**A guide to the neighborhoods of Düsseldorf, DE**

Oliver Möll

January 14, 2020

# Introduction

## Background

Düsseldorf is the 7th largest city of Germany and according to a study of the consulting company Mercer in 2019, one of the global top 10 cities it is worth living in.

Today, Düsseldorf is divided into 10 departments and 50 neighborhoods. Originally it was founded in the 12th/13th century next to the rivers Rhine and Düssel and has since grown into a city with a total population of ~620 thousands. In 2018 the population increased by ~14 thousands.

Due to its heritage and expansion over more than eight centuries, it is coined by a rather urbanized city center with a mix of historic and modernized buildings, accompanied by diverse neighborhoods surrounding the city centre.

## Problem

Imagine you would like to move to Düsseldorf. One of the first questions is, to decide where you would like to live within the city. So the following two questions might be of relevance for you: Which neighborhoods would suit your preferences with regards to urban design? Which neighborhoods are more popular, i.e. in higher demand than others?

This project aims to provide a clustering of the different neighborhoods based on selected social and urban design specific indicators (e.g. % of urbanised and non-urbanised area, population per square kilometre), as well as the type of venues located in the neighborhoods. Next to the clustering of neighborhoods, an analysis of the change of population per neighborhood over the last five years aims to provide an additional perspective on the neighborhoods and their development.

## Target Group

Anybody who is interested in the design and variety of the different neighborhoods of Düsseldorf is a potential interest in this report and might accompany any offical city guide. Thereby it does not matter, if you would like to move to Düsseldorf or if you just would like to visit the city and plan your city tour.

# Data Description

## Data sources

The primary source of this project is a dedicated resource for datasets of Düsseldorf:

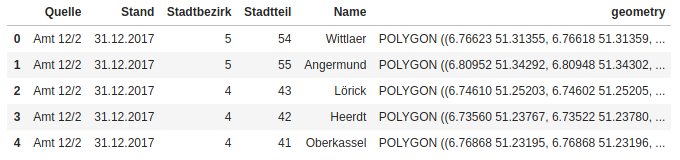
The city of Düsseldorf is publishing a large variety of datasets via the webpage [http://opendata.duesseldorf.de](http://opendata.duesseldorf.de/). For this project I will focus on the datasets that provide the geometrical coordinates of the neighborhoods and social areas of Düsseldorf, as well as the data on population changes and composition of each social area.

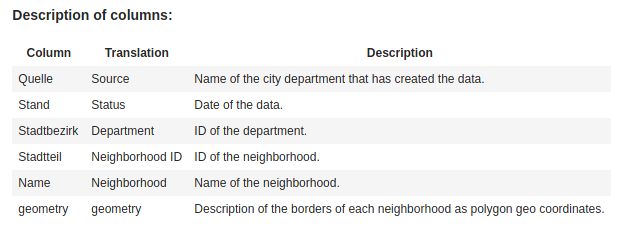
The secondary source of this project is Foursquare and its API, which allows to perform requests on recommended venues in a specific radius around a specific geo location. Each venue is thereby categorized in a hierarchical category structure and thereby allows to cluster the different locations by their type.

In the iPythonNotebook of this project, for each data source the relevant link is provided.

## **Data Examples**

### Neighborhoods Geo Location Data



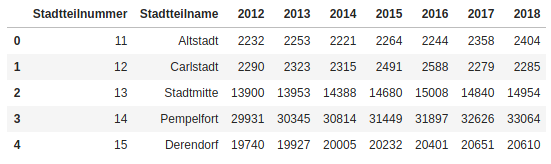


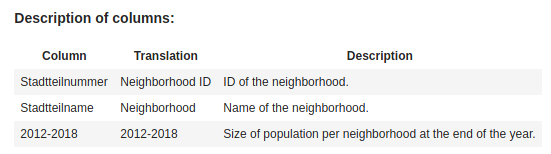
**Content:**

The data of the neighboorhoods has been created by the department 'Amt 12/2' of the citiy of Duesseldorf with status of December 31st, 2017. The city area of Duesseldorf is divided into 10 departments and 50 neighborhoods. Each neighborhood is assigned to one department. Departments only have an ID and no name, whereas neighborhoods have an ID and a name. The neighborhood ID is a merged key based on the ID of the department and a digit for the neighborhood.

As the analysis will be based on the neighborhoods only the columns 'Neighborhood ID', 'Neighborhood and 'geometry' will be used.

### Population



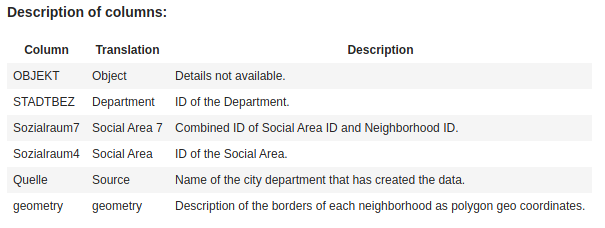
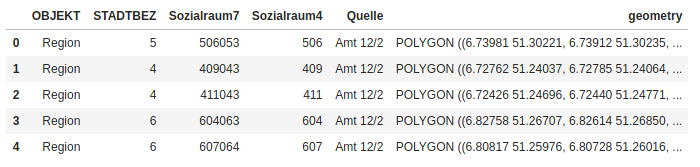


**Content:**

The data includes the total population of each neighborhood by the end of the years 2012 to 2018. The population data is provided for all 50 neighborhoods.

For the analysis the data for the past five years (2013 to 2018) will be used.

### Social Areas (Geo Location Data)

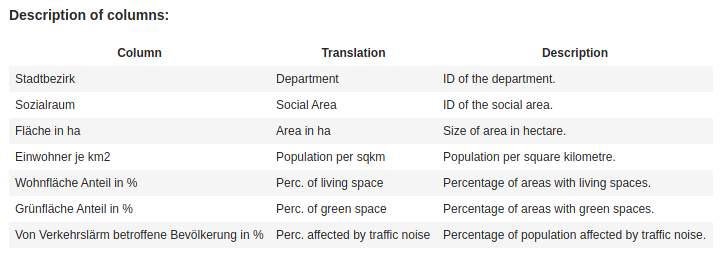
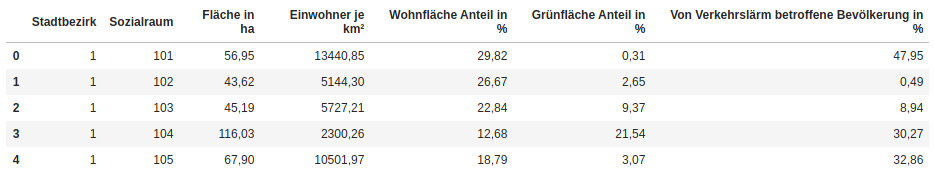


**Content:**

The data includes the geo location data for the 179 social areas in Düsseldorf with status 2018. The social areas are mapped to the departments and with three exceptions, also to neighborhoods. The key in column "Sozialraum7" is a merged key, in which the last three digits represent the corresponding Neighborhood ID and all digits before that represent the Social Area ID ('Sozialraum4').

As transparent in the 'Description' of the DataFrame, the value '0' is appearing three times in the column 'Sozialraum7'. All other values are unique. The three '0' entries need to be further analyzed and decided on how to cope with those.

### Social Areas (Geographical and Environmental Data)



**Content:**

The data includes the geographic and environmental data for the 179 social areas in Düsseldorf with status 2016. Besides the general information of the corresponding department, social area ID and total size in hectare, the data comprises of four different figures that can be summed up in three social characteristics:

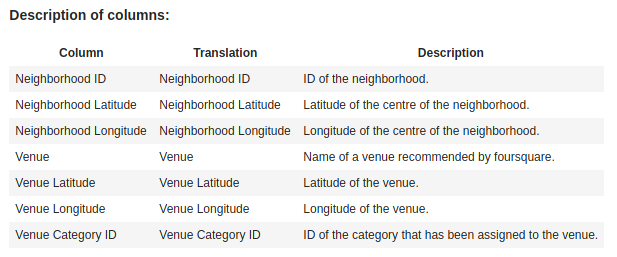
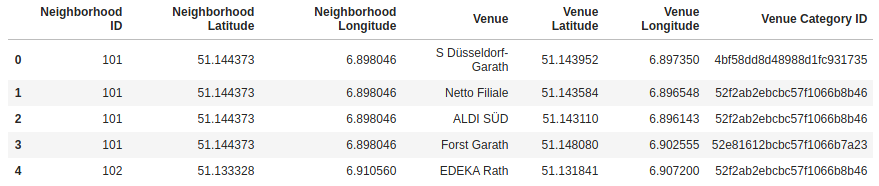
1) Population density: Population per square kilometre.

2) Land use: Percentage of area used as living spaces or green spaces.

3) Noise pollution: Percentage of population affected by traffic noise.

This data will be aggregated on neighborhood level and combined with the information about recommended venues in the neighborhood (provided by foursquare) to cluster the neighborhoods.

### Recommended Venues

**Content:**

The data consists of the Neighborhood ID and the geo coordinates of its centre as well as the venue name, its geo coordinates and the ID of the category it has been assigned to. This data will be aggregated on neighborhood and category level to get the count of the venues per category per neighborhood. This will provide a set of features that allows a clustering of similar neighborhoods based on the recommended venues.

As foursquare provides a hierarchical tree for the venue categories, it needs to analysed, if all venues have been categorized on the same hierarchical level, before performing the aggregation. This should provide sligthly less accurate but more consistent results, as e.g. a chinese restaurant could be categorized in one case as 'Asian' (level 2) and in another case as 'Chinese' (level 3). Both categorizations would be correct, but the clustering model would not recognize them as being the same. After the analysis on the categories provided, it will be decided on which level of the hierarchy all venues will be harmonized.