

## Capstone Project - Final Assignment



If you have time to visit one city in the Netherlands?  
What city would you visit?

# Contents

---

- Introduction
- Project objectives
- Data
- Methodology
- Results
- Discussion

# Introduction

---

## **Business problem**

The Netherlands is a great country to visit for a short break or a longer holiday. And there are many nice cities known for their restaurants, nightlife, museums, market, squares and parks. So deciding which Netherlands city to visit can be difficult. Especially, if you have time to visit one city. What city would you visit?

## **Target audience**

In this project data is collected to provide a recommendation to people who wants to visit the Netherlands for a short stay. The recommendation are a help to make a decision which city to visit in the Netherlands. They can choose a city by venue type and venues ratings to see which place and venue is recommended by other visitors.

# Objectives

---

The objective of this project is to compare the 20 most popular tourist cities in the Netherlands and determine how similar or dissimilar they are.

The aim is to determine the similarity or dissimilarity of the 20 cities in the Netherlands based on tourism venues like; food, outdoors & recreation, night live and travel & transport by performing a K-means clustering.

Analyze the diversity of the city venues:

- a. What are the most common venues types (venues count)
- b. What are the recommending scores of the top 10 venues based on 'ratings'
- c. Which 5 cities has the highest mean recommendation scores based on all venue 'ratings'

# Data

---

## **Foursquare**

For this project the data of foursquare (website: [www.foursquare.com](http://www.foursquare.com)) is used for information about venues and ratings. The Foursquare explore function will be used to get the most common venue categories in each city, and then use this feature to group the cities into clusters. Another venue query is performed to retrieve venue ratings for each city to find the most popular places. Note that there is max query limit of 50 ratings in FourSquare API for free subscription.

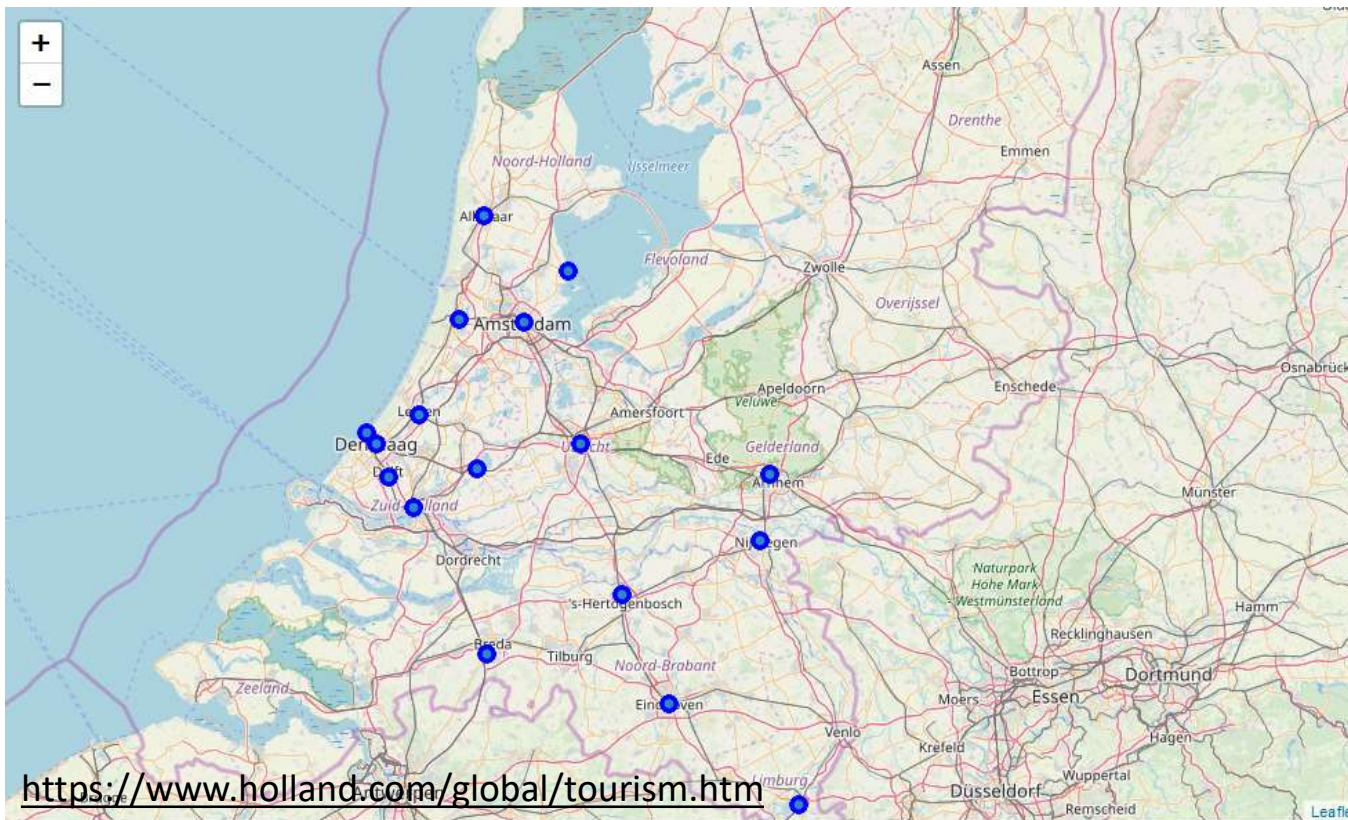
## **Popular cities**

A list of the most popular destinations in the Netherlands are taken from the website <https://www.holland.com/global/tourism.htm>. A list of cities is given at the next slide.



# Data

Map of the 20 popular destinations in the Netherlands



Amsterdam

Leeuwarden

Arnhem

Gouda

Rotterdam

Leiden

Breda

Groningen

The Hague

Haarlem

Delft

Utrecht

Volendam

Den Bosch

Maastricht

Alkmaar

Eindhoven

Nijmegen

Roermond

Scheveningen

# Methodology

---

## **Collect location data**

First the Foursquare API is used to retrieve data from the Foursquare database to explore a city and to get trending venues around a point of interest. Another venue query is performed to retrieve venue ratings for each city to find the most popular places.

## **Clustering**

After collect data with the Foursquare API we start analyzing the list of cities using location data. To explorer cities, segment them and group them into clusters to find similarity or dissimilarity an unsupervised k-means clustering model is used. K-means clustering is one of the popular unsupervised machine learning algorithms.

# Results

List of venues for every city by popularity and proximity

```
In [15]: NL_city_food_venues['Primary_category'] = "Food"
NL_city_food_venues.head()
```

Out[15]:

	City	City Latitude	City Longitude	VenueID	VenueName	score	category	catID	latitude	longitude	Primary_categor
0	Amsterdam	52.37454	4.897976	51d98583498e78da6626be60	De Koffieschenkerij	nan	Coffee Shops	4bf58dd8d48988d1e0931735	52.374043	4.898427	Foo
1	Amsterdam	52.37454	4.897976	4a2704b4f964a520f4851fe3	Manneken Pis	nan	Fast Food Restaurants	4bf58dd8d48988d16e941735	52.375708	4.896237	Foo
2	Amsterdam	52.37454	4.897976	4a26ffe4f964a5205a811fe3	The Bulldog Mack	nan	Bars	4bf58dd8d48988d116941735	52.373200	4.897321	Foo
3	Amsterdam	52.37454	4.897976	4af3e8e7f964a52099ef21e3	Starbucks	nan	Coffee Shops	4bf58dd8d48988d1e0931735	52.379117	4.899273	Foo

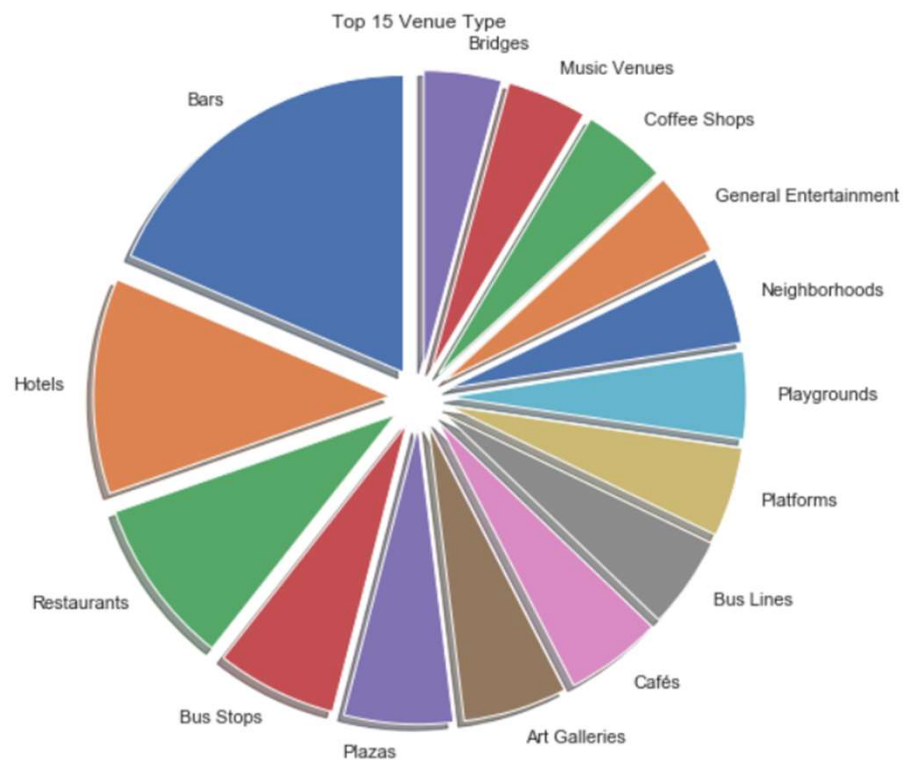
A list of venues is created with:

- Food Venues (Restaurants, Fastfoods, etc.)
- Outdoor and recreation venues (beach, castle, swimming pool)
- Arts & Entertainment venues (Aquarium, Movie Theater, Museum, etc)
- Nightlife Spot venues (Beer Bar, Karaoke Bar, Lounge etc)
- Travel & Transport venues (Airport, Hotel, Train Station etc)



# Results

The most common venue type in all cities are bars, hotels, restaurants and bus stops



	category	VenueName
0	Bars	141
1	Hotels	89
2	Restaurants	69
3	Bus Stops	50
4	Plazas	45
5	Art Galleries	43
6	Cafés	40
7	Bus Lines	38
8	Platforms	37
9	Playgrounds	36
10	Neighborhoods	36
11	General Entertainment	35
12	Coffee Shops	35
13	Music Venues	33
14	Bridges	32

# Results

Identify the top ten venues for each city

Out [123]:

	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	Alkmaar	Bars	Bus Lines	Beer Gardens	Art Galleries	Plazas	Parks	Bus Stops	Cafés	Restaurants	Coffee Shops
1	Amsterdam	Hotels	Bars	Bridges	Plazas	Fast Food Restaurants	Museums	Coffee Shops	Neighborhoods	Hotel Bars	Erotic Museums
2	Arnhem	Hockey Fields	Other Great Outdoors	Neighborhoods	Fields	Campgrounds	Bus Stops	Scenic Lookouts	Cities	Yoga Studios	Fish & Chips Shops
3	Breda	Restaurants	Bars	Pools	Coffee Shops	Dog Runs	Other Great Outdoors	Steakhouses	French Restaurants	Breakfast Spots	Hotels
4	Delft	Playgrounds	Bus Stops	Music Venues	Restaurants	Pubs	Bars	Gay Bars	Hotels	Other Great Outdoors	Bus Lines
5	Den Bosch	Harbors / Marinas	Boats or Ferries	Lakes	Other Great Outdoors	Diners	Bus Stops	Lounges	Skate Parks	Bed & Breakfasts	Fish & Chips Shops
6	Eindhoven	Bars	Hotels	Plazas	Movie Theaters	Multiplexes	Restaurants	Platforms	Neighborhoods	Cafés	Gyms
7	Gouda	Platforms	Bus Lines	Theaters	Dog Runs	Playgrounds	Snack Places	Bus Stops	Fields	Coffee Shops	Yoga Studios
8	Groningen	Bars	Restaurants	Bus Lines	Pubs	Cafés	Coffee Shops	Art Galleries	Concert Halls	Bed & Breakfasts	Bus Stops
9	Haarlem	Restaurants	Bars	Bridges	Bus Lines	Hotels	Nightclubs	Jazz Clubs	Museums	Athletics & Sports	Art Galleries

# Results

Another venue query is performed to retrieve venue ratings for each city

The first 50 venues get a score bases on the ratings in Foursquare. But due to a limit of 500 calls per day a fictive list of scores is created with the Python "random module"

```
In [98]: NL_city_food_venues['score'] = np.random.uniform(4.0,9.5, size=len(NL_city_food_venues)).astype(float).round(1)
```

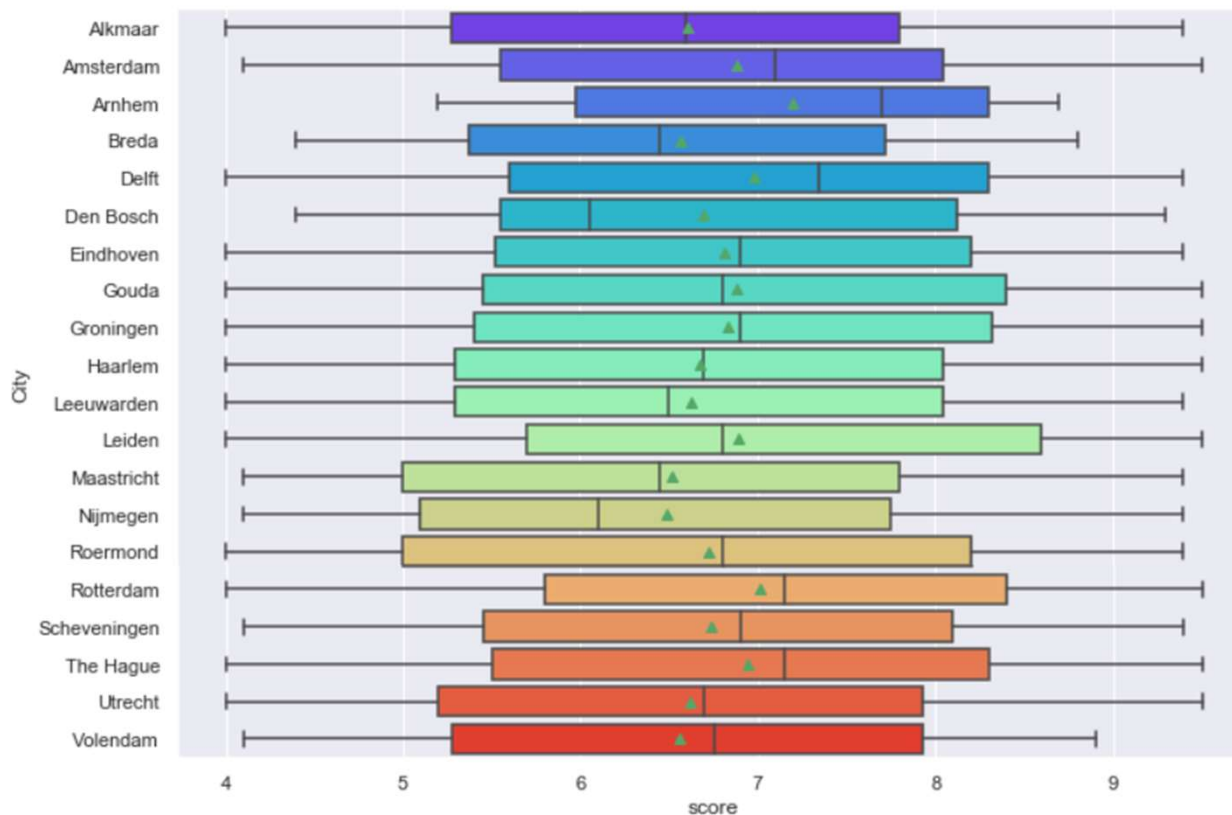
```
In [99]: NL_city_food_venues.head()
```

Out[99]:

	City	City Latitude	City Longitude	VenueID	VenueName	score	category	catID	latitude	longitude	Primary_categor
0	Amsterdam	52.37454	4.897976	51d98583498e78da6626be60	De Koffieschenkerij	7.6	Coffee Shops	4bf58dd8d48988d1e0931735	52.374043	4.898427	Foo
1	Amsterdam	52.37454	4.897976	4a2704b4f964a520f4851fe3	Manneken Pis	7.2	Fast Food Restaurants	4bf58dd8d48988d16e941735	52.375708	4.896237	Foo
2	Amsterdam	52.37454	4.897976	4a26ffe4f964a5205a811fe3	The Bulldog Mack	7.4	Bars	4bf58dd8d48988d116941735	52.373200	4.897321	Foo
3	Amsterdam	52.37454	4.897976	4af3e8e7f964a52099ef21e3	Starbucks	7.5	Coffee Shops	4bf58dd8d48988d1e0931735	52.379117	4.899273	Foo
4	Amsterdam	52.37454	4.897976	57910d45498eeaf3c1c08128	Cannibale Royale	9.0	Burger Joints	4bf58dd8d48988d16c941735	52.375031	4.898690	Foo

# Results

Recommendation scores based on all venue 'ratings' of all cities

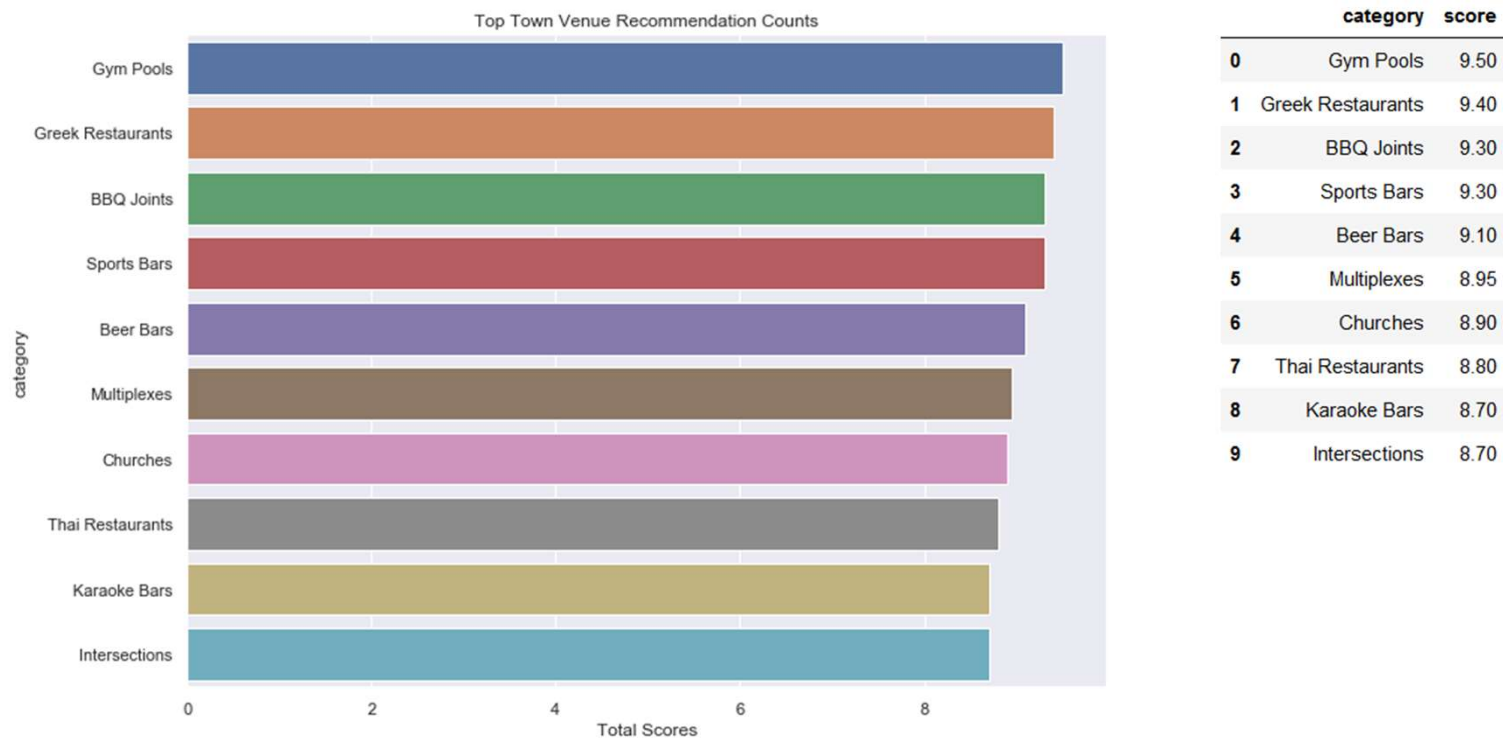


	City	score
0	Arnhem	7.200000
1	Rotterdam	7.011972
2	Delft	6.985000
3	The Hague	6.941892
4	Leiden	6.891463
5	Amsterdam	6.883673
6	Gouda	6.878505
7	Groningen	6.829545
8	Eindhoven	6.817808
9	Scheveningen	6.733858
10	Roermond	6.728696
11	Den Bosch	6.700000
12	Haarlem	6.671429
13	Leeuwarden	6.626984
14	Utrecht	6.617647
15	Alkmaar	6.603125
16	Breda	6.566667
17	Volendam	6.560000
18	Maastricht	6.519737
19	Nijmegen	6.484615

Top 5 cities

# Results

Recommendation scores based on all venue 'ratings' of the top 5 cities



The best rated venues are in the category "Gym pools", "Greek Restaurants" and "BBW Joints"

# Results

Cluster the popular cities in the Netherlands based on K-means

Out[148]:





# Results

Cluster 1: This cluster is the largest cluster with a mix of venues. The most common venues in this clusters are bars, hotels, restaurants

	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
0	Amsterdam	52.374540	4.897976	0	Hotels	Bars	Bridges	Plazas	Fast Food Restaurants	Museums	Coffee Shops	Neighborhoods	Hotels
1	Leeuwarden	53.199685	5.800691	0	Bars	Hotels	Restaurants	Bridges	Pubs	Music Venues	Cafés	Bus Lines	It Restaur
4	Rotterdam	51.922893	4.463179	0	Platforms	Art Galleries	Bars	General Entertainment	Gyms or Fitness Centers	Plazas	Nightclubs	Gay Bars	Cocktail
7	Groningen	53.221792	6.558245	0	Bars	Restaurants	Bus Lines	Pubs	Cafés	Coffee Shops	Art Galleries	Concert Halls	B Breakl
8	The Hague	52.079984	4.311346	0	Hotels	Bars	Plazas	Restaurants	Tram Stations	Cafés	Coffee Shops	Movie Theaters	Thea
9	Haarlem	52.383706	4.643560	0	Restaurants	Bars	Bridges	Bus Lines	Hotels	Nightclubs	Jazz Clubs	Museums	Athleti S
11	Utrecht	52.080986	5.127684	0	Bars	Parks	French Restaurants	Neighborhoods	Bus Stops	General Entertainment	Museums	Cafés	Restaur
14	Maastricht	50.857985	5.696988	0	Hotels	Art Galleries	Other Nightlife	Hotel Bars	Bed & Breakfasts	Lounges	Concert Halls	Movie Theaters	Music Ver
15	Alkmaar	52.633736	4.740203	0	Bars	Bus Lines	Beer Gardens	Art Galleries	Plazas	Parks	Bus Stops	Cafés	Restaur

# Results

Cluster 2: This cluster with 5 cities are different compare to the other cities in the Netherlands. The most common venues in these cities are Playgrounds, general entertainment and bus stops

	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
3	Gouda	52.018119	4.711122	1	Platforms	Bus Lines	Theaters	Dog Runs	Playgrounds	Snack Places	Bus Stops	Fields	Coffee Shops
5	Leiden	52.151816	4.481109	1	Bus Stops	General Entertainment	Playgrounds	Cafés	Bakeries	Bars	Parks	Neighborhoods	Campgrounds
10	Delft	51.999457	4.362725	1	Playgrounds	Bus Stops	Music Venues	Restaurants	Pubs	Bars	Gay Bars	Hotels	Other Great Outdoors
12	Volendam	52.500514	5.077833	1	Bus Stops	Bakeries	Playgrounds	Gyms	Snack Places	Cafés	Fields	Other Great Outdoors	Concert Halls
17	Nijmegen	51.842575	5.838961	1	Playgrounds	Bus Stops	Parks	Snack Places	Hookah Bars	Neighborhoods	Soccer Fields	Yoga Studios	Dance Studios

Cluster 3: This cluster is one city called Den Bosch. This city is different to other cities. This city have a harbor with boats and ferries, lakes and great outdoor venues

	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
13	Den Bosch	51.70997	5.291916	2	Harbors / Marinas	Boats or Ferries	Lakes	Other Great Outdoors	Diners	Bus Stops	Lounges	Skate Parks	Bed & Breakfasts	Fish & Chips Shops

# Results

Cluster 4: Cluster 4 is one city called Arnhem. This city is different to other cities. This city have nive neighborhoods, sportfields, campgrounds and great outdoor venues

	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
2	Arnhem	52.005664	5.876235	3	Hockey Fields	Other Great Outdoors	Neighborhoods	Fields	Campgrounds	Bus Stops	Scenic Lookouts	Cities	Yoga Studios	Fish & Chips Shops

Cluster 5: Cluster 5 is one city called Breda. This city is also different to other cities in the Netherlands. The most common venues in this city are Restaurants, Pools, Bars and nice Coffee shops

	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
6	Breda	51.564448	4.75123	4	Restaurants	Bars	Pools	Coffee Shops	Dog Runs	Other Great Outdoors	Steakhouses	French Restaurants	Breakfast Spots	Hotels

# Discussion and Conclusion

---

In this notebook, a analysis of the different cities in the Netherlands is performed. The results may help visitors to choose a city based on their own interests. Using the Foursquare API, a good amount of venue recommendations is collected of 20 cities in the Netherlands. Sourcing from the venue recommendations from FourSquare for this assignment has its limitation; (1) limit of search distance of 500 meters and (2) there is a querying limit of 50 calls a day for venue ratings. So, because of the distance limit of 500 meters, the list of venues is not the complete list of all the available venues in the cities. Also because of the querying limit of 50 calls per day for venue ratings, a fictive list of score is calculated using the random module of Python. So the results of this assignment may significantly different, when more information are collected when changing the distance limit or using the real venue ratings.

# Discussion and Conclusion

---

The primary goal of this project is:

- 1. To compare the 20 cities and determine how similar or dissimilar they are using K-means clustering**

The result of the K-means clustering is that there is some similarity between the cities in the Netherlands. There are 5 clusters of cities, with one large clusters with mixed venue types. The common venues of this large cluster are bars, hotels and restaurants, which a visitor can find in each city. But there are some cities with different venues, which may be interesting to go to. The 8 cities with special venues are Gouda, Leiden, Delft, Volendam, Nijmegen, Den Bosch, Arnhem and Breda. The most common venues in those cities are outdoor venues, boats and ferries, lakes, pools and playgrounds.

# Discussion and Conclusion

---

The secondary goals are to analyze the diversity of the city venues:

## **2. What are the most common venues types (venues count)**

A visualization is created with the most common venue types of the cities in the Netherlands. The most common venues in the Dutch cities are bars, restaurants and Bus Stops. A visitor can find these venues in every city.

## **3. What are the recommending scores of the top 10 venues based on 'ratings'**

If a visitor want to know what are the venue categories with the highest recommending scores (based on rating scores) a top10 list is created with the categories with the highest mean ratings. The best rated venues are in the category "Gym pools", "Greek Restaurants" and "BBW Joints". This has is limitation, because the ratings score have a large variation, from low to high. Not all rating scores of food companies are good.



# Discussion and Conclusion

---

The secondary goals are to analyze the diversity of the city venues:

## **4. Which 5 cities has the highest mean recommending scores based on all venue 'ratings'**

The top 5 cities with the highest mean recommending scores are Arnhem', 'Rotterdam', 'Delft', 'The Hague', 'Leiden'. As a bonus a script is written to explorer the best food companies in a city.