Capstone project IBM

Crime rates and venues in Maastricht

Roel de Beer

1. Introduction

Maastricht is a town in southern part of the Netherlands. Its has gone through an immense transition the last twenty years as different international companies have landed in the **capital city in the southern province of Limburg**. The number of expats has increased exponentially, it now serves as a hub for expats.

1.1 Problem

The choice of location of settling down can be difficult one, but safety should be a key decision maker. This project will focus to select the safest neighborhood in Maastricht based on the total amount of crimes in Maastricht, explore the top 10 common venues and plot these results using the foursquare API.

2. Data acquisition

2.1 Neighborhoods GEOdata

The geolocatopm data of the neighbourhoods in Maastricht will be fetched from the API of Nominatim. Nominatim uses OpenStreetMap data to find locations on Earth by name and address (geocoding). It can also do the reverse, find an address for any location on the planet.

2.2 Crime rates

The data of crime rates in Maastricht are publicly available through the website of CBS (Central Bureau of Statistics in the Netherlands) and can be downloaded in csv format.

The crime data consists of the following columns

- 1. Naam of Neigborhood
- 2. Name of district / neighborhood
- 3. Code
- 4. Number of residents
- 5. Different kinds of crimes and their numbers (in dutch)
 - Vermogen, vernieling en geweld, absoluut
 - Totaal vermogen, vernieling en geweld
 - Vermogensmisdrijven
 - Totaal vermogensmisdrijven
 - Diefstal
 - Totaal diefstal
 - Fietsendiefstal
 - Diefstal overige vervoermiddelen
 - Diefstal uit/vanaf vervoermiddelen
 - Zakkenrollerij, straatroof en beroving
 - Totaal diefstal uit woning/schuur/e.d.
 - Diefstal uit niet -residentiele gebouwen
 - Overige diefstal/inbraak
 - Overige vermogensmisdrijven
 - Vernieling, misdrijf tegen openbare orde
 - Totaal vernieling, tegen openbare orde
 - Vernieling
 - Totaal vernieling
 - Vernieling aan auto

- Overige vernieling
- Overige vernieling, tegen openbare orde
- Gewelds- en seksuele misdrijven
- Totaal gewelds- en seksuele misdrijven
- Mishandeling
- Bedreiging en stalking
- Overige gewelds- en seksuele misdrijven
- · Vermogen, vernieling en geweld, relatief
- Totaal vermogen, vernieling en geweld
- Vermogensmisdrijven
- Totaal vermogensmisdrijven
- Totaal diefstal uit woning/schuur/e.d.
- Vernieling, misdrijf tegen openbare orde
- Gewelds- en seksuele misdrijven

https://opendata.cbs.nl/statline/#/CBS/nl/dataset/84468NED/table?ts=1593431131390

2.3 Venue data

The venue data is retrieved by passing in the required parameters to the FourSquare API, and creating a DataFrame with all the venue details along with the relevant neighbourhoods.

https://developer.foursquare.com

3. Data cleaning

The data consists of multiple columns that are redundant in for this assignment. We will only use the most recent data (2018) and focus on the total amounts of subcategories of crimes. The rows of the data have a hierarchy, City, District and neighborhood. The city consists of districts and the districts consists of neighborhoods. Since it is the battle of the neighborhoods, we will focus on the differences on neighborhood level and therefore remove the districts rows. Next to that, we removed several columns that are not important for our analyses. In order to make the data more relatable we created main subcategories of crime; Property theft, Vandalism and Violence & Sex crimes. Next to that, we translated everything from Dutch to English.

Below are all the neighborhoods in Maastricht, the residents, total crime and their crime rates per subsection.

Neighbourhoods	Region	Residents	Total_crimes	Property_theft	Vandalism	Violence_sex_crimes
		2070.0	205.0	742.0		
Binnenstad	Buurt	2070.0	896.0	743.0	63.0	90.0
Jekerkwartier	Buurt	1615.0	327.0	297.0	19.0	11.0
Kommelkwartier	Buurt	2360.0	108.0	80.0	19.0	9.0
Statenkwartier	Buurt	3925.0	286.0	188.0	47.0	51.0
Boschstraatkwartier	Buurt	1765.0	243.0	198.0	19.0	26.0
Sint Maartenspoort	Buurt	1700.0	175.0	136.0	21.0	18.0
Wyck	Buurt	5310.0	574.0	484.0	42.0	48.0
Villapark	Buurt	3280.0	82.0	69.0	9.0	4.0
Jekerdal	Buurt	1235.0	65.0	33.0	17.0	15.0
Biesland	Buurt	1665.0	73.0	62.0	7.0	4.0
Campagne	Buurt	1270.0	13.0	9.0	3.0	1.0
Wolder	Buurt	1415.0	38.0	27.0	7.0	4.0
Sint Pieter	Buurt	150.0	40.0	38.0	1.0	1.0
Brusselsepoort	Buurt	6665.0	270.0	210.0	32.0	28.0
Mariaberg	Buurt	4880.0	184.0	127.0	30.0	27.0
Belfort	Buurt	2620.0	45.0	37.0	5.0	3.0
Pottenberg	Buurt	2475.0	91.0	31.0	26.0	34.0
Malpertuis	Buurt	2180.0	89.0	39.0	31.0	19.
Caberg	Buurt	3385.0	165.0	107.0	28.0	30.0
Oud-Caberg	Buurt	1900.0	43.0	38.0	4.0	1.0
Malberg	Buurt	5305.0	181.0	128.0	30.0	23.0
ousberg-Hazendans	Buurt	1995.0	56.0	49.0	5.0	2.0
Daalhof	Buurt	6665.0	137.0	97.0	21.0	19.0
Boschpoort	Buurt	1620.0	51.0	39.0	4.0	8.0
Bosscherveld	I Buurt	45.0	47.0	42.0	4.0	1.0
Frontenkwartier	r Buurt	t 420.0	101.0	93.0	7.0	1.0
Belvédère	Buurt	t 25.0	14.0	11.0	2.0	1.0
Lanakerveld	I Buurt	t 70.0	2.0	2.0	0.0	0.0
Wyckerpoort	t Buurt	t 3855.0	149.0	109.0	32.0	8.1
Heugemerveld	I Buurt	t 2905.0	81.0	50.0	13.0	18.0
Wittevrouwenveld	I Buurt	5465.0	447.0	327.0	73.0	47.
Nazareth	Buurt	t 3250.0	240.0	166.0	43.0	31.0
Limmel	I Buurt	2070.0	103.0	73.0	16.0	14.
Scharn	Buurt	t 6510.0	256.0	193.0	28.0	35.0
Amby	, Buurt	6585.0	159.0	105.0	30.0	24.
Beatrixhaven	Buurt	t 20.0	58.0	53.0	4.0	1.
Borgharen	Buurt	1720.0	29.0	19.0	5.0	5.
Itteren					2.0	2.
Meerssenhoven					0.0	1.
Randwyck					21.0	13.
Heugem					13.0	21.
Heer					47.0	21.
De Heeg					32.0	17.
Vroendaal					4.0	0.0
Viociidadi	buult	1075.0	20.0	24.0	4.0	0.0

Figure 1.1 Crimes in different neighborhoods in Maastricht after cleaning

To find the longtitude and latitude of the neighborhoods we used Nominatim to find locations on Earth by name and address (geocoding).

Neighborhood	Latitude2	Longitude2
Binnenstad	50.849271	5.688756
Jekerkwartier	50.844920	5.690547
Kommelkwartier	50.847243	5.682764
Statenkwartier	50.852338	5.686977
Boschstraatkwartier	50.854042	5.691918
Sint Maartenspoort	50.852845	5.700712
Wyck	50.847290	5.702527
Villapark	50.838870	5.692645
Jekerdal	50.838847	5.684309
Campagne	50.836978	5.663174
Wolder	50.834464	5.654380
Sint Pieter	50.831634	5.693637
Brusselsepoort	50.851215	5.669519
Mariaberg	50.844138	5.671711
	50.846293	5.660309
Pottenberg	50.852655	5.657684
Malpertuis	50.858362	5.661858
_	50.859404	5.671216
Oud-Caberg		5.663661
-	50.861357	5.655768
Dousberg-Hazendans	50.849765	5.646829
	50.841316	5.653703
Boschpoort		5.690203
Bosscherveld		5.683250
Frontenkwartier	50.854165	5.682527

Figure 1.2 Neighborhood latitude and longtitude fetched by Nominatim

4. Methodology

4.1 Statistical details & explanatory data analyses

Pandas describe() is used to view some basic statistical details like percentile, mean, std etc. of a data frame or a series of numeric values. When this method is applied to a series of string, it returns a different output which is shown in the example below.

The count for each of the major categories of crime returns the value 44 which is the number of neighborhoods in Maastricht. Property Theft is the highest reported crime during the year 2018 followed by 'Violence & sex crimes. The lowest recorded crime category is Vandalism.

	Residents	Total_crimes	Property_theft	Vandalism	Violence_sex_crimes
count	44.000000	44.000000	44.000000	44.000000	44.000000
mean	2787.045455	154.340909	117.909091	19.681818	16.750000
std	2142.021790	166.578698	137.400250	17.355078	18.020821
min	20.000000	2.000000	2.000000	0.000000	0.000000
25%	1378.750000	46.500000	37.750000	4.750000	2.000000
50%	2070.000000	102.000000	76.500000	18.000000	13.500000
75%	4086.250000	200.250000	145.000000	30.000000	24.500000
max	7535.000000	896.000000	743.000000	73.000000	90.000000

Figure 1.3 statistical analysis of the dataset

High crime neighborhoods

The neighborhood "Binnenstad" (City Center) is the neighborhood of the highest crime rate, followed by the neighborhoods Wyck, Wittevrouwenveld, Jekerkwartier and Statenkwartier. The highest crime category in the 'Binnenstad' neighborhood is Property theft, followed by Violence & sex crimes and Vandalism which is similar to the overall distribution.

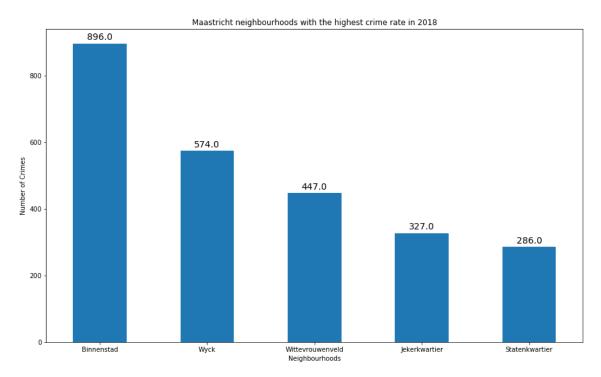


Figure 1.4 high crime neigborhoods in Maastricht in 2018

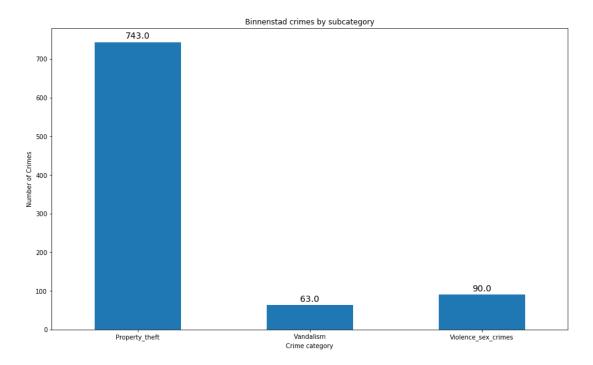


Figure 1.5 Distribution of crime in the neighborhood "Binnenstad", with the highest crime rate in Maastricht in 2018

Neighborhoods with the lowest crime rates

The neighborhood with the lowest crime rate is Lanakerveld, with only two crimes during the year 2018 (Both property theft), after that Meerssenhoven, Itteren, Campagne and Belvedere.

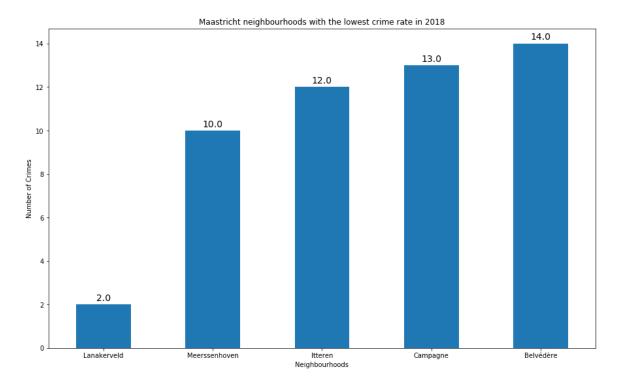


Figure 1.6 low crime neigborhoods in Maastricht in 2018

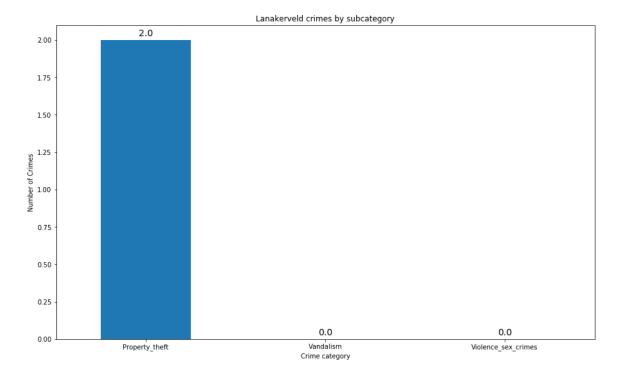


Figure 1.7 Distribution of crime in the neighborhood "Lanakerveld", with the lowest crime rate in Maastricht in 2018

Safest Neighborhoods in Maastricht

There are the five safest neighborhoods in Maastricht. Folium was used to visualize the neighborhoods.

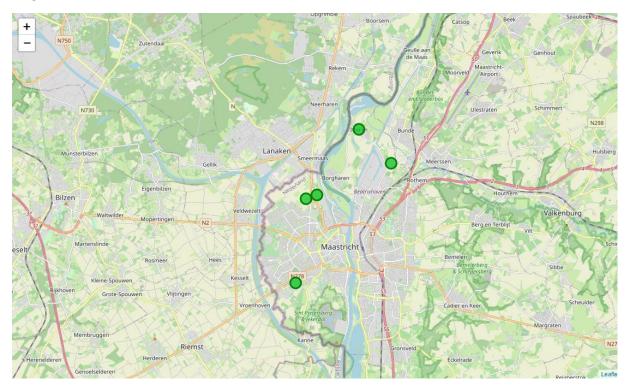


Figure 1.8 Safest neigborhoods plotted on map using Folium

5. Modelling

5.1 Using the dataset & Foursquare

We used the dataset with the longtitude and latitude of the safest neighborhoods in Maastricht to fetch the venues (within 500m) through the foursquare API and convert it to a dataframe shown below.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Lanakerveld	50.870430	5.669590	De Poshoof	50.866825	5.671815	Bed & Breakfast
- 1	Lanakerveld	50.870430	5.669590	Cour 8 Lofts	50.866442	5.671599	Bed & Breakfast
2	Meerssenhoven	50.884368	5.723069	Janssen Pluimvee en Poeliersbedrijf	50.883404	5.722990	Farm
3	Meerssenhoven	50.884368	5.723069	Attero BV	50.882806	5.720165	Business Service
4	Itteren	50.897733	5.702930	BrigidaHoeve	50.896401	5.704809	Ice Cream Shop
5	Itteren	50.897733	5.702930	Frietkar 't Weegelke Itteren	50.896604	5.704478	Friterie
6	Itteren	50.897733	5.702930	Brigidahoeve	50.897276	5.703777	Restaurant
7	Itteren	50.897733	5.702930	Wandelroute Itteren	50.896498	5.704726	Trail
8	Itteren	50.897733	5.702930	Bours Machineverhuur	50.896805	5.696287	Construction & Landscaping
9	Campagne	50.836978	5.663174	Bloemen en Planten van Den Boorn	50.838596	5.663030	Flower Shop
10	Campagne	50.836978	5.663174	Harmoniezaal "De Greune"	50.837249	5.659434	Concert Hall
11	Campagne	50.836978	5.663174	La Mulata	50.837030	5.657486	Salsa Club
12	Campagne	50.836978	5.663174	The Kasteel	50.837836	5.657560	Hotel Bar
13	Belvedere	50.871837	5.676515	Camperplaats Maastricht	50.874138	5.681763	Campground

Figure 1.9 Top5 Safest neigborhoods venues

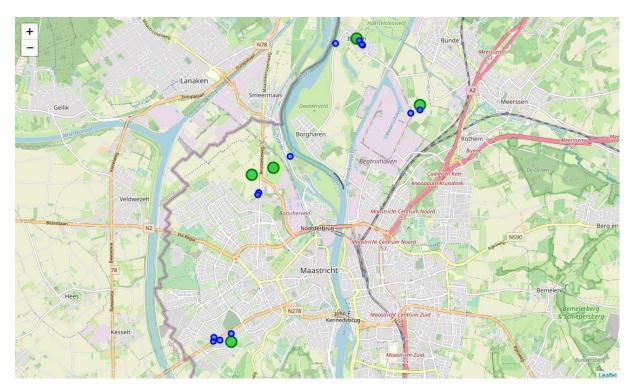


Figure 2.0 Safest neigborhoods with venues plotted on map using Folium

As we can see, there are not many venues in the safest neighborhoods of Maastricht. It is the safest, but not very appealing to the expats since almost no venues are nearby. Comparing this to the total

amount of venues in Maastricht there are 13 venues in the safest neighborhoods compared to a total of 472 in entire city of Maastricht.

:	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Binnenstad	50.849271	5.688756	Boekhandel Dominicanen	50.850129	5.689706	Bookstore
1	Binnenstad	50.849271	5.688756	Vrijthof	50.849295	5.689001	Plaza
2	Binnenstad	50.849271	5.688756	II Bacaro	50.848495	5.689203	Italian Restaurant
3	Binnenstad	50.849271	5.688756	La Bodega Tapas	50.848803	5.690020	Tapas Restaurant
4	Binnenstad	50.849271	5.688756	Theater aan het Vrijthof	50.850147	5.688010	Theater
		***		***	***	***	***
467	Biesland	50.839914	5.672686	Saint Servais	50.837345	5.671984	French Restaurant
468	Biesland	50.839914	5.672686	il fiore Mariaberg	50.844178	5.671180	Gym
469	Biesland	50.839914	5.672686	Annelies' Place to B & B	50.840440	5.675798	Bed & Breakfast
470	Biesland	50.839914	5.672686	Prix De Rome	50.837430	5.671901	Café
471	Biesland	50.839914	5.672686	Café Backstage	50.842956	5.673934	Pub

472 rows × 7 columns

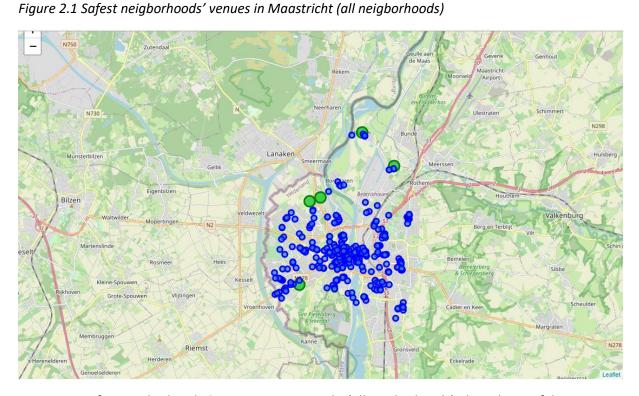


Figure 2.1 Safest neigborhoods & venues in Maastricht (all neigborhoods) plotted using folium

The figure above shows that most venues are outside the safest neighborhoods. So, the decision was made to broaden the terms of safest in Maastricht. The neighborhoods were selected that were below the mean of 154 total crime incidents.

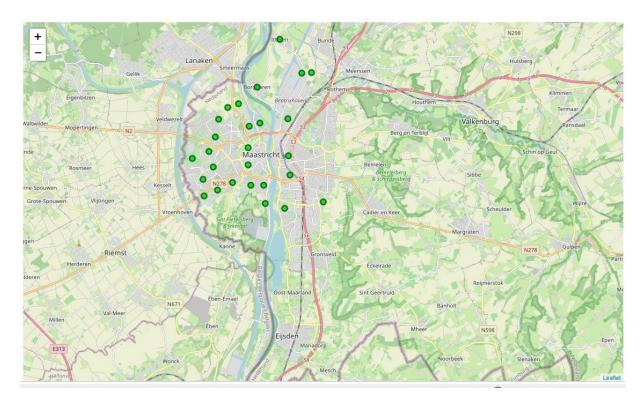


Figure 2.2 Neighborhoods (27) with a crime rate of less than 154 crime incidents plotted using folium

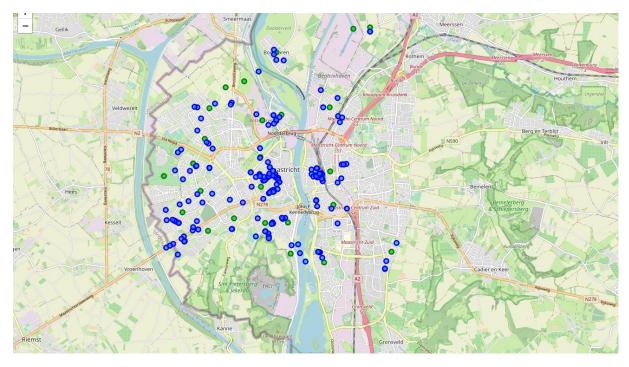


Figure 2.3 Neighborhoods (green) with a crime rate of less than 154 crime incidents and their venues (blue)

5.2 K-Means Clustering

On the venues data we used One hot encoding. One-hot encoding is essentially the representation of categorical variables as binary vectors. These categorical values are first mapped to integer values. Each integer value is then represented as a binary vector that is all 0s (except the index of the integer which is marked as 1).

The mean of the venues is calculated, and the venues will be grouped by the neighborhood and the ten venues most common are calculated. We will be using K-means clustering to find the neighborhoods with the same characteristics. A cluster of 5 is used for the 33 neighborhoods.

6. Results

We are now going to see which neighborhoods are clustered together.

First Cluster

The first cluster has only one neighborhood: 'Belvedere' which has venues like Campgrounds, cave excursions, Chinese restaurant, concert hall and some stores.



Figure 2.4 first cluster

Cluster 2

The second cluster has two neighborhoods: Wolder and Dousberg-Hazendans which has venues like soccer field, bars. Fast food restaurants and a convenience store.

	Neighborhood	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
5	Wolder	50.834464	5.654380	1	Border Crossing	Friterie	Bar	Middle Eastern Restaurant	Plaza	Soccer Field	Fast Food Restaurant	Gas Station	Farm	Country Dance Club
11	Dousberg- Hazendans	50.849765	5.646829	1	Soccer Field	Food Stand	Women's Store	Chinese Restaurant	Clothing Store	Coffee Shop	Concert Hall	Construction & Landscaping	Convenience Store	Cosmetics Shop

Figure 2.5 second cluster

Cluster 3

The third cluster has one neighborhood which has some stores, a coffee shop and a dance club.

	Neighborhood	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
17	Lanakerveld	50.87043	5.66959	2	Bed &	Women's Store	Department	Clothing Store	Coffee Shop	Concert Hall	Construction &	Convenience	Cosmetics	Country Dance

Figure 2.6 third cluster

Cluster 4

The cluster four consists of 27 neighborhoods and is the largest cluster. This is the most vibrant cluster of Maastricht which has all sorts of restaurants, bars, sporting activities, shopping malls and stores.



Figure 2.7 fourth cluster

Cluster 5

The fifth cluster has two neighborhoods which is considered a farming area with some stores and restaurants / deli's.



Figure 2.7 fifth cluster

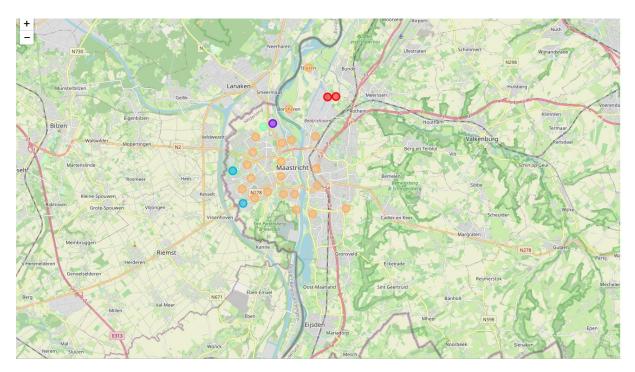


Figure 2.8 Clusters visualized: First cluster: Purple, Second cluster: Blue, Third cluster: Light Green, Fourth cluster: Orange, fifth cluster: Red

7. Discussion

The type of neighborhood of choice is heavily dependent on the preferences of the person; Cluster one's Belvedere was in the top 5 of the safest neigborhoods, outdoor activities and has limited venues, but you could consider it as a top choice if these venues are appealing your personal preferences, Cluster two has very limited venues but if you are into soccer this is a good place, cluster three is the safest area of Maastricht, but not so appealing venue-wise. The fourth cluster is has the most facilities, but the crime rate is higher then the other clusters. If you like quiet neighborhoods of Maastricht and do not care about the amount of venues, cluster five is your choice.

8. Conclusion

With this analysis, expats can make a better choice of location based on crime rates and the amount and type of venues. The next step of giving an even better analysis is to plot the available houses based in price range. This will be incorporated in the second version of the report at a later stage.