

Datenanalyse

Jan-Philipp Kolb

08 Mai, 2019

Aufgabe Import von Daten

- Gehe auf das Portal für **offene Daten der Stadt Frankfurt**



Figure 1:

- Importiere den Datensatz für Beschäftigte mit einer geeigneten Funktion.

Die Daten einlesen

```
l1 <- "http://offenedaten.frankfurt.de/dataset/"
l2 <- "50968551-b445-42a9-9288-563784f5768e/resource/"
l3 <- "c5e0a891-42c0-4259-84c7-afa01621477f/download/"
l4 <- "zprojekteopen-datadatenamt-12"
l5 <- "beschäftigtebeschäftigte-2009-2012.xls"

dat <- rio::import(paste0(l1,l2,l3,l4,l5))
```

Einen Überblick über die Daten bekommen

`head(dat)`

Stadtteil	Wirtschaft Betriebe 2009	Wirtschaft Betriebe 2010	Wirtschaft Beschäftigte in Betrieben 2009	Wirtschaft Beschäftigte in Betrieben 2010
Altstadt	708	704	11111	21850
Innenstadt	2733	2825	43455	36682
Bahnhofsviertel	1267	1301	23266	17442
Westend-Süd	3190	3186	47092	48202
Westend-Nord	747	759	3597	3871
Nordend-West	2325	2356	16050	16948
Nordend-Ost	1655	1636	3238	3226

Figure 2:

Einen ersten Eindruck der Daten bekommen

```
library(dplyr)
glimpse(dat)

## Observations: 46
## Variables: 54
## $ Codes
## $ Stadtteil
## $ `Wirtschaft Betriebe 2009`
## $ `Wirtschaft Betriebe 2010`
## $ `Wirtschaft Beschäftigte in Betrieben 2009`
## $ `Wirtschaft Beschäftigte in Betrieben 2010`
## $ `Wirtschaft Beschäftigte in Betrieben im Produzierende
## $ `Wirtschaft Beschäftigte in Betrieben im Produzierende
## $ `Wirtschaft Beschäftigte in Betrieben im Dienstleistun
## $ `Wirtschaft Beschäftigte in Betrieben im Dienstleistun
## $ `Wirtschaft Gewerbeanmeldungen 2011`
## $ `Wirtschaft Gewerbeanmeldungen 2012`
```

Indizieren

Nur die erste Spalte

```
dat[,2]
```

```
## [1] "Altstadt"      "Innenstadt"      "Bahnhofsvierte  
## [5] "Westend-Nord"    "Nordend-West"
```

Gleiches Ergebnis

```
dat$Stadtteil
```

```
## [1] "Altstadt"      "Innenstadt"      "Bahnhofsvierte  
## [5] "Westend-Nord"    "Nordend-West"
```

Eine Beobachtung anschauen

Nur die erste Reihe

```
dat[1,]
```

```
## Codes Stadtteil Wirtschaft Betriebe 2009 Wirtschaft I
## 1 1 Altstadt 708
## Wirtschaft Beschäftigte in Betrieben 2009
## 1 11111
## Wirtschaft Beschäftigte in Betrieben 2010
## 1 21850
## Wirtschaft Beschäftigte in Betrieben im Produzierenden
## 1
## Wirtschaft Beschäftigte in Betrieben im Produzierenden
## 1
## Wirtschaft Beschäftigte in Betrieben im Dienstleistung
## 1
## Wirtschaft Beschäftigte in Betrieben im Dienstleistung
## 1
```

Die Daten zusammenfassen

```
summary(dat[3])
```

```
##  Wirtschaft Betriebe  2009  
##  Min.      : 118.0  
##  1st Qu.: 429.8  
##  Median : 695.0  
##  Mean    : 1810.5  
##  3rd Qu.: 1090.0  
##  Max.    :42126.0
```


Eine Auswahl treffen

Wirtschaft Betriebe 2009 - mehr als 3000

```
dat$Stadtteil[dat[,3]>2000]
```

```
## [1] "Innenstadt"          "Westend-Süd"          "Nordend-Westend"
## [4] "Ostend"              "Bockenheim"           "Sachsenhausen"
## [7] "Frankfurt am Main"
```

Wirtschaft Betriebe 2010 - weniger als 1000

```
dat$Stadtteil[dat[,4]<1000]
```

```
## [1] "Altstadt"          "Westend-Nord"          "Gutleutviertel"
## [4] "Oberrad"          "Niederrad"            "Schwanheim"
## [7] "Griesheim"        "Rödelheim"            "Hausen"
## [10] "Praunheim"        "Heddernheim"          "Niederursel"
## [13] "Ginnheim"         "Dornbusch"            "Eschersheim"
## [16] "Eckenheim"        "Preungesheim"         "Bonames"
```

Mehrere Bedingungen miteinander verknüpfen

Spalte 23 - Wirtschaft Gewerbeabmeldungen 2011

```
dat$Stadtteil[dat[,4]<1000 & dat[,23] > 300]
```

```
## [1] "Gutleutviertel" "Griesheim"      "Fechenheim"      "
```