussels that targets travellers or people who quickly need an necessity (e.g. eggs, milk, one person meals to heat up in the microwave etc.)



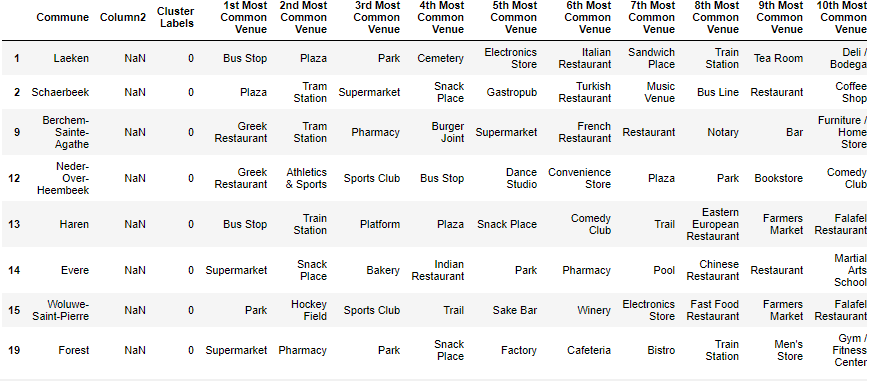


We used the K-means algorithm to cluster the communes by using the elbow method ot find the optimum value of k for the given data. This resulted in 2 clusters.





**CLUSTER 1:**



**CLUSTER 2:**



1. **Result section and discussion**

Based on the elbow method we were able to identify 2 big clusters in brussels that identified the main travel zone and the frequency of restaurants in those communes. We can conclude that the second identified cluster is suitable for further market investigation to open a press shop and/or convenience store that targets travellers.

1. **Conclusion and further recommendations**

Using the K-means clustering algorithm allowed us to divide the city of brussels into two clusters 1) travel area and 2) tourist and living area. The first cluster identified fits our description of the area which fits the criteria to set up a press shop or convenience store that targets the travellers.

Future possible research could make use of other significant factors that are business drivers to set up this type of business which includes foot traffic, accessibility, competition, pricing analysis. These analysis could be an added value to help make our analysis and conclusion more accurate.