Investigation on Population Migration within Cologne

# 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.

## 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.

## 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.