BKA_Kriminalstatistik_2019_data_analysis

Peter von Bodelschwingh, Joshua Gawenda

12 12 2020

Introduction

This dataset is taken from the Bundeskriminalamt - the leading institution of crime detection in germany. The original dataset can be found here: BKA Dataset

It contains data about all crimes recorded in 2019. This analysis divides the statistics into all 16 german provinces. The goal of this analysis is to present the differences between these provinces by categories of crime, population, sucessful and non successfull enlisted crimes. Further documentation is provided in the "Interpretationshilfen" subfolder of this project.

Import

The first step we did was to divide the Overview Sheet into subsheets that contain all data by the province. In order to make sure the importing into R is done correctly, we removed line one til seven so only the Nr of the column is recognized as header. We also outsourced the description of the keys into a separate sheet to prevent a multiple import of the description column. The advantage of this practice is: we can use the total values of the origin sheet in order to test our sums which will be calculated by the province sheets.

Origin datasheet:

```
oSet <- read_excel("DataSets/OriginalDataSet.xlsx", sheet = "Origin")
```

```
## New names:
## * `` -> ...2
## * `` -> ...3
## * `` -> ...4
## * `` -> ...5
## * `` -> ...6
## * ...
```

Baden-Württemberg datasheet:

```
bawSet <- read_excel("DataSets/OriginalDataSet.xlsx", sheet = "BadenWuerttemberg")
```

Bayern datasheet:

```
bavSet <- read_excel("DataSets/OriginalDataSet.xlsx", sheet = "Bayern")
```

Berlin datasheet:

```
berSet <- read_excel("DataSets/OriginalDataSet.xlsx", sheet = "Berlin")
```

Brandenburg datasheet:

```
braSet <- read_excel("DataSets/OriginalDataSet.xlsx", sheet = "Brandenburg")
```

Bremen datasheet:

```
breSet <- read_excel("DataSets/OriginalDataSet.xlsx", sheet = "Bremen")
```

Hamburg datasheet:

```
hamSet <- read_excel("DataSets/OriginalDataSet.xlsx", sheet = "Hamburg")
```

Hessen datasheet:

```
hesSet <- read_excel("DataSets/OriginalDataSet.xlsx", sheet = "Hessen")
```

Mecklenburg-Vorpommern datasheet:

```
mvpSet <- read_excel("DataSets/OriginalDataSet.xlsx", sheet = "MecklenburgVorpommern")
```

Niedersachsen datasheet:

```
ndsSet <- read_excel("DataSets/OriginalDataSet.xlsx", sheet = "Niedersachsen")
```

Nordrhein-Westfalen datasheet:

```
nrwSet <- read_excel("DataSets/OriginalDataSet.xlsx", sheet = "NordrheinWestfalen")
```

Rheinland-Pfalz datasheet:

```
rpfSet <- read_excel("DataSets/OriginalDataSet.xlsx", sheet = "RheinlandPfalz")
```

Saarland datasheet:

```
salSet <- read_excel("DataSets/OriginalDataSet.xlsx", sheet = "Saarland")
```

Sachsen datasheet:

```
saxSet <- read_excel("DataSets/OriginalDataSet.xlsx", sheet = "Sachsen")
```

Sachsen-Anhalt datasheet:

```
saaSet <- read_excel("DataSets/OriginalDataSet.xlsx", sheet = "SachsenAnhalt")
```

Schleswig-Holstein datasheet:

```
swhSet <- read_excel("DataSets/OriginalDataSet.xlsx", sheet = "SchleswigHolstein")
```

Thüringen datasheet:

```
thuSet <- read_excel("DataSets/OriginalDataSet.xlsx", sheet = "Thueringen")
```

Tidy

Transform

Visualise

Model