

Daten beschaffen und importieren

Jan-Philipp Kolb

11 Januar 2018

EU Initiative INSPIRE



Ziele:

- Räumliche Information zugänglicher und interoperabel machen
- Nachhaltige Entwicklung in diesem Bereich unterstützen

Entwicklung

- Aufgrund der Richtlinie sind mehr Daten frei verfügbar.

- 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

- Die Regionaldatenbank **Genesis**



- Daneben gibt es Angebote der Landesämter bspw das Angebot des **statistischen Landesamts Rheinland-Pfalz**


KREISEBENE



Mein Kreis, meine kreisfreie Stadt

Informationen direkt aus der Datenbank für alle 24 Landkreise und 12 kreisfreien Städte


VERBANDSGEMEINDEEBENE



Meine Verbandsgemeinde

Informationen direkt aus der Datenbank für alle Verbandsgemeinden und verbandsfreie Gemeinden

GEMEINDEEBENE



Mein Dorf, meine Stadt

Informationen direkt aus der Datenbank für alle Städte und Gemeinden in Rheinland-Pfalz

GEMEINDEEBENE



Infrastruktur

Wie sieht es mit der Infrastruktur in den Kommunen des Bundeslandes Rheinland-Pfalz aus? Wo ist die nächste Schule, wo sind die nächsten Einkaufsmöglichkeiten und vieles mehr...

Eurostat Daten

Beispiel: Principal European Economic Indicators

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

```
library(xlsx)
HHsr<-read.xlsx2(paste0(eurostatpath,"HHsavingRate.xls"),1)
```

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

A3A Übung: Download von Eurostat Daten

- Gehe auf die Website mit den *Principal European Economic Indicators* und lade die Statistik der Sparquote **hier** herunter.
- Importiere die Daten in R mit einem geeigneten Befehl.

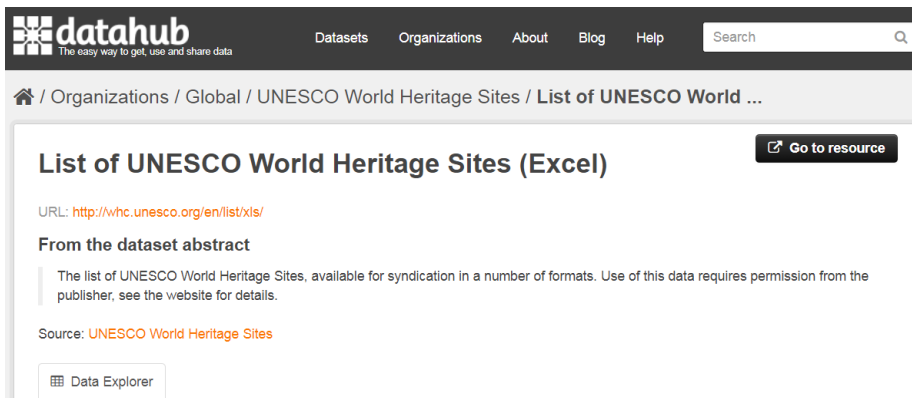
Daten - Institut für ökologische Raumforschung (IÖR)



- Hier gibt es bspw. Indikatoren zu Nachhaltigkeit, Siedlung, Gebäuden, Verkehr etc.
- Es könnte also interessant sein, diese Daten an das Gesis Panel anzuspiesen
- Aber dazu später mehr

Datahub.io

- Auf dieser Plattform sind sehr viele Daten vorhanden,
- bspw. der UNESCO **Weltkulturerbe** Datensatz, den ich in der Folge auch in Beispielen verwenden werde.



The screenshot shows the Datahub.io interface. At the top is a dark navigation bar with the Datahub logo and tagline 'The easy way to get, use and share data', followed by links for Datasets, Organizations, About, Blog, and Help, and a search bar. Below this is a breadcrumb trail: Home / Organizations / Global / UNESCO World Heritage Sites / List of UNESCO World ... The main content area features the title 'List of UNESCO World Heritage Sites (Excel)' with a 'Go to resource' button. Below the title is the URL 'http://whc.unesco.org/en/list/xls/' and a section 'From the dataset abstract' containing a paragraph about syndication and a 'Source: UNESCO World Heritage Sites' link. At the bottom left of the content area is a 'Data Explorer' button.

datahub
The easy way to get, use and share data

Datasets Organizations About Blog Help

Search

Home / Organizations / Global / UNESCO World Heritage Sites / List of UNESCO World ...

List of UNESCO World Heritage Sites (Excel)

[Go to resource](#)

URL: <http://whc.unesco.org/en/list/xls/>

From the dataset abstract

The list of UNESCO World Heritage Sites, available for syndication in a number of formats. Use of this data requires permission from the publisher, see the website for details.

Source: [UNESCO World Heritage Sites](#)

Data Explorer

name_en

longitude latitude category_short

American Community Survey

Die Daten des *American Community Survey*:

American Community Survey (ACS)

About the Survey

Respond to the Survey

News & Updates

Data

Guidance for Data Users

Geography & ACS

Technical
Documentation

Methodology

Library

Operations and
Administration

Contact Us



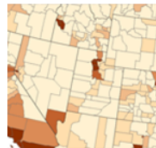
The [American Community Survey](#) (ACS) helps local officials, community leaders, and businesses understand the changes taking place in their communities. It is the premier source for detailed population and housing information about our nation.



How do I respond to the ACS?



Why do you ask each question?




Where can I get ACS data?


The Humanitarian Data Exchange

Zum Beispiel Daten zur Ebola Epedemie

The Humanitarian Data Exchange
WHERE YOUR DATA COMES TO LIFE

FIND DATA

Search HDX Repository 
e.g. preparedness data for Kenya or world health statistics

Choose Country 

1,335
DATASETS

243
COUNTRIES

81
SOURCES

- AG.AGR.TRAC.NO - **Agricultural machinery, tractors**

```
library(WDI)
WDI_dat <- WDI(country="all",
  indicator=c("AG.AGR.TRAC.NO",
    "TM.TAX.TCOM.BC.ZS"),
  start=1990, end=2000)
```

- Es gibt auch eine Funktion `WDIsearch` mit der man nach Indikatoren suchen kann

```
head(WDI_dat)
```

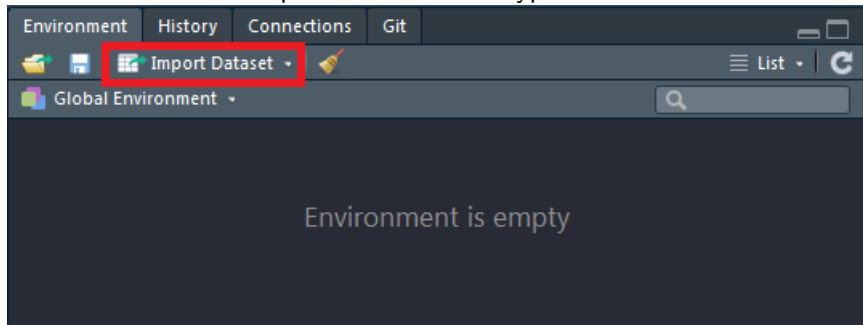
Datenimport



Daten mit RStudio importieren

RStudio Funktionalität um Daten zu importieren

- Environment - Import Dataset - Filetyp auswählen



Wo findet man die Daten?

Browse Button in RStudio

Import Excel Data

File/Url:

Data Preview:

Code Vorschau in Rstudio

Code Preview:

```
library(readxl)
ee_recode_questionnaire_coded <- read_excel("data/ee_recode_questionnaire_coded.xls")
View(ee_recode_questionnaire_coded)
```

Zensus Atlas

Für den Zensus 2011 kann man Daten herunterladen und/oder diese in Karten visualisieren.

<https://ergebnisse.zensus2011.de/>

| Ergebnisse einfach und schnell | | Ergebnisse dynamisch und individuell | | Ergebnisse kartografisch und visuell | |
|---|--|---|--|--|--|
|  | |  | |  | |
| Abruf von Einwohnerzahlen, zusammengefassten Ergebnissen nach Regionen sowie vordefinierten Tabellen und Diagrammen nach Themen | | Erstellung von variablen Tabellen, variablen Diagrammen sowie Gemeinde- und Regionalvergleichen | | Darstellung verschiedener Indikatoren auf Landes-, Kreis- und Gemeindeebene mit der Möglichkeit zur Anpassung der Klassenzahl und Klassengrenzen | |

Figure 1: Zensus Datenbank

Zensus Gemeindeergebnisse

```
setwd(zen_path)
bev_dat <- readxl::read_excel("xlsx_Bevoelkerung.xlsx",3)
DT::datatable(bev_dat[1:1000,1:20])
```

| | 12-stelliger Amtlicher Gemeindeschlüssel (AGS) | Bundesland | Regierungsbezirk/ Ebene NUTS-2 | Kreisfreie Stadt/ Stadtkreis/ Landkreis | Gemeinde- verband | Gemeinde | Gebiet | Regional- ebene | Einwohnerzahl zum 09. Mai 2011 | Bevölkerung nach Geschlecht | X_1 | X_2 |
|----|---|------------|-----------------------------------|--|----------------------|----------|------------------------|--------------------|--------------------------------------|-----------------------------------|------|------|
| 31 | 010515163037 | 01 | 0 | 51 | 5163 | 037 | Großenrade | Gemeinde | 493 | 491 | 238 | 253 |
| 32 | 010515163051 | 01 | 0 | 51 | 5163 | 051 | Hochdonn | Gemeinde | 1193 | 1193 | 572 | 621 |
| 33 | 010515163064 | 01 | 0 | 51 | 5163 | 064 | Kuden | Gemeinde | 637 | 637 | 337 | 300 |
| 34 | 010515163089 | 01 | 0 | 51 | 5163 | 089 | Quickborn | Gemeinde | 201 | 200 | 99 | 101 |
| 35 | 010515163097 | 01 | 0 | 51 | 5163 | 097 | Sankt Michaelisdonn | Gemeinde | 3618 | 3618 | 1790 | 1828 |

A3A Aufgabe: Zensus Ergebnisse für Gemeinden downloaden

- Lade die Zensus Gemeinde Ergebnisse **hier** herunter.
- Importiere die Daten mit einer geeigneten Funktion in R.
- Welche Information ist in den Daten enthalten?

Der amtliche Gemeindeschlüssel

03 2 54 021 = Hildesheim

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

12 0 64 340 = Neuhausen

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

AGS - Bundesländer

| # ↕ | Land ↕ |
|-----|------------------------------|
| 01 | Schleswig-Holstein |
| 02 | Freie und Hansestadt Hamburg |
| 03 | Niedersachsen |
| 04 | Freie Hansestadt Bremen |
| 05 | Nordrhein-Westfalen |
| 06 | Hessen |
| 07 | Rheinland-Pfalz |
| 08 | Baden-Württemberg |
| 09 | Freistaat Bayern |
| 10 | Saarland |
| 11 | Berlin |
| 12 | Brandenburg |
| 13 | Mecklenburg-Vorpommern |
| 14 | Freistaat Sachsen |

A3B Aufgabe

- Nutze die Gemeindeergebnisse für den Zensus 2011 und erzeuge einen Datensatz, der nur die Ergebnisse für die Saarländischen Gemeinden enthält.
- Ermittle aus diesem Datensatz die Gemeinde 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