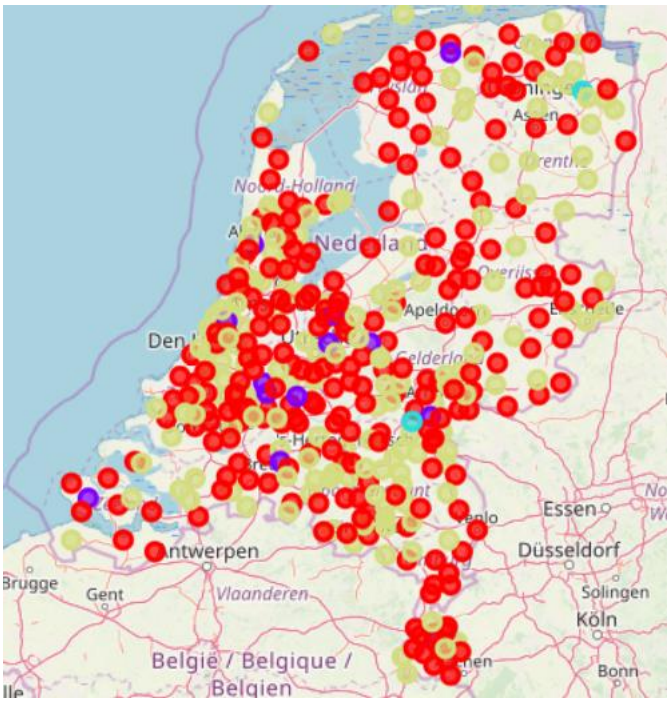


OPENING A RESTAURANT IN NETHERLANDS.

ABSTRACT

Pablo is going to Netherlands to open a new restaurant, where should he go?

Pablo Esquerdo

IBM DATA SCIENCE-FINAL CAPSTONE

1. Introduction

1.1 Background

Netherlands is one of the country's with the best work-life balance, most of Spanish people move out to places like Germany or Netherlands because their balance. Exhausted of working 60 hours in Spain, getting poor salaries at the end of the month they go abroad to find new opportunities and start their new life.

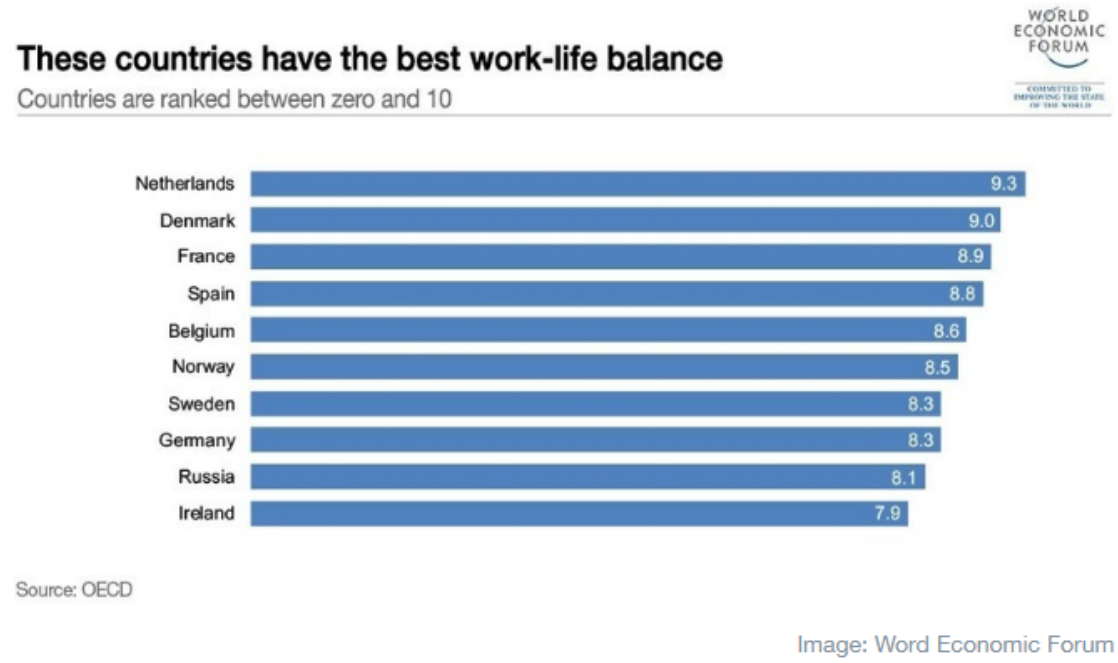


Image: Word Economic Forum

1.2 Problem

Pablo is thinking to open a new restaurant in Netherlands and cook typical dishes in some city in the Netherlands, but the obviously question is, where in the Netherlands?

1.3 Interest

This will be interesting for any person planning to move to Netherlands for opening a restaurant, but also will be interesting for those persons thinking to open any business related thing in the Netherlands because I will display the most common venues in each of the 409 different cities in Netherlands, so maybe someone is planning to open a Supermarket, or an IT business and they can find interesting where it is the best place for their purpose.

2. Data acquisition and cleaning

2.1 Data sources

I will use the .csv file provided by the following page: https://simplemaps.com/data/nl-cities consisting on all the different cities within Netherlands.

In addition, this dataset will provide me the different latitude and longitude of all of those cities, which will be useful for printing them in a map using Foursquare API. Also, the dataset

will provide the population of the 14 first cities (or the ones with more population like The Hague or Amsterdam).

2.2 Data cleaning and how will be processed

First of all, I sorted the different cities by their population to have the most populated at first because those are usually the most common chosen by the people who is going to Netherlands.

After, I will delete the rows that will not be useful for this purpose, or at least for this project, having at the end only a table with the name of the city, their latitude and longitude that will be necessarily for showing into the map

	City	Latitude	Longitude
0	The Hague	52.083333	4.300000
1	Amsterdam	52.350000	4.916667
2	Rotterdam	51.916667	4.500000
3	Utrecht	52.093813	5.119095
4	Eindhoven	51.450000	5.466667

Once I have the different cities with their latitude and longitude will be the time to print them in the map, for this purpose I will use the folium library of Python to display it in the map.

Note: At this point I have not worked with the data yet, this is only for showing the different cities into the map.

Once I have the different data displayed into the map, now is time to start applying the Data Analysis and fetch the data in different Clusters with “Similar data”

3. Exploratory Data Analysis

Now that we have all the cities represented in the map with their latitude and longitude, I will use my Foursquare credentials to be able to fetch all the venues surrounding that city with an radius of 500, so I will be only focused in the center on the city because there is where usually people go for restaurants.

Once I have every venue in the city, I will use One Hot Encoding to extract all the dummy variables that will be necessarily to implement in the future the K means clustering.

Hockey Field	Hockey Rink	Home Service	Hookah Bar	Hostel	Hot Dog Joint	Hot Spring	Hotel	Hotel Bar	Hunting Supply	IT Services	Ice Cream Shop	Res
0	0	0	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	1	0	0	0	0	
0	0	0	0	0	0	0	0	0	0	0	0	

Now that I have the different venues like Hotels, Ice Cream shops, Restaurants and so on, I will obtain the mean of all of them to extract the mean of each venue in that city.

```
----Amsterdam----
                venue  freq
0         Coffee Shop  0.10
1                Café  0.07
2          Restaurant  0.07
3        Train Station  0.03
4  French Restaurant  0.03


----Apeldoorn----
                venue  freq
0         Coffee Shop  0.11
1          Restaurant  0.08
2  French Restaurant  0.08
3               Diner  0.06
4      Clothing Store  0.06
```

With this data we can observe that Amsterdam the most common thing is Coffee shops, but also I can see that Apeldoorn is also coffee shop, that means that probably in the future when I will apply clustering, those cities will be in the same cluster.

Once I get the frequency of each venue in each city, I will make a table to make it more user friendly and easily understandable.

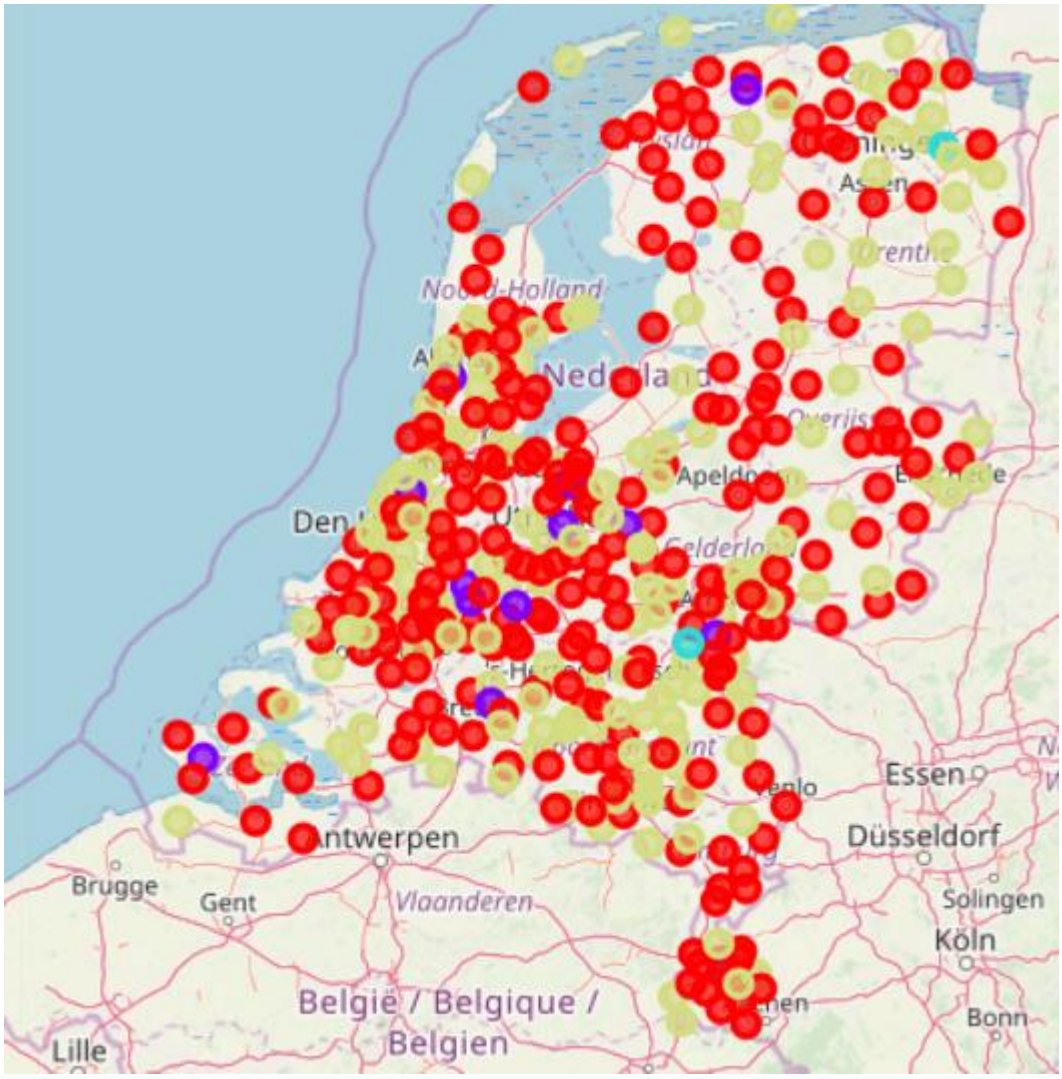
	City	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Aalsmeer	Restaurant	Drugstore	Fish Market	Sporting Goods Shop	Fast Food Restaurant	Park	Bar	Falafel Restaurant	Gym / Fitness Center	Chinese Restaurant
1	Aalten	Supermarket	German Restaurant	Pub	Drugstore	Discount Store	Department Store	Bar	English Restaurant	Event Space	Factory
2	Abbenbroek	Supermarket	Café	Coffee Shop	Yoga Studio	English Restaurant	Event Space	Factory	Falafel Restaurant	Farm	Farmers Market
3	Alblasserdam	Supermarket	Pizza Place	Grocery Store	Shopping Mall	Bus Stop	Snack Place	Pool	Creperie	Farm	English Restaurant
4	Alkmaar	Furniture / Home Store	Supermarket	Playground	Sporting Goods Shop	Electronics Store	Beach	Thrift / Vintage Store	Discount Store	Sports Bar	Gym / Fitness Center

With that table, I will Cluster the data in 4 different tables (the optimum number of clusters) and finally, to be able to represent it again I will need to join the latitude and longitude in that table, having as a result the following one:

	City	Latitude	Longitude	Cluster Labels	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	The Hague	52.083333	4.300000	0.0	Italian Restaurant	Restaurant	French Restaurant	Hotel	Coffee Shop	Bar	Wine Bar	Gym	Indian Restaurant	Supermarket
1	Amsterdam	52.350000	4.916667	0.0	Coffee Shop	Café	Restaurant	Platform	Bus Station	Breakfast Spot	Bus Stop	Sandwich Place	French Restaurant	Roof Deck
2	Rotterdam	51.916667	4.500000	0.0	Boat or Ferry	Bar	Breakfast Spot	Pool	Bistro	Island	Bus Stop	Field	Factory	Falafel Restaurant
3	Utrecht	52.093813	5.119095	0.0	Restaurant	Bar	Coffee Shop	Burger Joint	Vietnamese Restaurant	Plaza	Indonesian Restaurant	Gay Bar	Sandwich Place	Italian Restaurant
4	Eindhoven	51.450000	5.466667	0.0	Turkish Restaurant	Supermarket	Bakery	Kids Store	Fruit & Vegetable Store	Restaurant	Spanish Restaurant	Department Store	Grocery Store	Falafel Restaurant

Note: As you can see, the cluster labels are floats, I will need to make them int for be able to represent them with different colors in the map with each city separated by their cluster.

Now that I have the most common venue, the City, the longitude and which Cluster is inside in, it is time to plot it in the map to see the results and the different clusters created in Netherlands.



And if I have a look to the data:

Cluster 0 (Brown cities)

	City	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	The Hague	Italian Restaurant	Restaurant	French Restaurant	Hotel	Coffee Shop	Bar	Wine Bar	Gym	Indian Restaurant	Supermarket
1	Amsterdam	Coffee Shop	Café	Restaurant	Platform	Bus Station	Breakfast Spot	Bus Stop	Sandwich Place	French Restaurant	Roof Deck
2	Rotterdam	Boat or Ferry	Bar	Breakfast Spot	Pool	Bistro	Island	Bus Stop	Field	Factory	Falafel Restaurant
3	Utrecht	Restaurant	Bar	Coffee Shop	Burger Joint	Vietnamese Restaurant	Plaza	Indonesian Restaurant	Gay Bar	Sandwich Place	Italian Restaurant
4	Eindhoven	Turkish Restaurant	Supermarket	Bakery	Kids Store	Fruit & Vegetable Store	Restaurant	Spanish Restaurant	Department Store	Grocery Store	Falafel Restaurant
5	Haarlemmerliede	Bar	Sporting Goods Shop	Monument / Landmark	Comedy Club	Farm	Yoga Studio	Fast Food Restaurant	Event Space	Factory	Falafel Restaurant
9	Leeuwarden	Furniture / Home Store	Snack Place	Fish & Chips Shop	Bus Stop	Music Venue	Yoga Studio	Event Space	Factory	Falafel Restaurant	Farm
10	Maastricht	Hotel	Italian Restaurant	Pub	French Restaurant	Coffee Shop	Supermarket	Asian Restaurant	Clothing Store	Theater	Shoe Store
12	Assen	Restaurant	Shopping Mall	Coffee Shop	Diner	Italian Restaurant	Supermarket	Bar	Ice Cream Shop	Bakery	Sandwich Place

These cities are mostly the touristic ones and as you can see, they have plenty of restaurants, Bars or food related business. In addition, Apeldoorn is also in this cluster as we predicted.

Cluster 1 (Red cities)

	City	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
159	Oude Pekela	Bar	Bus Stop	Yoga Studio	Fast Food Restaurant	Event Space	Factory	Falafel Restaurant	Farm	Farmers Market	Field
199	IJmuiden	Bus Stop	Middle Eastern Restaurant	History Museum	Volleyball Court	Paper / Office Supplies Store	Farmers Market	English Restaurant	Event Space	Factory	Falafel Restaurant
205	Montfoort	Bus Stop	Bookstore	Shoe Store	Tennis Court	Department Store	Supermarket	Event Space	Factory	Falafel Restaurant	Farm
246	Appingedam	Department Store	Bus Stop	Fast Food Restaurant	Field	Event Space	Factory	Falafel Restaurant	Farm	Farmers Market	Yoga Studio
296	Nieuwerkerk aan den IJssel	Motorcycle Shop	Bus Stop	Yoga Studio	Fast Food Restaurant	English Restaurant	Event Space	Factory	Falafel Restaurant	Farm	Farmers Market
308	Bergen	Bus Stop	Yoga Studio	Fast Food Restaurant	English Restaurant	Event Space	Factory	Falafel Restaurant	Farm	Farmers Market	Field
316	Leusden	Bus Stop	Park	Yoga Studio	Farmers Market	English Restaurant	Event Space	Factory	Falafel Restaurant	Farm	Fast Food Restaurant
329	Rhenen	Bus Stop	Hockey Rink	Snack Place	Gym Pool	Farmers Market	English Restaurant	Event Space	Factory	Falafel Restaurant	Farm

This cluster seem to have the cities that are in between tourist places, or probably for working companies or fabrics because most of them have lot of bus tops, factories or Farms.

Cluster 2 (Blue cities)

	City	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
127	Tilburg	Basketball Court	Yoga Studio	Fast Food Restaurant	English Restaurant	Event Space	Factory	Falafel Restaurant	Farm	Farmers Market	Field
366	Son en Breugel	Basketball Court	Yoga Studio	Fast Food Restaurant	English Restaurant	Event Space	Factory	Falafel Restaurant	Farm	Farmers Market	Field

About the Blue cities seem to be the same city and in the data that we fetched they had different name because they have the same Venues.

Cluster 3 (Violet cities)

	City	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
6	Groningen	Supermarket	Pizza Place	Indian Restaurant	Fast Food Restaurant	Bus Stop	Beer Garden	Sandwich Place	Bar	Bakery	Indonesian Restaurant
7	Arnhem	Bus Stop	Supermarket	Pizza Place	Snack Place	Liquor Store	Restaurant	Fast Food Restaurant	Farm	English Restaurant	Event Space
8	's-Hertogenbosch	Bus Stop	Electronics Store	Discount Store	Supermarket	Sports Club	Grocery Store	Gym / Fitness Center	Shoe Store	Plaza	Fast Food Restaurant
11	Zwolle	Hotel	Gym	Arcade	Coffee Shop	Supermarket	Yoga Studio	Farmers Market	English Restaurant	Event Space	Factory
17	Bergeijk	Supermarket	Bar	Snack Place	Department Store	Middle Eastern Restaurant	Drugstore	Restaurant	Chinese Restaurant	Diner	Shoe Store
19	Grootegast	Supermarket	Fast Food Restaurant	Electronics Store	English Restaurant	Event Space	Factory	Falafel Restaurant	Farm	Farmers Market	Yoga Studio
27	Diever	Gastropub	Supermarket	Bar	Diner	Dessert Shop	English Restaurant	Factory	Falafel Restaurant	Farm	Farmers Market
28	Uithuizen	Plaza	Fast Food Restaurant	Discount Store	Train Station	Chinese Restaurant	Supermarket	Department Store	Electronics Store	Field	Fireworks Store
31	Zundert	Bar	Supermarket	Friterie	Restaurant	Gym	Shawarma Place	Clothing Store	Drugstore	Bus Station	Farm
32	Nunspeet	Bowling Alley	Supermarket	Snack Place	Pool	Electronics Store	English Restaurant	Event Space	Factory	Falafel Restaurant	Farm

Finally, on this cluster it is mainly Supermarkets and Markets, so it seems to be a good place to buy food for a restaurant.

CONCLUSIONS

The question is: Where Pablo should open a Restaurant? Answering this question I would say that Pablo should go to a tourist place like Amsterdam, The Hague, Rotterdam or one of those because it is inside the Cluster 0 (Brown color) meaning that it is plenty of restaurants (Which means that they really work there) And also those cities because is where there is more population.

Also, I would suggest to Pablo to go to a city plenty of Supermarkets or markets like in the last cluster because probably there he can finds better quality food, and way cheaper.

FUTURE IMPROVEMENTS.

- Fetching Data about rent prices for each city and compare them
- Fetching Data about benefits of different restaurants in each city
- Fetching Data about cost of living in each city
- Analyze each city by their cities surrounding