

A New Food Warehouse in Leuven (Business Recommendation)



1. Introduction: Business Problem

Leuven is the capital of the province of Flemish Brabant in Belgium. It is located about 25 kilometers east of Brussels. The municipality itself comprises the historic city and the former neighboring municipalities of Heverlee, Kessel-Lo, Wilsele and Wijkmaal. It is the eighth largest city in Belgium and the fourth in Flanders with more than 100,244 inhabitants (Federal Ministry of Home Affairs, 1/11/2016).

Leuven is home to the KU Leuven, the largest and oldest university of the Low Countries and the oldest Catholic university still in existence. The related university hospital of UZ Leuven is one of the largest hospitals in Europe. The city is also known for being the headquarters of Anheuser-Busch InBev, the world's largest brewer and one of the five largest consumer-goods companies in the world.¹

With the knowledge of issues mentioned above over Leuven, an investor in the food sector would like to build up a new supermarket/food warehouse in Leuven (Belgium). The investor wants to deliver fresh and high-quality food products to places as Restaurants, Fast Food Stores, Bakeries, Breakfast Venues, Breweries and Cafés. The investor aims to build a food warehouse for the groceries and products bought from local villagers and farmers, in this sense he also would like to get local acceptance and support of people that will also provide more and satisfied customers. Customer satisfaction and Service quality are always the main stones in the philosophy of the company.

To get a prestigious name in the sector, besides more profit in terms of money and time, he targets with a quick and effective response to the customers. So, the location of the food warehouse has utmost importance that will provide smooth and timely delivery of products from local farmers and deliver to the related customers (venues) in the city.

The investor would also like to minimize transportation cost by building the warehouse in a close and optimized location to its customers. There are several (5) neighborhoods (Leuven Center, Heverlee, Kessel-Lo, Wilsele and Wijkmaal) in city Leuven that our investor set the focus. Determining the right neighborhood is our goal, and we will use a Machine Learning algorithm to cluster the neighborhoods and make a recommendation to the investor as a solution to this problem.

2. Data

a. Data We Need

The investor would like to invest in a food warehouse in Leuven but which neighborhood is not determined. The Postal Codes of Leuven (3000,3001,3010,3012,3018) that provides us with the necessary information to find which neighborhoods (3000 Leuven Center (Old City), 3001 Heverlee, 3010 Kessel-Lo, 3012 Wilsele and 3018 Wijkmaal) are our candidates for the location of the warehouse.

In this sense, we can get the data including postal codes and names of neighborhoods in Leuven from open source internet web pages. For this case, I got the information from the official site of Flemish Government (vlaanderen.be).²

We needed the geo-locational data of this specific borough (Leuven city) and the five neighborhoods for our analysis. This information can be gathered in several ways, such as using Geocoder package and by transferring pre-prepared open data source.³ Finally, I would get the below sample data for each neighborhood:

PostalCode/ Borough/ Neighborhood/ Latitude/ Longitude/
3001/ Leuven/ Heverlee/ 50.851729/ 4.693131/

Further on, we would need data about different venues in the various neighborhoods of city Leuven. "Foursquare" application can provide us with the locational information of the sites that we needed. By locational information of the center of each neighborhood and by inputting a self-determined distance from city center we can get detailed information about the venues that lays inside the circular interest area we have described. For each venue, we can gather information including its specific location information (latitude, longitude), category and popularity of the venue.

After gathering and merging necessary information, we can do clustering to see the similarities and differences of the neighborhoods. Then we can sum up the number of food product customers per neighborhood to decide which one is potentially a better choice for investment.

Foursquare will provide us data such as the following information:

Neighborhood/ Neighborhood Latitude/ Neighborhood Longitude/ Venue/ Venue Latitude/ Venue Longitude/ Venue Distance/ Venue Summary/ Venue Category/
Wijkmaal 50.926428/ 4.700121/ Halte Waterstraat/ 50.922755/ 4.691789/ 713/ This spot is popular/ Bus Stop/

b. Data Gathering

To get the necessary data from several sources, I used different methods and libraries. I scrape the postal codes and the names of the neighborhoods from the web sites by using Beautiful Soup and Requests and transfer into data frames by using Pandas.

After building a data frame of the postal code of each neighborhood along with the borough name and neighborhood name, we should get the coordinates of each neighborhood to be able to utilize Foursquare API. I had several alternatives here.

(1) Google Maps Geocoding API (not free): In an older version of this course, we were leveraging the Google Maps Geocoding API to get the latitude and the longitude coordinates of each neighborhood. However, recently Google started charging for their API.⁴ I skipped this option.

(2) Geocoder Python package⁵: The problem with this Package is that you have to be persistent sometimes to get the geographical coordinates of a given postal code.

So, you can make a call to get the latitude and longitude coordinates of a given postal code, and the result would be None, and then make the call again, and you would get the coordinates. So, to make sure that you get the coordinates for all of our neighborhoods, you can run a while loop for each postal code.

(3) Open Data source as Json/csv: Given the fact that Geocoder package can be very unreliable, I tried the third alternative. I searched and found an open source json and a csv file that has the geographical coordinates of each postal code in Belgium⁶. I set the necessary parts out of this data into another data frame. After some preprocessing my Leuven Data Frame was ready for further steps.

	PostalCode	Borough	Neighborhood	Latitude	Longitude
0	3000	Leuven	Leuven	50.881253	4.692990
1	3001	Leuven	Heverlee	50.851729	4.693131
2	3010	Leuven	Kessel-Lo	50.889915	4.730761
3	3012	Leuven	Willesele	50.909536	4.713629
4	3018	Leuven	Wijgmaal	50.926428	4.700121

Figure 1- Leuven Neighborhoods Location Data

I continued data gathering here to explore, analysis and cluster the neighborhoods in Leuven. Firstly, I created a map of Leuven with neighborhoods superimposed on top by using Folium.

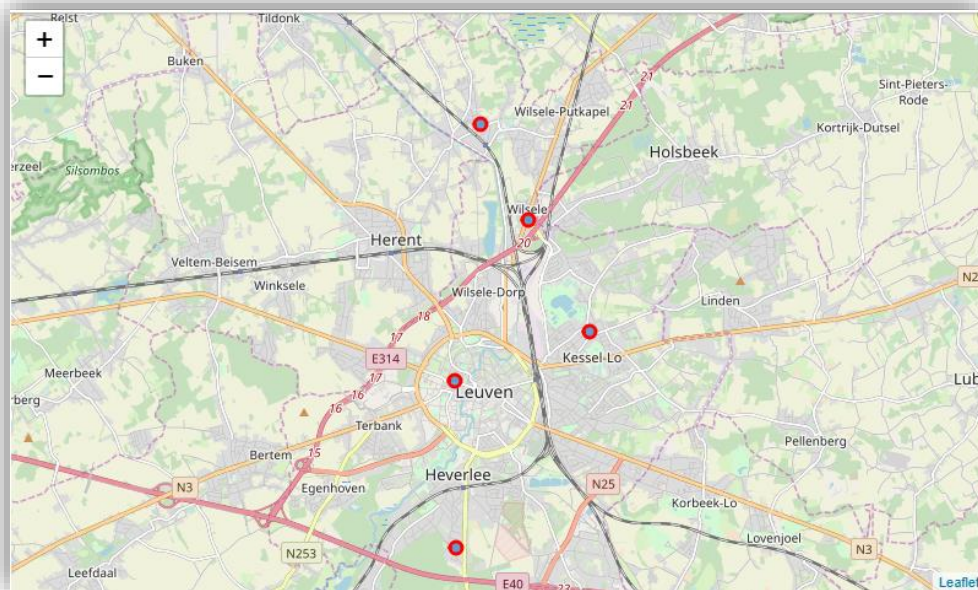


Figure 2-Neighborhoods in Leuven

In this step, I would like to show the map of Leuven to get an overall idea how could we get venues data from Foursquare and how could I determine optimum values of the radius values from the center points of neighborhoods. With the information from Leuven Data Frame and the intuition from the overview map of Leuven, I was ready to utilize Foursquare API to get venues data.

After connecting to the Foursquare API to gather information about venues in each neighborhood, I applied the radius value as 1500 meter. It means that we have asked Foursquare to find venues that are at most 1500 meter far from the center of the neighborhood. And I saved the information into a data frame. I got as a result comprising 294 venues in the data set. I had 110 unique venues in the data.

```
# Let's check the size of the resulting dataframe
print(leuven_venues.shape)
leuven_venues.tail()
```

(294, 9)

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Distance	Venue Summary	Venue Category
289	Wijgmaal	50.926428	4.700121	Sportlokaal	50.924938	4.686548	966	This spot is popular	Bar
290	Wijgmaal	50.926428	4.700121	Lijnloperspad	50.917226	4.706344	1113	This spot is popular	Bike Trail
291	Wijgmaal	50.926428	4.700121	Apotheek Haegemans	50.917074	4.685834	1445	This spot is popular	Pharmacy
292	Wijgmaal	50.926428	4.700121	Apotheek Adriaens	50.922932	4.720527	1483	This spot is popular	Pharmacy
293	Wijgmaal	50.926428	4.700121	Nachtwinkel Euro	50.924335	4.721062	1487	This spot is popular	Convenience Store

Figure 3- Venues per Neighborhoods Data Frame

References:

- [1 https://en.wikipedia.org/wiki/Leuven](https://en.wikipedia.org/wiki/Leuven)
- [2 https://www.vlaanderen.be/gemeenten-en-provincies/provincie-vlaams-brabant/leuven](https://www.vlaanderen.be/gemeenten-en-provincies/provincie-vlaams-brabant/leuven)
- [3 https://raw.githubusercontent.com/jief/zipcode-belgium/master/zipcode-belgium.csv](https://raw.githubusercontent.com/jief/zipcode-belgium/master/zipcode-belgium.csv)
- [4 http://geoawesomeness.com/developers-up-in-arms-over-google-maps-api-insane-price-hike/](http://geoawesomeness.com/developers-up-in-arms-over-google-maps-api-insane-price-hike/)
- [5 https://geocoder.readthedocs.io/index.html](https://geocoder.readthedocs.io/index.html)
- [6 https://raw.githubusercontent.com/jief/zipcode-belgium/master/zipcode-belgium.json](https://raw.githubusercontent.com/jief/zipcode-belgium/master/zipcode-belgium.json)

Levent BINGOL

levobelgium@gmail.com
<https://www.linkedin.com/in/levent-bingöl-957160139/>
<https://github.com/LeventBINGOL>