# Investigation on Population Migration within Cologne

## 1 Introduction

Cologne is the 4th biggest city in Germany. It is known for it's very individual culture and people. The city seems to be very heterogeneous in different aspects.
In recent years, the costs of living and rents have increased by a lot all over Germany, especially in bigger cities. This is of course also true for Cologne.

As a consequence, people are forced to move in these days. In this report, we want to investigate this migration.

### 1.1 Problem/Question

Moving is always a big decision which you will only do after a lot of thinking, because it's a big step personally, socially and also financially. We want to understand what makes people leaving their place. We want to find out if there are groups of people who are more probable to move than others.

Also, we want to identify growing and shrinking districts within Cologne. For this we take as a measure the net migration *within* Cologne. People from outside moving to Cologne are neglected in this analysis. After analyzing the population in those districts, we want to attempt to understand why people are coming/leaving.

### 1.2 Data

The city of Cologne has a portal for open data which is located at https://www.offenedaten-koeln.de.

For this analysis, I will use some of the data files provided there. Most of the files are given in geojson format. This means, that data is given in an array of features. Each item in the feature list corresponds to a district in Cologne and has attributes and geometries. The attributes dictionary contains the data which depends on the file. The geometry dictionary contains a list of points which correspond to the edges of the wrapping polygon of that district on a map.

I will use data for the following data files:

| Description | Filename | Source | |
|---|---|---|---|
| **Districts Location** | bezirke.json | https://www.offenedaten-koeln.de/dataset/stadtbezirke | Used for extracting the center of each district |
| **Inhabitants** | einwohner.json | https://www.offenedaten-koeln.de/dataset/einwohner-statistik-koeln | Used for finding growing or shrinking districts |
| **Area Usage** | flaeche.json | https://offenedaten-koeln.de/dataset/flaechennutzung-koeln-anzahl | Used for finding area usages in districts like water/park/… |

| | | | |
|---|---|---|---|
| **Age Distribution** | altersgruppen.csv | https://offenedaten-koeln.de/dataset/einwohner-nach-altersgruppen | Distribution of age distribution |
| **Household distribution** | haushalte.json | https://offenedaten-koeln.de/dataset/haushaltsgr%C3%B6%C3%9Fe | Distribution of household size |

I will also use the Foursquare API to retrieve venues close to the districts' center points.

# 2 Methodology

The analysis will be done in several steps. The final target is to find districts that are growing/shrinking and to connect that to the districts' inhabitants' characteristics. Most of the effort will be used to categorize these characteristics.

We will use the following aspects to characterize districts:

- Most Popular Venue Categories
- Area Usage
- Age Distribution among Inhabitants
- Household Size Distribution

For all of these aspects, a cluster analysis will be done in order to find similar districts.

In each of the next four subchapters, one of the bullet points from above item list will be investigated.

## 2.1 Venue Categories

Let's start with the first aspect: Venue Categories. We want to perform a similar analysis like we did in the previous capstone project assignments for NY and Toronto. Therefore, we need a data set containing Cologne's districts with center points around which we can search for venues by using Foursquare API calls.

We use the file bezirke.json which is a geojson file with the following attributes per district:

```
dict_keys(['OBJECTID', 'NUMMER', 'NAME', 'NR_STADTBEZIRK', 'STADTBEZIRK', 'FLAECHE', 'LINK'])
```

After a thorough data preparation (see Jupyter Notebook for details) we end up with a data frame containing all the districts and their corresponding center points:

| | District | Latitude | Longitude | Polygon |
|---|---|---|---|---|
| 0 | Godorf | 50.852610 | 6.982218 | [[6.994359341598065, 50.85835989408827], [6.99... |
| 1 | Lövenich | 50.948533 | 6.823829 | [[6.835101297207516, 50.957260493877406], [6.8... |
| 2 | Weiden | 50.934514 | 6.824246 | [[6.849501526802686, 50.94220412958467], [6.84... |
| 3 | Junkersdorf | 50.923891 | 6.859688 | [[6.854198181651118, 50.94052346246431], [6.85... |
| 4 | Widdersdorf | 50.967660 | 6.841605 | [[6.851763526592514, 50.97718504868292], [6.85... |

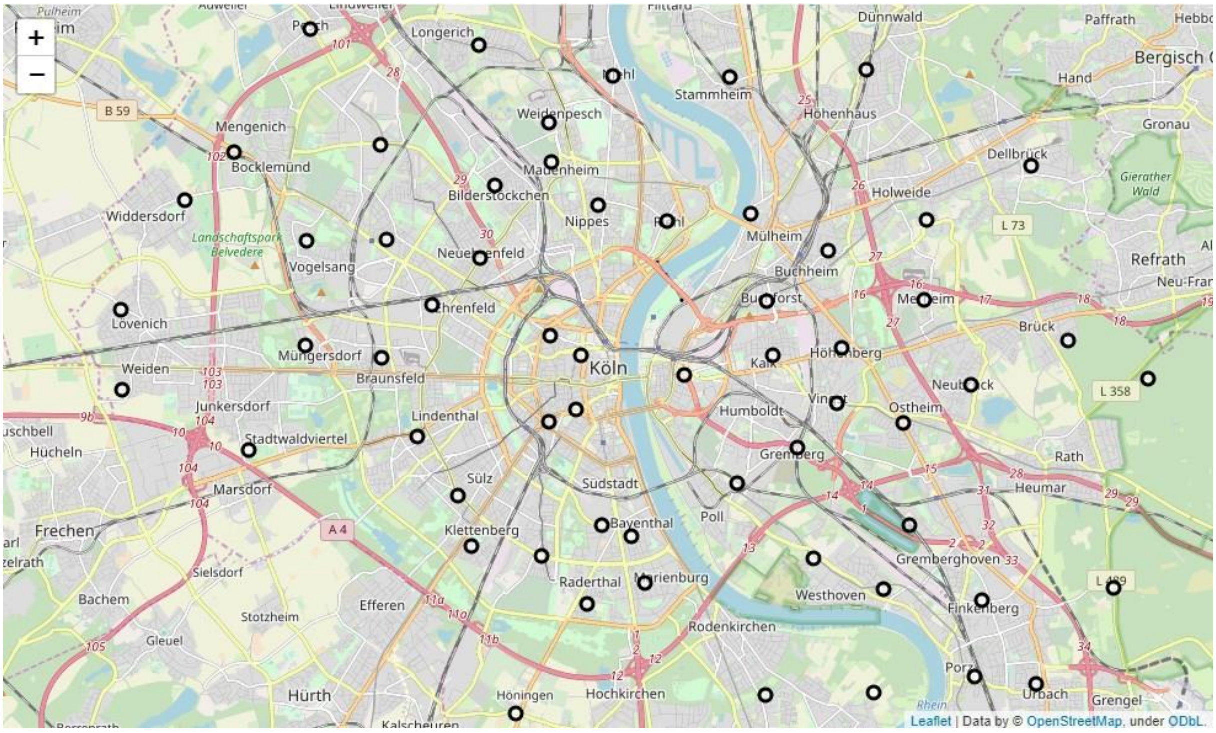



*Figure 1 Cologne Map with Districts' Center Points*

After having extracted all districts' center points, we can assemble requests to the Foursquare API. We limit the number of venues per call to 100 and search in a radius of 500m around each district's location.

As a result, we get a data frame of all retrieved venues in the district's search radius:

| | Neighborhood | Neighborhood Latitude | Neighborhood Longitude | Venue | Venue Latitude | Venue Longitude | Venue Category |
|---|---|---|---|---|---|---|---|
| 0 | Godorf | 50.852610 | 6.982218 | Godorf Hafen | 50.848477 | 6.982683 | Harbor / Marina |
| 1 | Godorf | 50.852610 | 6.982218 | Bude Pausenraum | 50.852996 | 6.988761 | Breakfast Spot |
| 2 | Godorf | 50.852610 | 6.982218 | Godorfer Imbiss | 50.852567 | 6.975184 | Snack Place |
| 3 | Lövenich | 50.948533 | 6.823829 | Bäckerei Kraus - Lövenicher Café | 50.946507 | 6.829737 | Bakery |
| 4 | Lövenich | 50.948533 | 6.823829 | Alte Schmiede | 50.946739 | 6.829642 | Gastropub |

Also, we can retransform the data such that we get the most frequent venue categories for each district (see Appendix: Most frequent venue categories per district).

In order to find similar districts, we will compare them by their ten most common venues. Therefore, a data frame is created that holds in each row a district and its ten most common venues as columns. For instance, the row for the district <Nippes> looks like this:

| | Neighborhood | 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 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 51 | Nippes | Café | Italian Restaurant | Bar | German Restaurant | Supermarket | Bakery | Plaza | Modern European Restaurant | Gym / Fitness Center | Greek Restaurant |

With this data frame, it is now possible to compare neighborhoods. For this we check for overlaps in the venue frequency of the districts and group districts with a similar venue environment.
As a method for this, the K-Means algorithm is used. This will try to find clusters of elements by means of the distance between them. Here, the distance is measured on basis of the venue categories. We perform the algorithm with 10 clusters on the ten most common venues. In turn, we

get a cluster label for each district. This can be color-coded to the district's location on the map (cf. Figure 2).

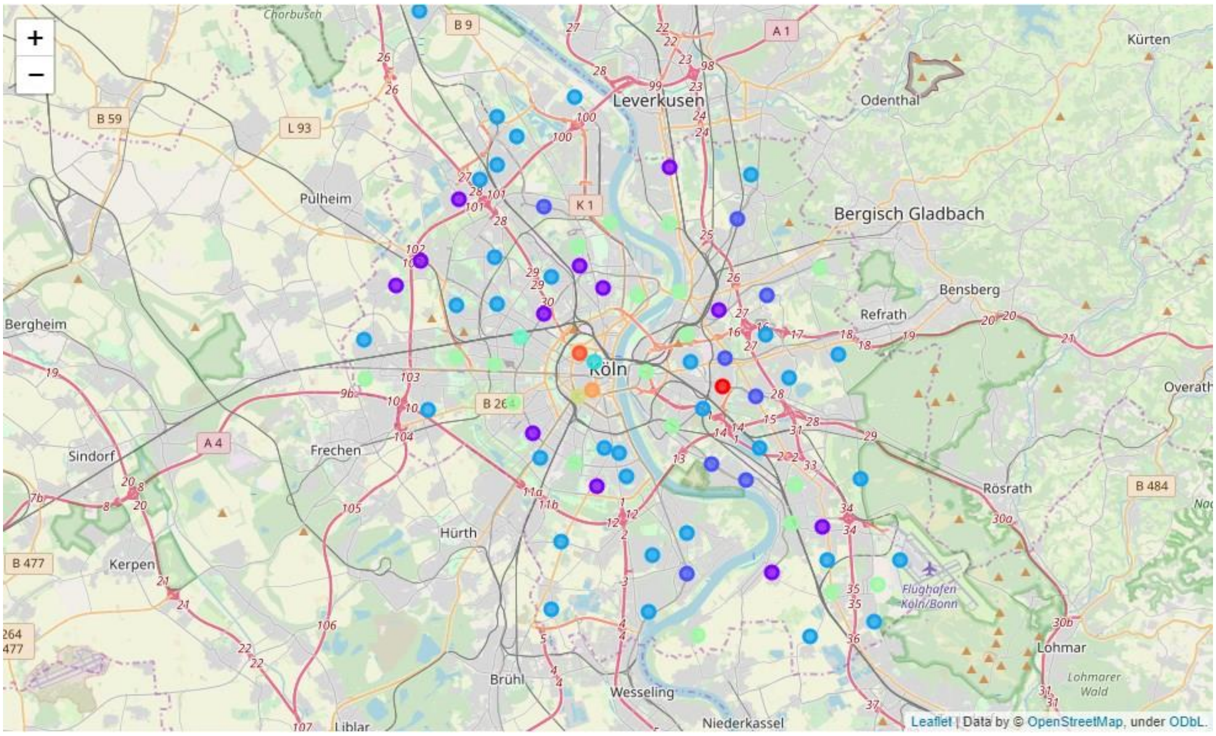



*Figure 2 Venue Clustering: The color code among the districts shows districts with similar venue environment.*

We can see the different clusters of the Cologne venue environment. One interpretation[1] of the result could be that in the city center the venue distribution is very heterogeneous. The clusters are not similar to any other cluster. Each district in the city center forms its own cluster and thus does not resemble any other district in the Cologne area. If we go to the outer rings, we see more similarities between districts. For instance, the districts close to but right of the Rhein river seem to be quite similar (cluster 6 → green). They also show similarities to clusters on the left side of the rivers, in districts where industry governed. If we look at cluster 3 (blue), we again see a lot of food and drink venues but now in form of restaurants and also shops. This could be an area with many households. Cluster 1 (violet) shows many activity locations.

This concludes the section about Cologne's venue environment analysis. There are some hints about whether or not people might want to migrate to other districts but for sure further analyses have to be performed.

## 2.2 Land Utilisation Among Districts

In this subchapter we look at the land use within each district. The provided data file lists several categories for land use. Here, four of those will be used:

- FN_BEBAUT_AP : Ratio of built-up area
- FN_PARK_AP : Ratio of park area
- FN_VERKEHR_AP : Ratio of area assigned to traffic
- FN_WASSER_AP : Area ratio of water

[1] The interpretation is based on looking at similarities in the categories of the venues of each cluster. The corresponding tables can be found in the Jupyter Notebook.

By writing a simple extraction script, a data frame can be extracted which assigns the above-mentioned categories for land use to each of the districts:

| | District | built-up | park | traffic | water |
|---|---|---|---|---|---|
| 0 | Mülheim | 52.1 | 10 | 24.6 | 10 |
| 1 | Braunsfeld | 68.5 | 1.2 | 26.8 | 0 |
| 2 | Ossendorf | 41.4 | 24.5 | 15.1 | 0 |
| 3 | Porz | 61.2 | 6.1 | 19.9 | 9 |
| 4 | Altstadt-Süd | 54.7 | 4.2 | 30.9 | 10.2 |

In this table, the values are given in percent.

Next, a cluster analysis is performed on this data set. It again assigns labels to each district considering similarities in the land utilization, i.e. districts with similar land use environments will fall into the same cluster. The result of the clustering can be found in Figure 3.

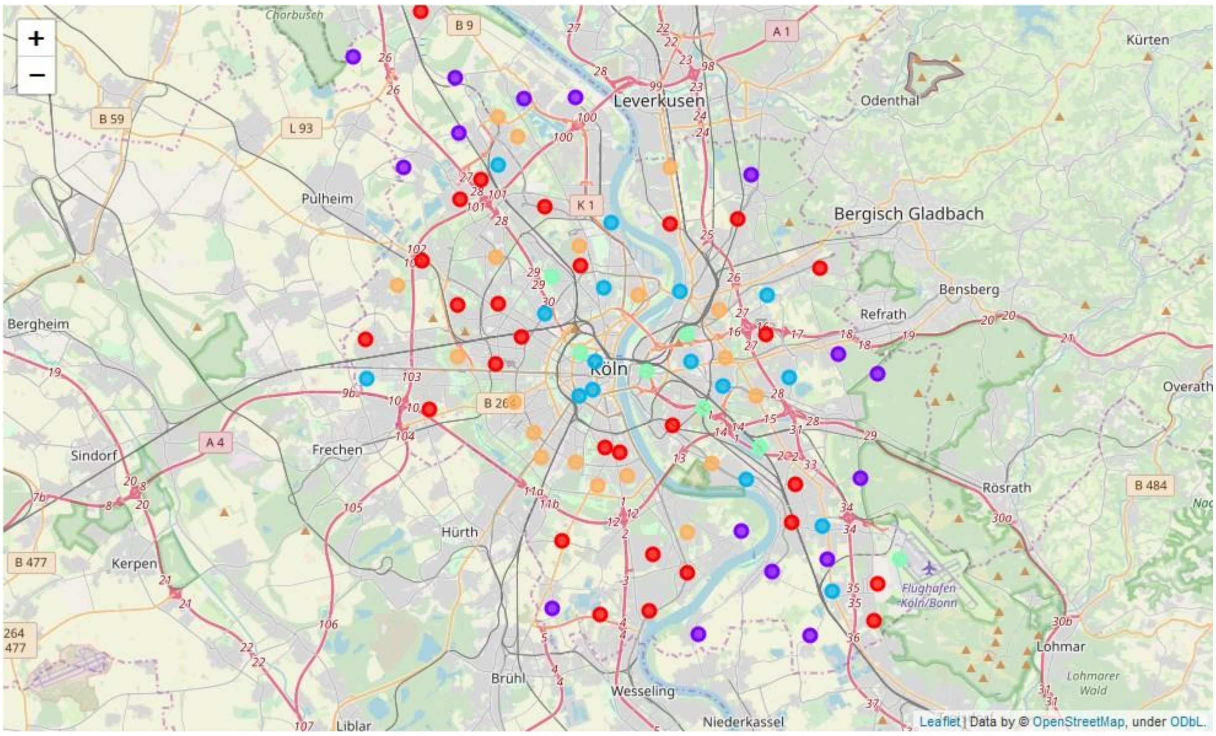
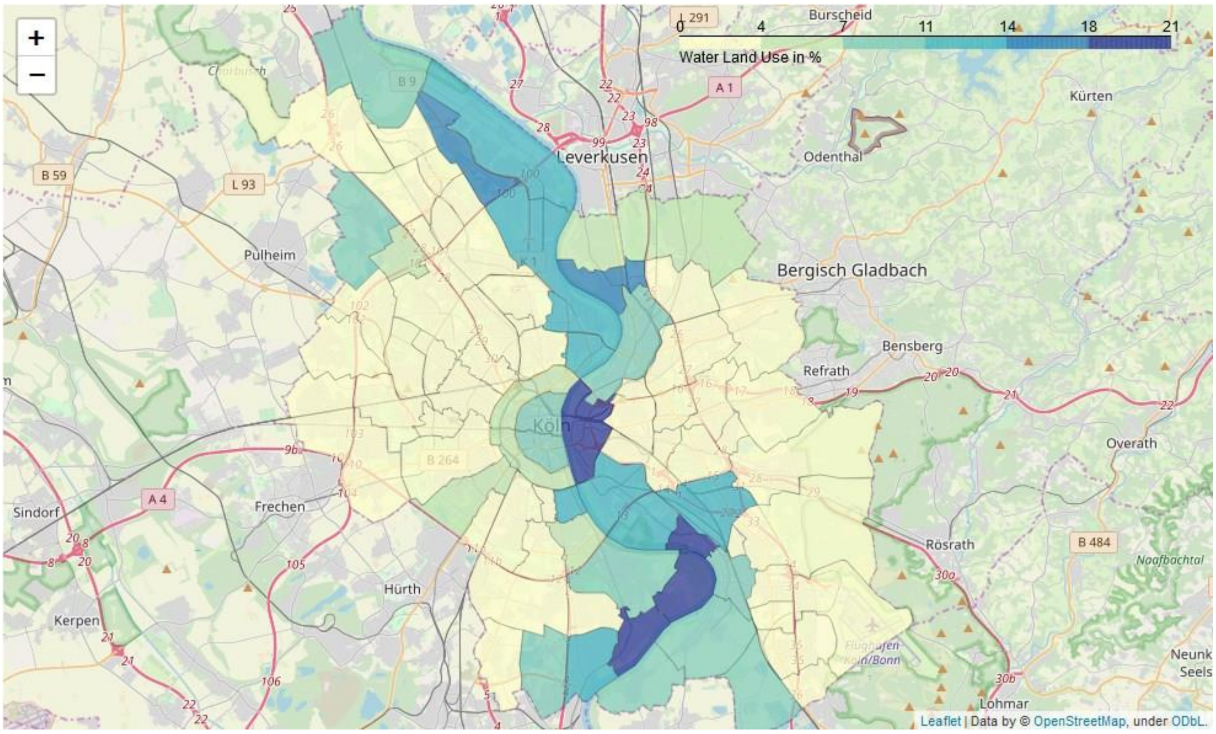
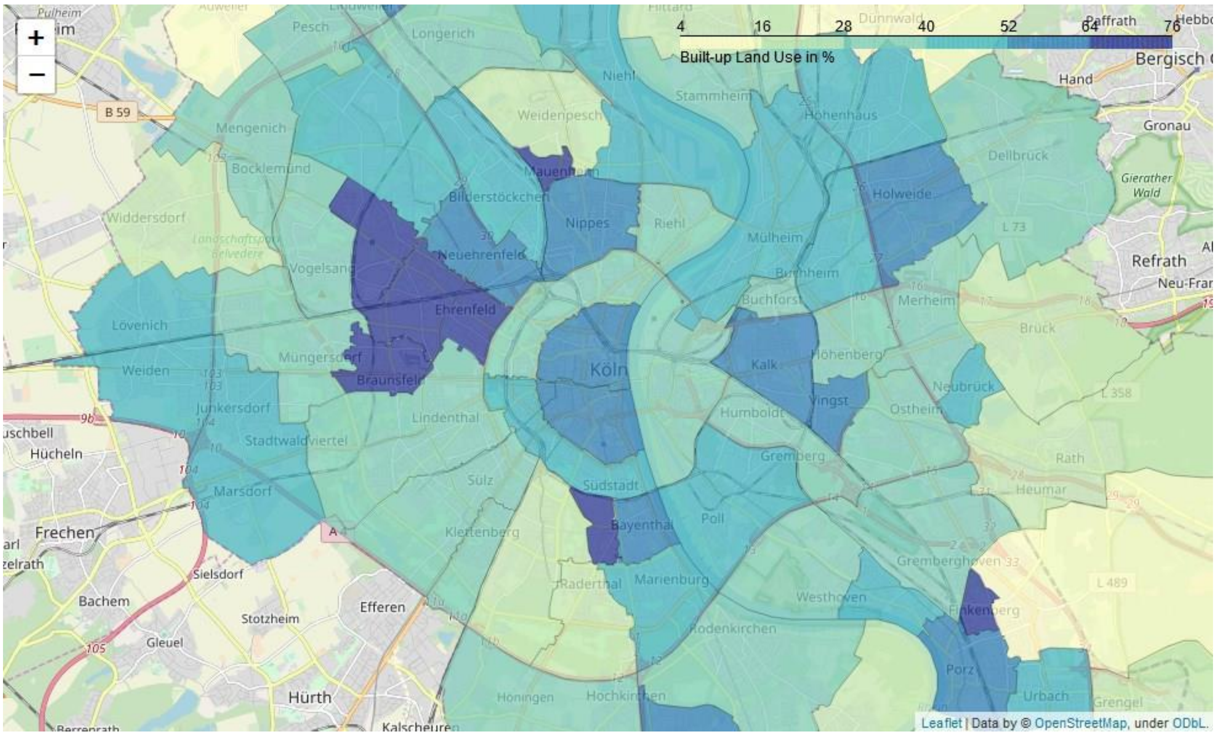



*Figure 3 Land Use Clustering: Districts with similar land use environments fall into the same cluster, i.e. they have the same color code.*

As expected we see the distribution of clusters matching the open street map's color code. Areas which are close to rivers belong to one cluster category and then the clusters are ring-shaped with their common center being the city center. This makes again sense, since built-up areas and traffic decreases radially. Let's try to use a Choropleth map to make this more apparent:




In this Choropleth map we see the water density in the certain districts. This semi-explains the distribution of clusters. Let's also look at the built-up density:




Here, we see how built-up areas are more dense towards the city center, but we also find some dense areas towards the south-east near the airport. This could - for instance - explain, why districts in that area fall into the same cluster label class as districts in the city center.

## 2.3 Household Size Distribution

In the following, we use the very same analysis as for the land utilization, but now for the distribution of households. The data has the same format, i.e. the same processes are used to extract the following columns:

- District: Name of district
- density: Density of households per ha
- ratio-<x>P: Ratio of people living in households of <x> persons
- ratio-5P+: Ratio of people living in households of 5 persons or above

The result is a data frame that shows for each district the household size distribution. That is the ratio of people living in a household of certain size.

| | District | density | ratio-1P | ratio-2P | ratio-3P | ratio-4P | ratio-5P+ |
|---|---|---|---|---|---|---|---|
| 0 | Stammheim | 10.4 | 40.7 | 32 | 14.2 | 9 | 4.2 |
| 1 | Mülheim | 31.8 | 54.3 | 24.4 | 10.3 | 7.2 | 3.8 |
| 2 | Braunsfeld | 40.6 | 56.0 | 26.5 | 9.5 | 6.1 | 1.9 |
| 3 | Ossendorf | 6.6 | 38.1 | 27.9 | 15.6 | 12.7 | 5.8 |
| 4 | Porz | 22.7 | 44.0 | 29.3 | 12.6 | 9.5 | 4.6 |

We cluster also this data set and retrieve a map showing the cluster distribution by household size in Figure 4.

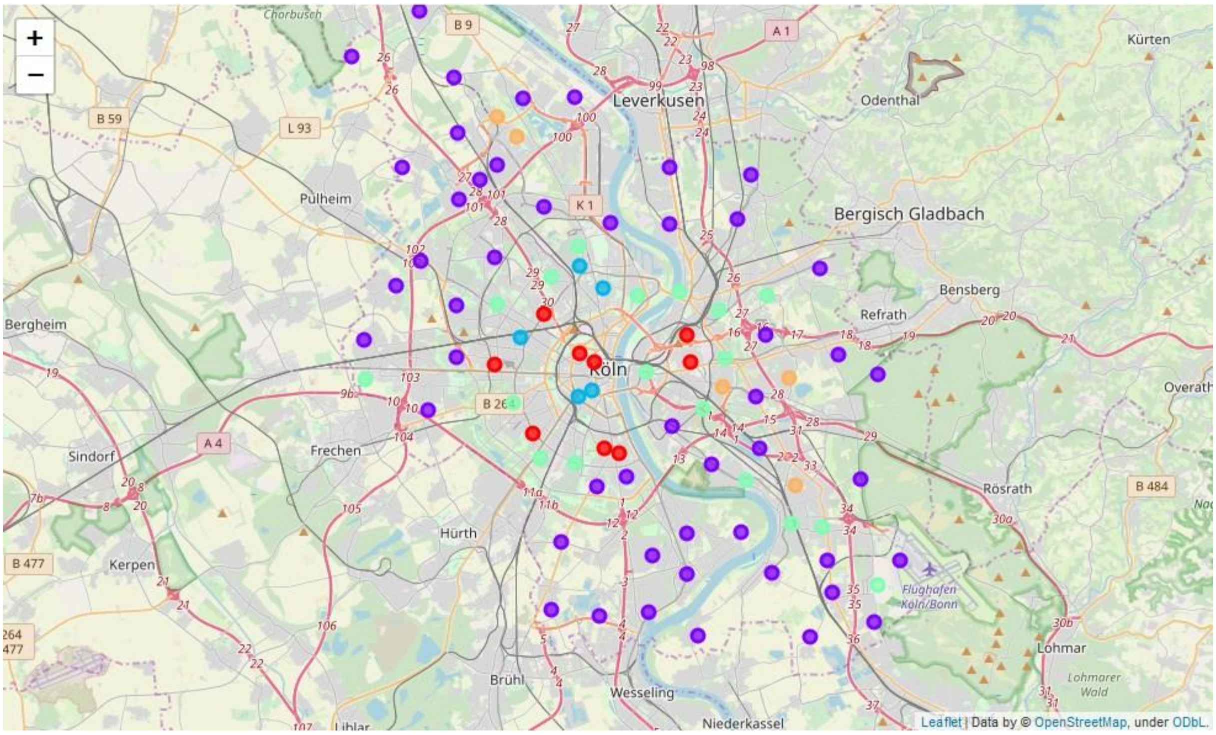



*Figure 4 Household Size Clustering: Districts with household size distribution will show the same color code.*

Here, it is very obvious, that there is one cluster which contains the outer areas of Cologne. The mean values of the household sizes in that cluster are as follows:

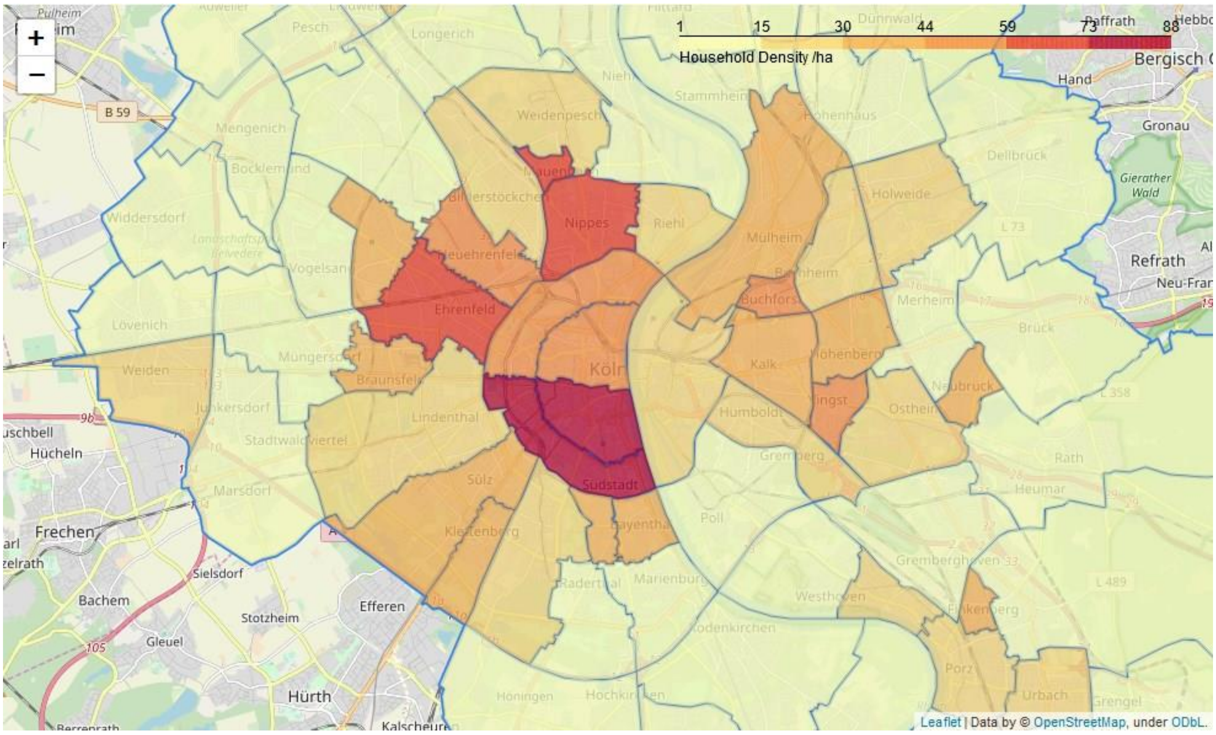
```
ratio-1P      38.726531
ratio-2P      31.510204
```

```
ratio-3P      14.404082
ratio-4P      10.777551
ratio-5P+      4.581633
```

It can be seen, that the household distribution is rather equal. This speaks for larger families living in these areas. Compared to households near the city center, the difference is apparent:

```
ratio-1P      61.86
ratio-2P      22.50
ratio-3P       8.32
ratio-4P       5.44
ratio-5P+      1.88
```

To confirm those interpretation, a Choropleth map of the household density per hectare is produced:




This result matches the expectation as districts which show higher amounts of single person households will also show a higher density of households.

## 2.4 Age Distribution

Lastly, we look at the age distribution in the districts. We have the data in `.csv` format and can easily scrape them and put them into a data frame:

| | District | 0-2 | 3-5 | 6-14 | 15-17 | 18-20 | 21-34 | 35-59 | 60-64 | 65-74 | 75-79 | 80+ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | Altstadt-Süd | 0.018752 | 0.015342 | 0.034855 | 0.013202 | 0.024011 | 0.334482 | 0.351565 | 0.044830 | 0.081245 | 0.038519 | 0.043198 |
| 1 | Neustadt-Süd | 0.026606 | 0.020046 | 0.043228 | 0.013251 | 0.026684 | 0.361220 | 0.363651 | 0.038864 | 0.058727 | 0.023287 | 0.024437 |
| 2 | Altstadt-Nord | 0.018887 | 0.012629 | 0.029392 | 0.009890 | 0.025089 | 0.338064 | 0.365724 | 0.040847 | 0.073201 | 0.033136 | 0.053140 |
| 3 | Neustadt-Nord | 0.027228 | 0.020315 | 0.041941 | 0.012905 | 0.019074 | 0.291073 | 0.405127 | 0.046302 | 0.077856 | 0.030206 | 0.027973 |
| 4 | Deutz | 0.024880 | 0.020128 | 0.046393 | 0.016366 | 0.028905 | 0.290899 | 0.364944 | 0.041444 | 0.080644 | 0.040586 | 0.044810 |

Here, we have for each district the relative amount of people of a certain range of ages. Based on that data, another cluster analysis can be performed:

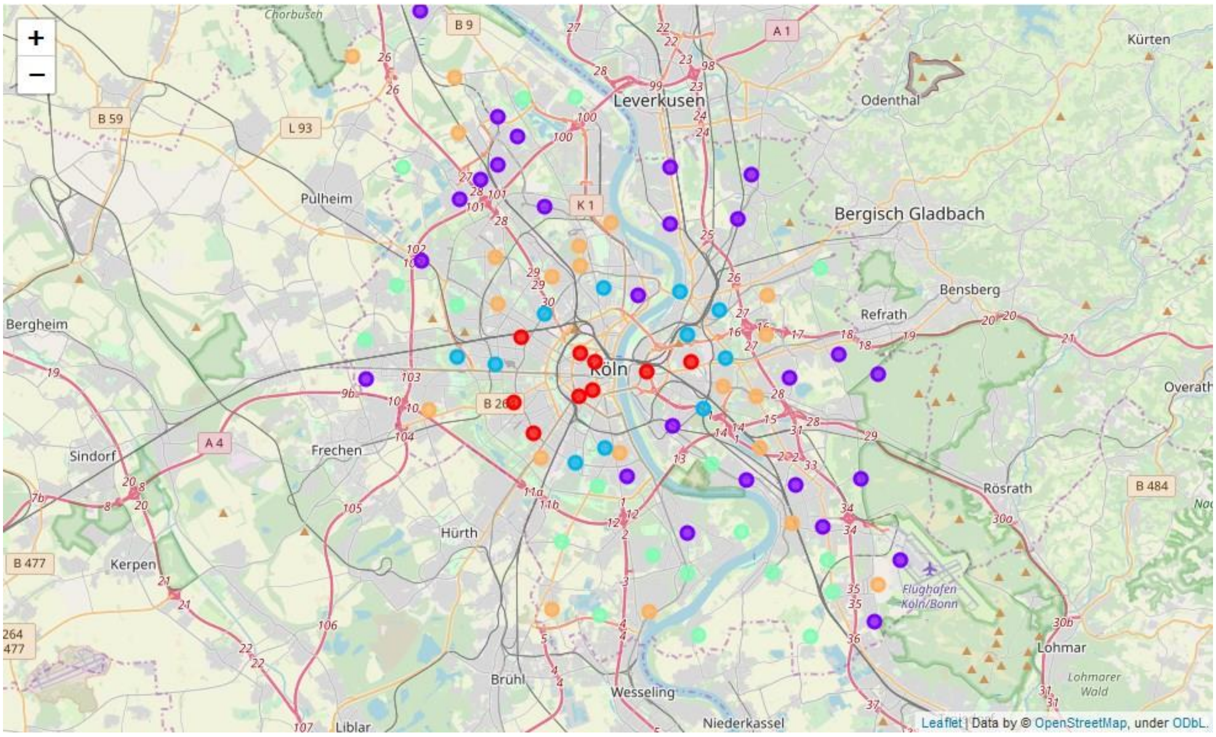



*Figure 5 Age Distribution Clustering: Districts with people of a similare age distribution structure have the same color.*

We see again radial clustering. But here, we can see a clearer distinction than before. Let's make this distinction visible at one example: Note how all the districts in the center of Cologne (horizontal center) are in the same cluster class (cluster 0 → red). Please look at the age distribution there:

| Red cluster | |
|---|---|
| 0-2 | 2.636131 |
| 3-5 | 2.030705 |
| 6-14 | 4.782664 |
| 15-17 | 1.557625 |
| 18-20 | 2.658425 |
| 21-34 | 31.367353 |
| 35-59 | 36.322382 |
| 60-64 | 4.182191 |
| 65-74 | 7.392915 |
| 75-79 | 3.271608 |
| 80+ | 3.798002 |

In comparison let's look at the age distributions from two outer cluster families:

| Violet cluster | | Orange cluster | |
|---|---|---|---|
| 0-2 | 2.755135 | 0-2 | 3.208497 |
| 3-5 | 2.790196 | 3-5 | 3.121043 |
| 6-14 | 8.350272 | 6-14 | 9.346335 |
| 15-17 | 3.048606 | 15-17 | 3.198538 |
| 18-20 | 3.060187 | 18-20 | 3.399559 |
| 21-34 | 16.454788 | 21-34 | 19.583120 |
| 35-59 | 34.952961 | 35-59 | 37.059361 |
| 60-64 | 5.952040 | 60-64 | 5.087842 |
| 65-74 | 10.995371 | 65-74 | 8.474782 |
| 75-79 | 5.543783 | 75-79 | 3.713653 |
| 80+ | 6.096660 | 80+ | 3.807268 |

We see that in the center live people in working age. This may be because, it's rather expensive to live there but at the same time it's a very busy place. People who get older, or who want to raise children move more to the outer districts. Which is visible from the higher amount of people in age 6-14 (children at school) and 65-74 (retired).

## 2.5 Investigating the Migration based on the different Aspects of Living

Finally, we want to put everything together to explain migration among districts in Cologne. Therefore, in the last sections we have tried to characterize districts based on the following aspects of living:

- Most Popular Venue Categories
- Area Usage
- Age Distribution among Inhabitants
- Household Size Distribution

For each of those aspects, a cluster analysis was performed, labelling districts based on similarities in those aspects. As a result, we can now assemble a data frame that holds all districts and their respective cluster label for each aspect:

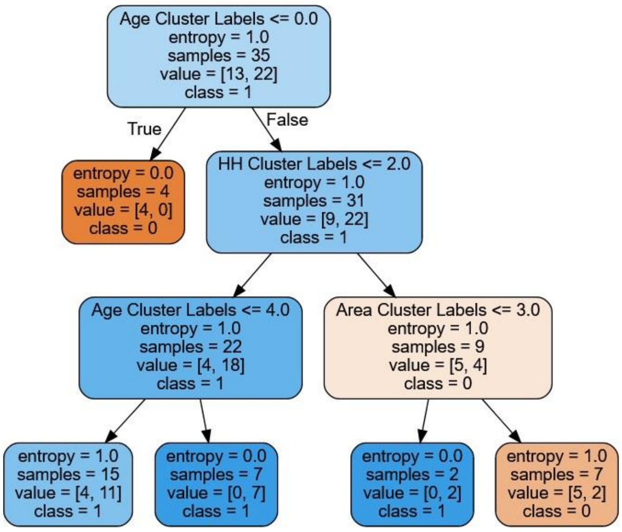
| | District | Latitude | Longitude | Venue Cluster Labels | Age Cluster Labels | Area Cluster Labels | HH Cluster Labels |
|---|---|---|---|---|---|---|---|
| 0 | Godorf | 50.852610 | 6.982218 | 3 | 4 | 5 | 1 |
| 1 | Lövenich | 50.948533 | 6.823829 | 3 | 3 | 0 | 1 |
| 2 | Weiden | 50.934514 | 6.824246 | 6 | 1 | 2 | 3 |
| 3 | Junkersdorf | 50.923891 | 6.859688 | 3 | 4 | 0 | 1 |
| 4 | Widdersdorf | 50.967660 | 6.841605 | 1 | 3 | 4 | 1 |

This forms our *independent data.* The dependent variable will be the migration. In order to get a measure for growing or shrinking districts, we use another data set that gives the net migration of each district. We encode this value in a new variable churn in the following way:

$$\text{Churn} = \begin{cases} 1 \text{ if net migration} > 0 \\ 0 \text{ if net migration} \leq 0 \end{cases}$$

| | District | Latitude | Longitude | Venue Cluster Labels | Age Cluster Labels | Area Cluster Labels | HH Cluster Labels | Churn |
|---|---|---|---|---|---|---|---|---|
| 0 | Godorf | 50.852610 | 6.982218 | 3 | 4 | 5 | 1 | 1 |
| 1 | Weiden | 50.934514 | 6.824246 | 6 | 1 | 2 | 3 | 1 |
| 2 | Junkersdorf | 50.923891 | 6.859688 | 3 | 4 | 0 | 1 | 1 |
| 3 | Widdersdorf | 50.967660 | 6.841605 | 1 | 3 | 4 | 1 | 1 |
| 4 | Vogelsang | 50.960676 | 6.875571 | 3 | 3 | 0 | 1 | 0 |

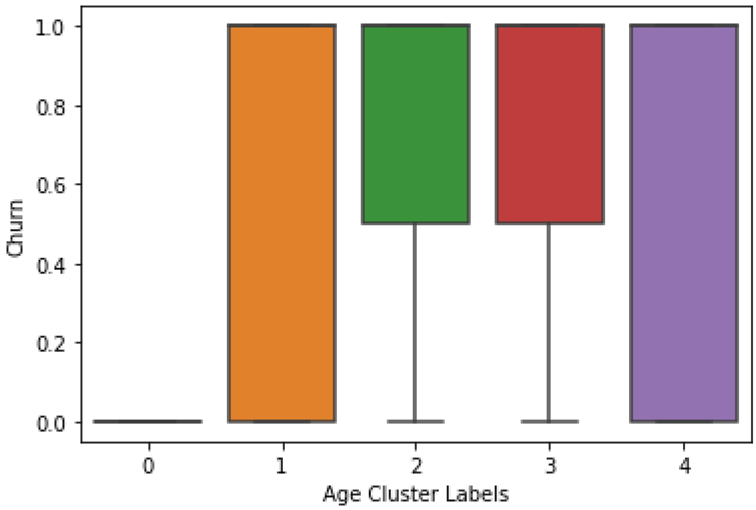
This will be the final data set. After splitting the data set, a decision tree is built having the columns with the cluster labels as independent variable and the churn data as dependent variable. The tree is restricted to a level depth of 3. The resulting tree looks like this:




Looking at the top features of the decision tree, we can see which categories seem to have the biggest impact on whether or not a district is growing or shrinking:

| | index | Feature | Importance |
|---|---|---|---|
| 0 | 1 | Age Cluster Labels | 0.602413 |
| 1 | 3 | HH Cluster Labels | 0.202085 |
| 2 | 2 | Area Cluster Labels | 0.195502 |
| 3 | 0 | Venue Cluster Labels | 0.000000 |

It can be seen, that primarily, the age distribution governs whether a district is growing or shrinking. Let’s therefore look at the age cluster labels that correspond to a high/low churn:




We see that for instance districts with age cluster 0 labels have a shrinking population. We remember, these where the ones in the city center, where most probably, costs of living are very high and where we had a high ratio of younger people living in 1Person households. As soon as the found a family, they will move to outer district, i.e. this makes sense. If we look for instance at the age distribution of cluster 3:

```
0-2         2.624339
 3-5        2.921833
 6-14       9.284933
```

```
15-17     3.396935
18-20     2.971762
21-34    13.878731
35-59    39.165455
60-64     6.060903
65-74    10.416385
75-79     4.601962
80+       4.676762
```

We see, that in those districts, older people are living that have already a family or they are retired. In both cases they probably have a lower interest in leaving their place, i.e. they stay while people from other districts enter.

# 3 Conclusion

In the last chapters, we have characterized each district of Cologne by virtue of several aspects:

- Most Popular Venue Categories
- Area Usage
- Age Distribution among Inhabitants
- Household Size Distribution

For each of those aspects, a cluster analysis was performed in order to find similar districts based on their environment on basis of the respective aspects.

Finally, this data was associated to the growth/shrinkage of districts by using a decision tree algorithm. This showed, that the by far the biggest impact on the growth of a district was given by the age distribution of the district.

Looking more profoundly on the age distribution of growing or shrinking districts, this result also made perfect sense: A negative net migration can be associated to districts with younger (pre-family-founding) persons, positive net migration can be found in districts with older age distributions. This interpretation also matches the household size distribution of those districts which made the 2$^{nd}$ biggest impact on net migration: In shrinking districts, there is a vast majority of people living in single person households.
As soon as people decide to found a family, they will move to the outer districts of Cologne, where the costs of living are lower whereas people already living there, will stay. This leads to a growth of the outer districts of Cologne and a negative net migration of people in the inner districts of Cologne.

Please note, that the migration investigated here, is only the migration within Cologne. A shrinking net migration here does not mean, that the absolute population of those districts sink since people from outside Cologne may move to the inner districts of Cologne and balance the flux to the outer districts. In fact, they probably over-compensate it, but that may be investigated in a separate report.

# 4 Appendix

## 4.1 Most frequent venue categories per district

```
--------Altstadt-Nord---------
                venue  freq
0  Italian Restaurant  0.07
1                Café  0.06
2              Bakery  0.05


---------Altstadt-Süd---------
              venue  freq
0             Hotel  0.16
1  Sushi Restaurant  0.08
2              Café  0.08


----------Bayenthal-----------
         venue  freq
0  Supermarket  0.16
1       Bakery  0.08
2     Bus Stop  0.08


----------Bickendorf----------
         venue  freq
0          Pub  0.17
1  Supermarket  0.17
2         Bank  0.08


-------Bilderstöckchen--------
       venue  freq
0   Bus Stop  0.50
1  Drugstore  0.25
2       Café  0.25


------Bocklem./Mengenich------
              venue  freq
0  Greek Restaurant  0.25
1      Intersection  0.25
2      Soccer Field  0.25


----------Braunsfeld----------
               venue  freq
```

```
0            Bakery  0.18
1         Drugstore  0.12
2  German Restaurant  0.12


------------Brück-------------
               venue  freq
0             Forest   0.5
1             Bakery   0.5
2  African Restaurant   0.0


----------Buchforst-----------
          venue  freq
0         Hotel  0.33
1  Tram Station  0.17
2   Supermarket  0.17


-----------Buchheim-----------
             venue  freq
0  Greek Restaurant  0.25
1       Supermarket  0.25
2     Big Box Store  0.25


----------Chorweiler----------
              venue  freq
0  Electronics Store   0.2
1         Restaurant   0.2
2        Supermarket   0.2


----------Dellbrück-----------
            venue  freq
0       Drugstore  0.11
1          Bakery  0.11
2  Farmers Market  0.11


------------Deutz-------------
            venue  freq
0  Ice Cream Shop  0.05
1          Bakery  0.05
2          Bistro  0.05


-----------Dünnwald-----------
           venue  freq
```

```
0     Restaurant  0.33
1  Shopping Mall  0.33
2           Park  0.33


----------Ehrenfeld-----------
         venue  freq
0         Café  0.09
1          Bar  0.09
2  Supermarket  0.06


-------------Eil--------------
                venue  freq
0  Athletics & Sports   1.0
1  African Restaurant   0.0
2        Perfume Shop   0.0


-----------Elsdorf------------
                        venue  freq
0  Construction & Landscaping   1.0
1       Portuguese Restaurant   0.0
2                Optical Shop   0.0


------------Ensen-------------
         venue  freq
0  Gas Station  0.25
1        River  0.25
2       Bakery  0.25


----------Finkenberg----------
               venue  freq
0               Café   0.1
1  Food & Drink Shop   0.1
2                Gym   0.1


-----------Flittard-----------
          venue  freq
0    Food Truck  0.25
1      Bus Stop  0.25
2  Soccer Field  0.25


------------Godorf------------
             venue  freq
```

```
0      Snack Place  0.33
1   Breakfast Spot  0.33
2  Harbor / Marina  0.33


--------Gremberghoven---------
                venue  freq
0  Light Rail Station   1.0
1  African Restaurant   0.0
2        Perfume Shop   0.0


-----------Grengel------------
              venue  freq
0   Airport Service   0.5
1             Plane   0.5
2  Pedestrian Plaza   0.0


-----------Hahnwald-----------
                venue  freq
0           Racetrack   1.0
1  African Restaurant   0.0
2        Perfume Shop   0.0


---------Heimersdorf----------
                venue  freq
0          Restaurant   1.0
1  African Restaurant   0.0
2        Perfume Shop   0.0


-----------Holweide-----------
          venue  freq
0  Tram Station   0.5
1      Bus Stop   0.5
2  Perfume Shop   0.0


-------Humboldt/Gremb.--------
             venue  freq
0        Pet Store  0.25
1  Organic Grocery  0.25
2    Big Box Store  0.25


----------Höhenberg-----------
            venue  freq
```

```
0            Pool  0.14
1  Discount Store  0.14
2     Supermarket  0.14


----------Höhenhaus-----------
               venue  freq
0       Tram Station   0.5
1  Outdoor Sculpture   0.5
2          Nightclub   0.0


---------Junkersdorf----------
                    venue  freq
0             Supermarket  0.17
1  Furniture / Home Store  0.17
2                  Bakery  0.08


-------------Kalk-------------
                venue  freq
0              Bakery  0.14
1    Doner Restaurant  0.09
2  Turkish Restaurant  0.05


---------Klettenberg----------
                 venue  freq
0        Train Station   0.2
1         Dessert Shop   0.2
2  Rental Car Location   0.2


------------Langel------------
            venue  freq
0             Bar   0.2
1  Ice Cream Shop   0.2
2          Market   0.2


------------Libur-------------
                venue  freq
0         Golf Course   1.0
1  Israeli Restaurant   0.0
2              Office   0.0


-------------Lind-------------
              venue  freq
```

```
0  Business Service  0.33
1          Bus Stop  0.33
2           Theater  0.33


----------Lindenthal----------
               venue  freq
0  German Restaurant  0.17
1         Restaurant  0.17
2               Café  0.08


----------Lindweiler----------
                venue  freq
0         Flower Shop   1.0
1  African Restaurant   0.0
2        Perfume Shop   0.0


----------Longerich-----------
          venue  freq
0  Tram Station   0.4
1   Supermarket   0.2
2      Hospital   0.2


-----------Lövenich-----------
       venue  freq
0        Bar  0.25
1   Bus Stop  0.25
2  Gastropub  0.25


----------Marienburg----------
           venue  freq
0  Grocery Store   0.2
1      Wine Shop   0.2
2       Bus Stop   0.2


----------Mauenheim-----------
               venue  freq
0        Supermarket  0.33
1  German Restaurant  0.17
2       Soccer Field  0.17


-----------Merheim------------
                venue  freq
```

```
0   German Restaurant   0.5
1                Park   0.5
2  African Restaurant   0.0


----------Merkenich-----------
                  venue  freq
0                  Café   1.0
1     African Restaurant   0.0
2  Portuguese Restaurant   0.0


----------Meschenich----------
               venue  freq
0            Brewery   0.5
1            Stables   0.5
2  African Restaurant   0.0


-----------Mülheim------------
         venue  freq
0         Café  0.15
1  Supermarket  0.15
2    BBQ Joint  0.08


---------Müngersdorf----------
              venue  freq
0              Café  0.25
1  German Restaurant  0.25
2        Pizza Place  0.25


-----------Neubrück-----------
              venue  freq
0          Drugstore  0.25
1  Convenience Store  0.25
2              Plaza  0.25


---------Neuehrenfeld---------
               venue  freq
0  Italian Restaurant   0.1
1         Supermarket   0.1
2          Restaurant   0.1


--------Neustadt-Nord---------
             venue  freq
```

```
0              Hotel  0.10
1  Sushi Restaurant  0.07
2                Bar  0.05


---------Neustadt-Süd---------
                venue  freq
0  Italian Restaurant  0.09
1                 Bar  0.09
2              Bakery  0.05


------------Niehl-------------
               venue  freq
0  German Restaurant   0.2
1       Intersection   0.2
2      Ice Cream Shop   0.2


------------Nippes------------
                venue  freq
0                Café  0.15
1  Italian Restaurant  0.09
2                 Bar  0.09


----------Ossendorf-----------
              venue  freq
0         Racetrack  0.11
1  Indoor Play Area  0.11
2     Historic Site  0.11


-----------Ostheim------------
              venue  freq
0       Supermarket  0.75
1      Tram Station  0.25
2  Pedestrian Plaza  0.00


------------Pesch-------------
                venue  freq
0          Steakhouse  0.14
1  Italian Restaurant  0.14
2         Supermarket  0.14


-------------Poll-------------
            venue  freq
```

```
0    Tram Station  0.18
1       Drugstore  0.09
2  Ice Cream Shop  0.09


-------------Porz-------------
                venue  freq
0  Turkish Restaurant  0.12
1       Train Station  0.06
2           BBQ Joint  0.06


----------Raderberg-----------
                  venue  freq
0           Supermarket  0.31
1  Gym / Fitness Center  0.12
2             Drugstore  0.06


----------Raderthal-----------
              venue  freq
0  Greek Restaurant  0.25
1        Playground  0.25
2              Park  0.25


------------Riehl-------------
         venue  freq
0  Zoo Exhibit  0.62
1       Bakery  0.04
2   Playground  0.04


---------Rodenkirchen---------
          venue  freq
0          Pool   1.0
1  Perfume Shop   0.0
2  Optical Shop   0.0


-----------Rondorf------------
                venue  freq
0    Insurance Office   1.0
1  African Restaurant   0.0
2        Perfume Shop   0.0


-----------Seeberg------------
                venue  freq
```

```
0  Turkish Restaurant  0.50
1      Discount Store  0.25
2  Light Rail Station  0.25


----------Stammheim-----------
            venue  freq
0     Supermarket  0.25
1  Ice Cream Shop  0.25
2        Pharmacy  0.25


-------------Sülz-------------
                   venue  freq
0                 Bakery  0.14
1  Vietnamese Restaurant  0.14
2                    Pub  0.07


------------Sürth-------------
               venue  freq
0       Tram Station   0.2
1        Supermarket   0.2
2  German Restaurant   0.2


------------Urbach------------
               venue  freq
0          Drugstore  0.11
1        Pizza Place  0.11
2  Electronics Store  0.11


------------Vingst------------
           venue  freq
0    Supermarket  0.31
1         Bakery  0.23
2  Metro Station  0.08


----------Vogelsang-----------
                venue  freq
0              Lawyer   0.5
1              Bakery   0.5
2  African Restaurant   0.0


-------------Wahn-------------
         venue  freq
```

```
0     Platform  0.27
1  Supermarket  0.18
2        Hotel  0.09


----------Wahnheide-----------
                venue  freq
0  Italian Restaurant  0.33
1               Hotel  0.33
2    Greek Restaurant  0.17


------------Weiden------------
               venue  freq
0  German Restaurant   0.5
1      Historic Site   0.5
2             Office   0.0


---------Weidenpesch----------
               venue  freq
0   Greek Restaurant  0.25
1  Korean Restaurant  0.25
2      Ice Cream Shop  0.25


----------Westhoven-----------
                  venue  freq
0      Business Service  0.25
1  Gym / Fitness Center  0.25
2           Gas Station  0.25


---------Widdersdorf----------
           venue  freq
0     Food Truck  0.33
1    Supermarket  0.17
2  Soccer Field  0.17


----------Worringen-----------
                venue  freq
0         Supermarket  0.25
1             Taverna  0.25
2  Photography Studio  0.25


----------Zollstock-----------
                         venue  freq
```

```
0            German Restaurant  0.29
1               Ice Cream Shop  0.29
2  Eastern European Restaurant  0.14


-----------Zündorf------------
                venue  freq
0         Supermarket  0.25
1  Italian Restaurant  0.12
2        Soccer Field  0.12
```