

Roland Teschke

# Target areas to launch new gyms in Frankfurt, Germany

Capstone project, IBM Data Science Professional Certificate

1.2.2021

## Table of contents

1) Introduction and business problem .....	2
a) Frankfurt, Germany: A brief summary about the city .....	2
b) Growing fitness market in Germany .....	2
c) Target audience, business problem and benefits provided by this report .....	2
2) Data .....	3
a) Overview of the data sources.....	3
b) GeoJSON data .....	4
c) Location data for gyms in and around Frankfurt.....	5
d) Statistical data .....	7
e) Other computed data, ratios and indices.....	13

# 1) Introduction and business problem

## a) Frankfurt, Germany: A brief summary about the city<sup>1</sup>

Frankfurt, officially Frankfurt am Main, is the most populous city in the German state of Hesse. Its 763,380 inhabitants as of December 31, 2019 make it the fifth-most populous city in Germany. The city is at the centre of the larger Rhine-Main Metropolitan Region, which is Germany's second-largest metropolitan region after the Rhine-Ruhr Region. Frankfurt is a global hub for commerce, culture, education, tourism and transportation. Frankfurt Airport is Germany's busiest. Frankfurt is one of the major financial centres of the European continent, with the headquarters of the European Central Bank, Deutsche Bundesbank, Frankfurt Stock Exchange, Deutsche Bank, DZ Bank, KfW, Commerzbank, to name just a few.

## b) Growing fitness market in Germany<sup>2</sup>

Fitness market revenues and number of fitness-related venues in Germany have been increasing significantly in the previous years. Due to this major trend, it may be assumed that additional facilities are planned to be established in the near future.

As we will discover within this research, Frankfurt already has many gyms as of January 2021.

## c) Target audience, business problem and benefits provided by this report

### **Target audience:**

This report is targeted at

- owners and managers of already established gyms, gym franchises who intend to expand their business and set a new gyms in Frankfurt am Main, and
- people interested at setting up a new gym (i.e. who have previously not owned and/or managed a gym).

### **Business problem:**

Given the vast existing competition and the potential desire to establish new gym, insights are desired as to which areas / boroughs might be most suitable in promising business success when setting up new gyms in Frankfurt am Main, Germany.

### **Benefits of this report:**

As Frankfurt is a large and diversity city, this report will

- provide selected information about Frankfurt boroughs with regards to existing gyms, alternative facilities, average age, population, income data, and
- gain insights which factors might also be considered to establish a gym in Frankfurt (focusing on selected data, not being fully exhaustive), and
- provide a proposal which areas might be most interesting to launch new gyms in Frankfurt.

---

<sup>1</sup> Source: Abstract, based on <https://en.wikipedia.org/wiki/Frankfurt> - slightly amended for this report

<sup>2</sup> Source: <https://de.statista.com/statistik/daten/studie/6228/umfrage/umsatz-der-fitness-branche-in-deutschland/#statisticContainer>

## 2) Data

### a) Overview of the data sources

This report is based on several data sources as follows:

- GeoJSON data for Frankfurt boroughs
- Gym locations in Frankfurt and surrounding areas, as offered by Foursquare
- Publicly available statistical data for boroughs of Frankfurt, Germany
- Self-computed data which is derived from processing the above mentioned data sources further

The following sub-chapters are going to provide more detailed information for each of these data sources.

## b) GeoJSON data

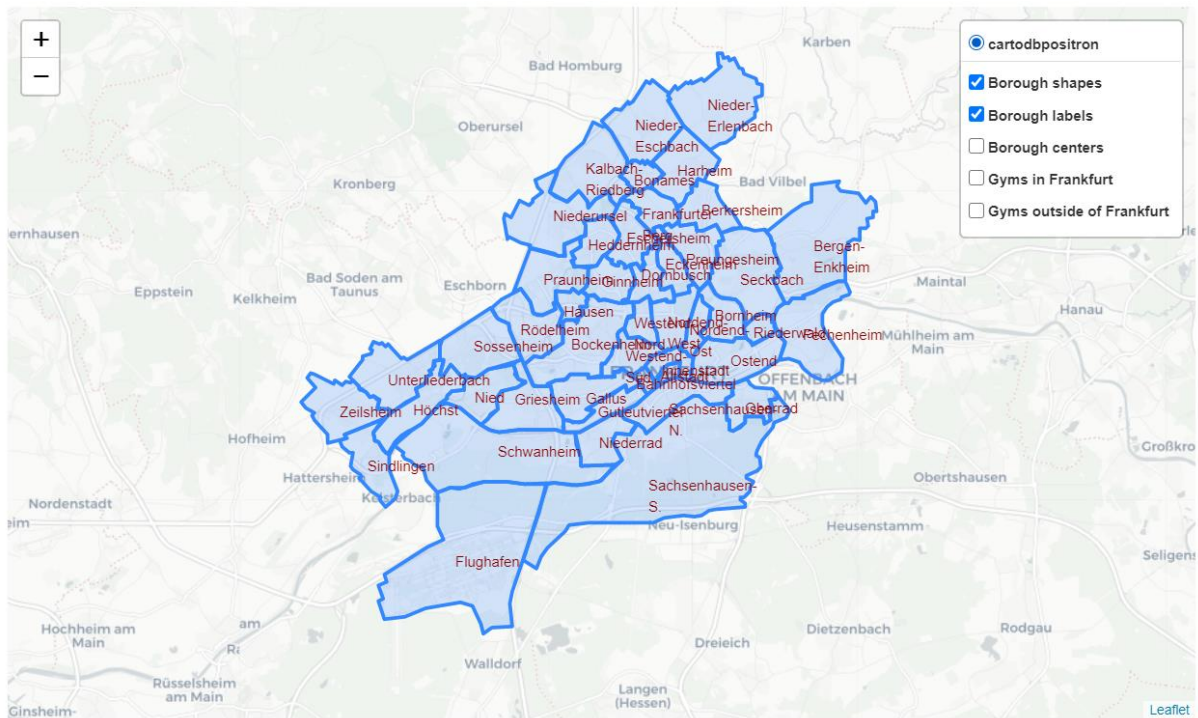
Source:

[https://opendata-esri-de.opendata.arcgis.com/datasets/ca64da7abad04c0eb8717ca3ec486cae\\_0?geometry=7.594%2C49.967%2C9.680%2C50.275](https://opendata-esri-de.opendata.arcgis.com/datasets/ca64da7abad04c0eb8717ca3ec486cae_0?geometry=7.594%2C49.967%2C9.680%2C50.275)

Publically and officially available GeoJSON data has been retrieved for Frankfurt am Main in order to be able to properly visualize selected information in maps and choropleth maps for each of it's 46 boroughs.

Each borough center's coordinates have been computed in order to request gym information from Foursquare subsequently.

Example:



### c) Location data for gyms in and around Frankfurt

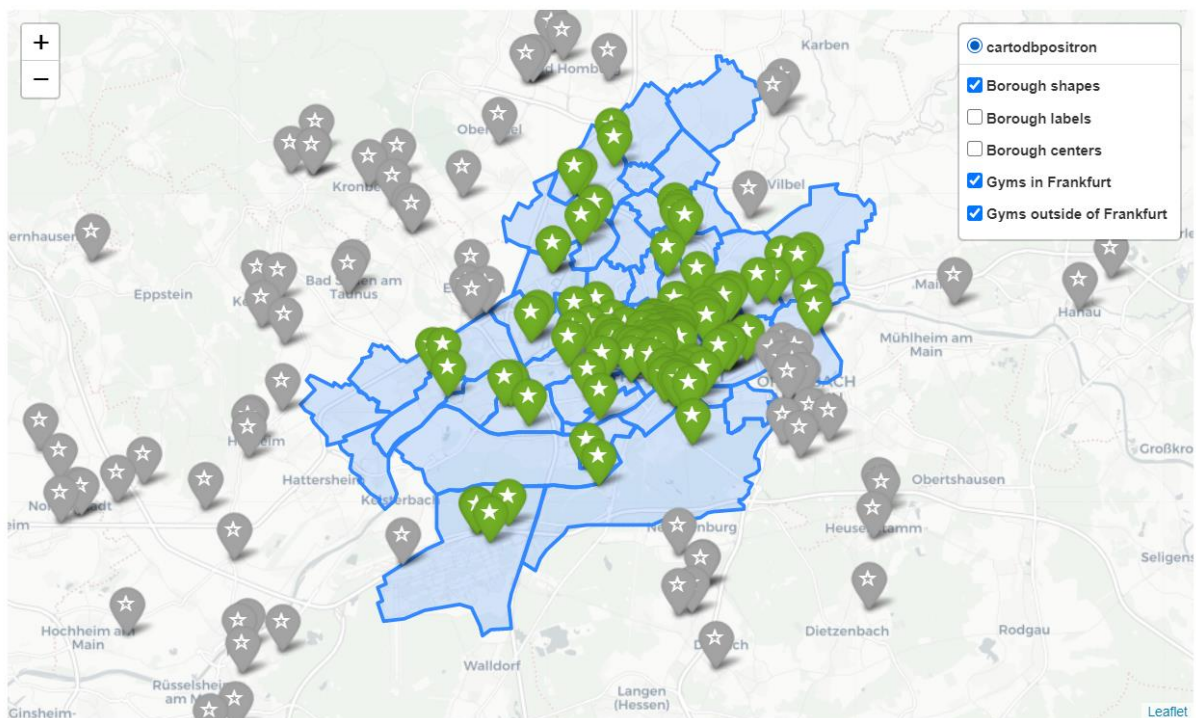
Source: <https://developer.foursquare.com/> , <https://api.foursquare.com>

The Foursquare API has been used in order to retrieve all gyms in an around Frankfurt for each of the 46 boroughs.

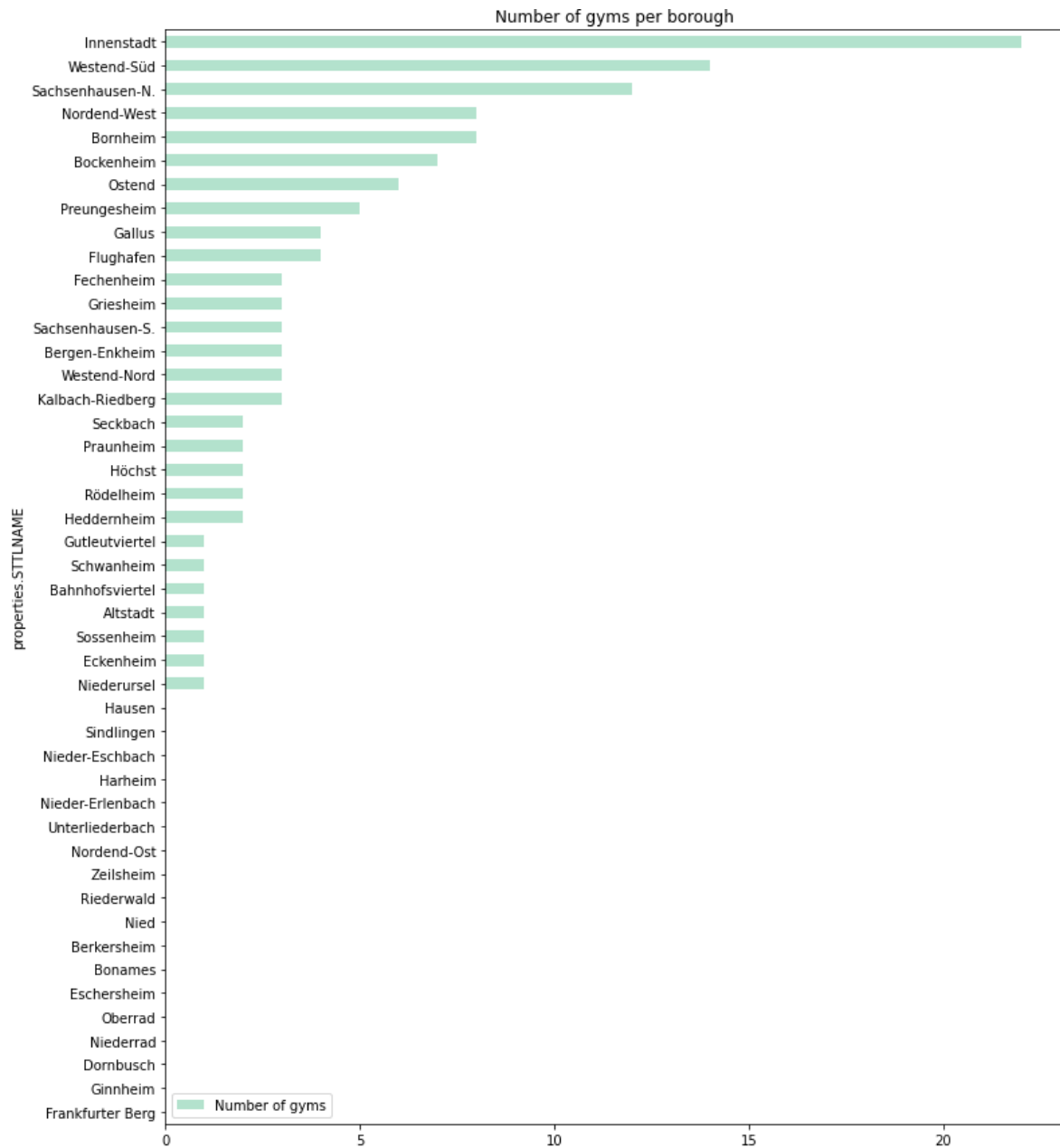
The data was processed further as follows:

- Location retrieval was consciously limited to include gyms only for each borough (based on the borough center locations)
- Duplicates had to be removed since the borough area sizes (i.e. also borough center distances) vary significantly. Approximately 200 gyms have been retrieved in total<sup>3</sup>.
- Gyms were classified as either to be within or outside of Frankfurt (**entire city**)
- Gyms were classified as either to be within the given **boroughs**

Examples:



<sup>3</sup> very slightly varying depend on when the real-time request was performed



## d) Statistical data

Source: [https://statistik.stadt-frankfurt.de/strukturdatenatlas/downloads/Stadtteile\\_FFM.xlsx](https://statistik.stadt-frankfurt.de/strukturdatenatlas/downloads/Stadtteile_FFM.xlsx)

The city of Frankfurt offers a vast array of statistical data. The data selected for this research is the most up-to-date publicly available - that is, data from end of Dec 2019.

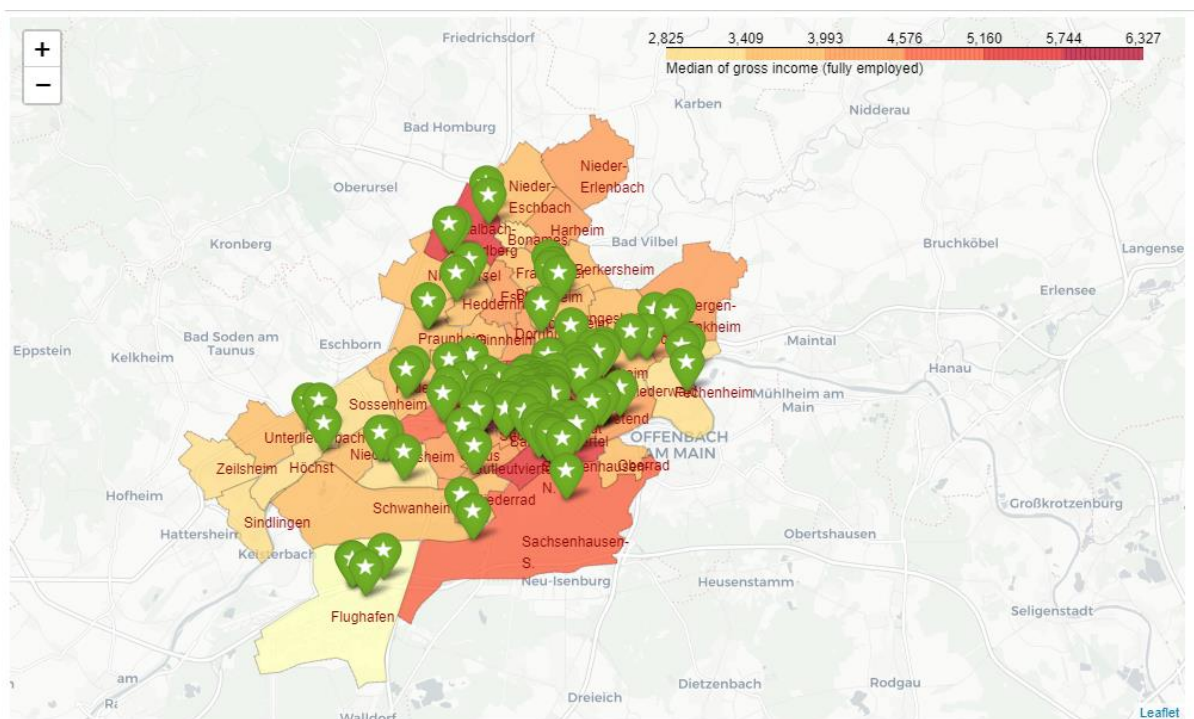
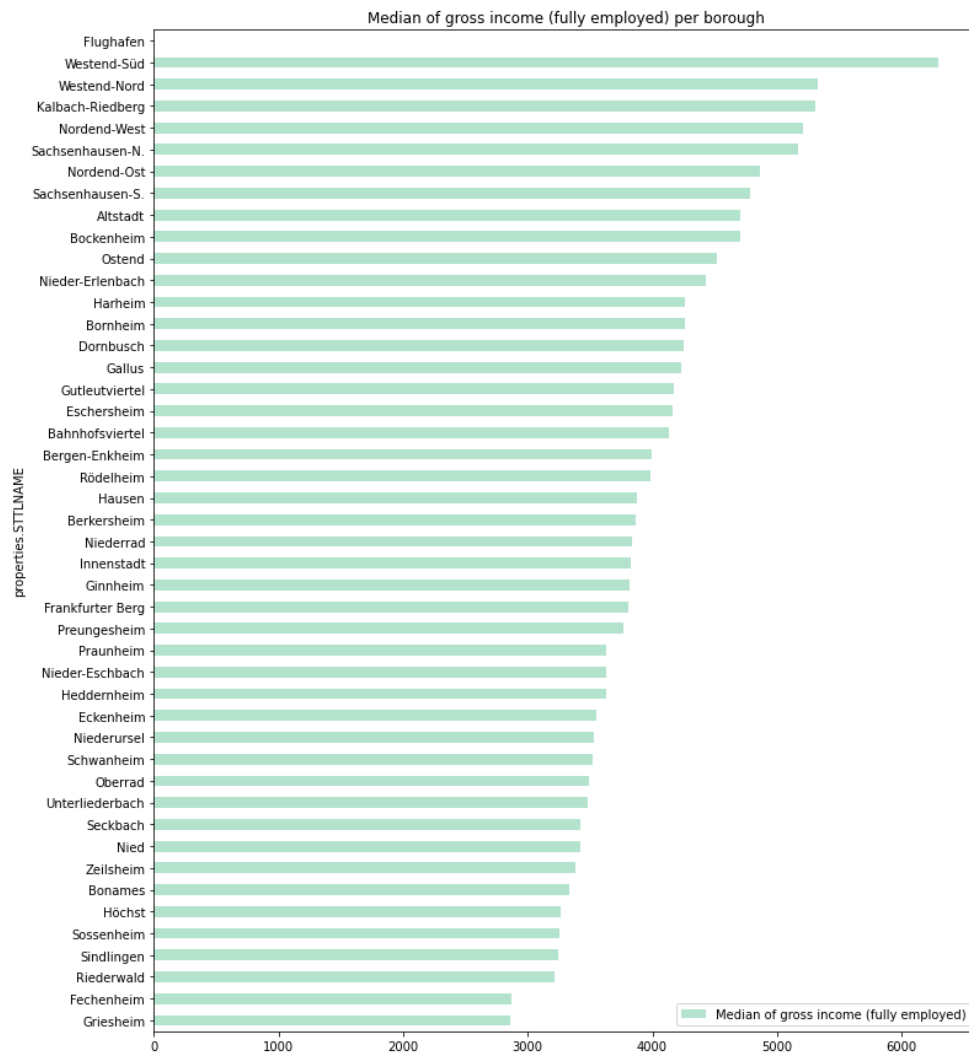
- The raw data was downloaded as an Excel file
- Potentially relevant metrics have been selected in the Excel file
- The excel cells did not contain clean header data as it contained several grouping mechanisms, so the data was slightly preprocessed in Excel in order to support a clean import to the Python notebook
- Subsequently been imported into Python notebook for further data processing as a Pandas dataframe

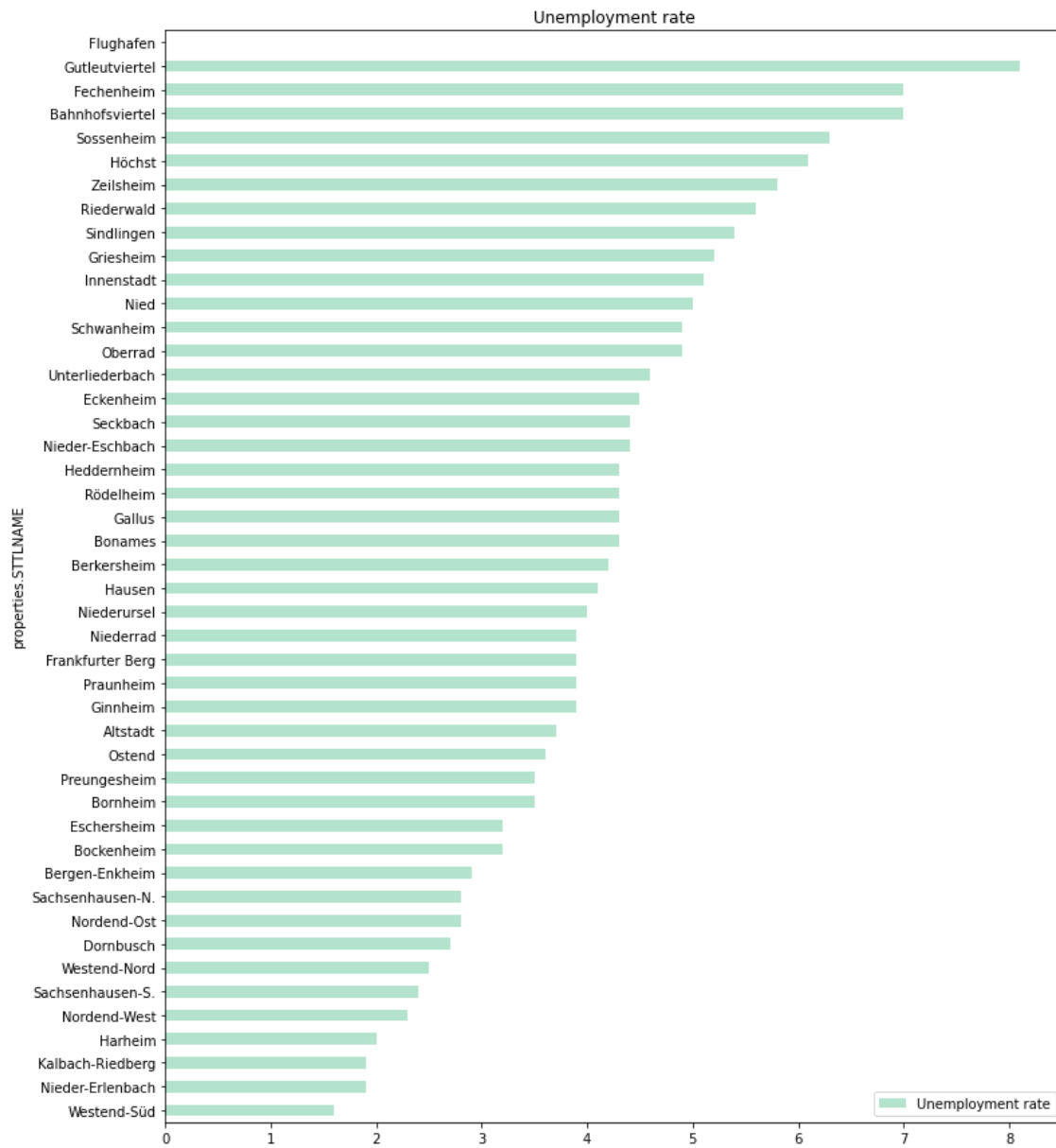
Several statistics have been selected for further research as follows:

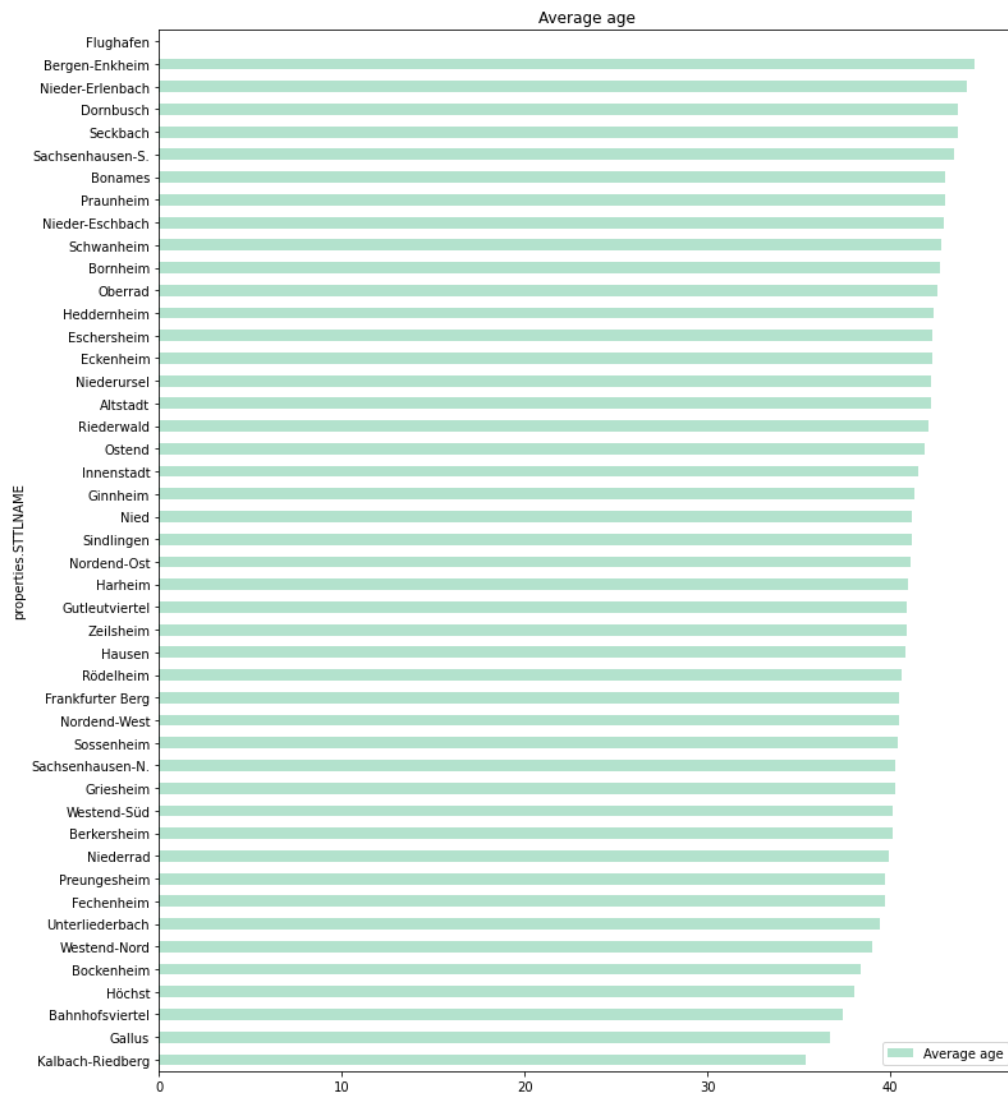
- Area information
  - Borough size
  - Size of sports, leisure and recreation areas
- Population details
  - Households
  - Inhabitants
  - People per household
  - Population density per ha
- Socio-demographic
  - Average age
  - Gender (female, male)
- Employment related
  - Employment total
  - Employment rate
  - Unemployment rate
- Financial
  - Median of gross income (fully employed)

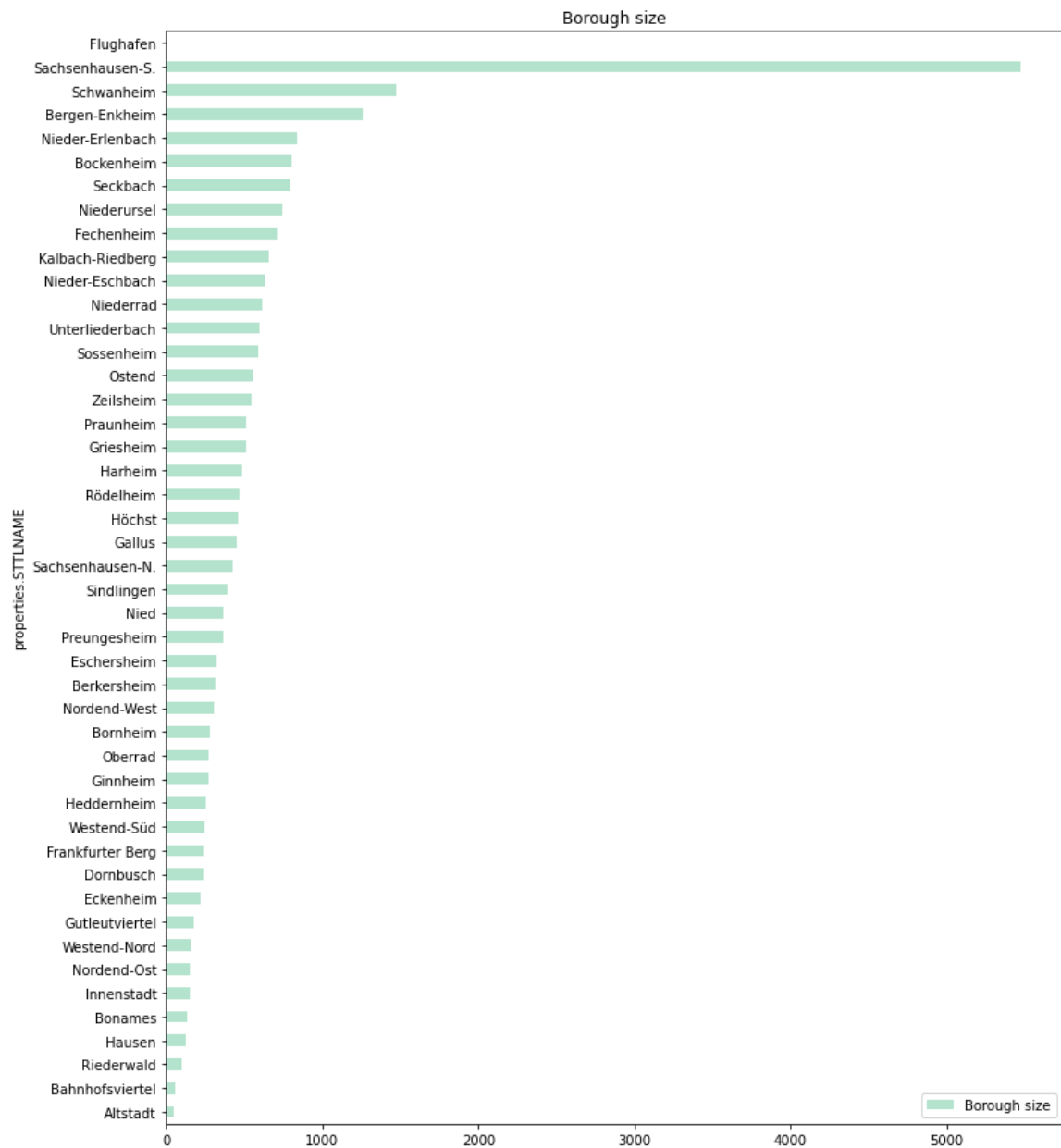
The following examples illustrate variances for the data retrieved for the various boroughs.

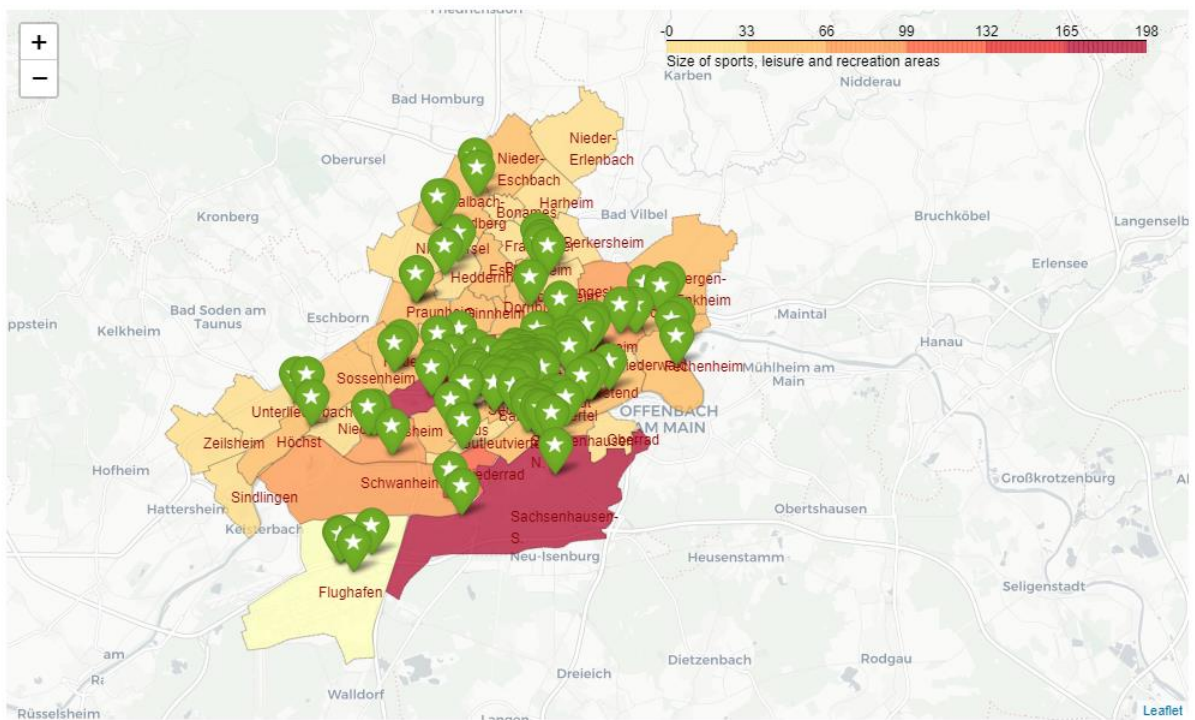
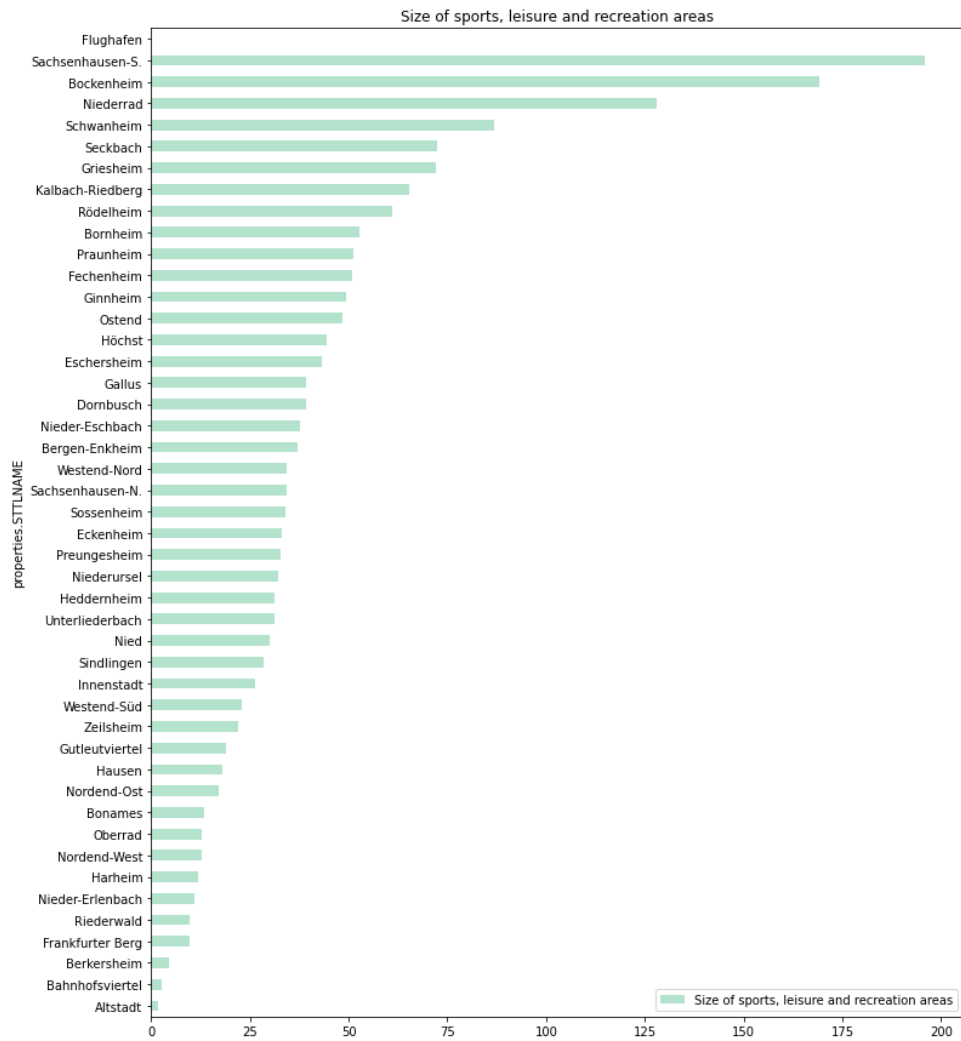












## e) Other computed data, ratios and indices

The following data was computed based on processing the various data sources and as either a prerequisite for further processing, or as a finding of some analyses described in more detail in the upcoming methodology and results sections:

- **Totals:**
  - o Amount of total income of employed inhabitants for each borough (= inhabitants \* employment rate \* median gross income)
  - o Number of gyms per borough
- **Boolean:**
  - o Determine whether gyms are inside or outside of Frankfurt (entire city)
  - o Determine whether gyms are inside specific boroughs of Frankfurt
- **Categories:**
  - o Gym competition level (based on number of gyms)
  - o Clusters of similarity based on a few selected variables (see cluster in analyses in subsequent report sections)
- **Continuous:**
  - o GridSearch results to help determine the number of clusters to choose for a cluster analysis
- **Gym expansion index**
  - o Gym expansion target index for each borough. It is important to note that the variables had to be scaled upfront to ensure the different ranges of the 3 variables can be used to create a meaningful target expansion index.  
The index has been computed **from normalized values** as
    - + Amount of total income of employed inhabitants for each borough
    - 50% of the ratio of sports, leisure and recreation areas
    - 50% of the number of gyms in the borough

Here are some examples of the computed data:



