## Daten beschaffen

Jan-Philipp Kolb

22 Oktober 2018

## **Datenzugang**

- Public-Use-File (PUF) Datei zur öffentlichen Nutzung meist stark anonymisierte Daten (Beispiele: FDZ, Statistik Portal, Meine Region )
- Scientific-Use-File (SUF) Datei zur wissenschaftlichen Nutzung anonymisierte Daten, die zu wissenschaftlichen Zwecken und zur Sekundäranalyse genutzt werden können.
- On-Site-Nutzung Arbeitsplätze für Gastwissenschaftler Kontrollierte Datenfernverarbeitung

#### **EU** initiative INSPIRE



**Targets:** 

#### **Zensus Atlas**

## https://ergebnisse.zensus2011.de/



Jan-Philipp KolbDaten beschaffen22 Oktober 20184 / 15

# A3A Aufgabe: Zensus Ergebnisse für Gemeinden downloaden

- Lade die Zensus Gemeinde Ergebniss hier herunter
- Importiere die Daten mit einer geeigneten Funktion

## Der amtliche Gemeindeschlüssel

03 2 54 021 = Hildesheim

- 03 Niedersachsen
- 2 ehemaliger Regierungsbezirk Hannover
- 54 Landkreis Hildesheim
- 021 Stadt Hildesheim

12 0 64 340 = Neuhardenberg

- 12 Brandenburg
- 0 (in Brandenburg gibt es die Verwaltungseinheit Regierungsbezirk nicht)
- 64 Landkreis Märkisch-Oderland
- 340 Gemeinde Neuhardenberg

6 / 15

#### AGS - Bundesländer



## A3B Aufgabe

- Nutze die Gemeindeergebnisse für den Zensus 2011 und erzeuge einen Datensatz, der nur die Ergebnisse für die Saarländischen Gemeinden enthält.
- Gib die Gemeinde an, in der der Anteil der unter 1-jährigen am höchsten ist.
- Speichere einen Datensatz ab, in dem die folgenden Variablen enthalten sind:
  - der amtliche Gemeindeschlüssel,
  - die Gemeindenamen,
  - die Bevölkerungszahl insgesamt
  - die Zahl der Einjährigen und
  - die Zahl der Zwanzigjährigen

8 / 15

## Forschungsdatenzentren

• Bspw. FDZ der statistischen Ämter:

http://www.forschungsdatenzentrum.de/

- Es werden hauptsächlich Public Use Files angeboten,
- teilweise können Gewichtungsfaktoren verwendet werden um regionale Ergebnisse zu bekommen
- In der Regel ist Darstellung in Karten aber schwierig

## Weitere Amtliche Datenquellen

Genesis

https://www-genesis.destatis.de/genesis/online

• Daneben gibt es Angebote der Landesämter bspw:

https://www.statistik.rlp.de/regionaldaten/

#### **Eurostat Daten**

Sie können eine Statistik der Sparquote bei Eurostat downloaden.

http://ec.europa.eu/eurostat/web/euro-indicators/peeis

```
library(xlsx)
```

HHsr <- read.xlsx2("HHsavingRate.xls",1)</pre>

geo	X2012Q3	X2012Q4	X2013Q1	X2013Q2	X2
Euro area (19 countries)	9.82	11.86	11.37	16.28	10
EU (28 countries)	8.67	10.92	9.42	14.63	8.3
Belgium	12.52	9.33	13.99	19.03	12
Czech Republic	10.16	14.81	9.46	10.44	10

## Datahub.io

- Viele Daten vorhanden,
- bspw. zum UNESCO Weltkulturerbe
- Funktioniert nur im Windows Explorer ?????!!!!

longitude	Large da	
iongitude	iatitude	category_short
20.03	39.75	С
nmad 4.79	35.82	C
3.68	32.48	C
5.74	36.32	C
6.63	35.45	С
1	20.03 nmad 4.79 3.68 5.74	20.03 39.75 mmad 4.79 35.82 3.68 32.48 5.74 36.32

# **American Community Survey**

## **Data for the American Community Survey:**



# The Humanitarian Data Exchange

data.hdx.rwlabs.org



Zum Beispiel Ebola Fälle

Jan-Philipp Kolb Daten beschaffen 22 Oktober 2018 14 / 15

#### **London Datastore**

