

Nutzung von GeoDaten in den Sozialwissenschaften - Daten verbinden

Jan-Philipp Kolb

08 April 2016

Die wrld_simpl Daten

```
library("maptools")  
data("wrld_simpl")
```

Datensatz in die Daten-Registerkarte von Rstudio schreiben

```
df_ws <- data.frame(wrld_simpl@data)
```

Rstudio Data Browser

Jetzt können wir den Rstudio Daten-Browser verwenden

	FIPS	ISO2	ISO3	UN	NAME	AREA	POP2005	REGION	SUBREGION	LON	LAT	co2_90
ATG	AC	AC	ATC	28	Antigua and Barbuda	44	83039	19	29	-61.783	17.078	4.9
DZA	AG	DZ	DZA	12	Algeria	238174	32854159	2	15	2.632	28.163	3.1
AZE	AJ	AZ	AZE	31	Azerbaijan	8260	8352021	142	145	47.395	40.430	-
ALB	AL	AL	ALB	8	Albania	2740	3153731	150	39	20.068	41.143	2.3
ARM	AM	AM	ARM	51	Armenia	2820	3017661	142	145	44.563	40.534	-
AGO	AO	AO	ACO	24	Angola	124670	16095214	2	17	17.544	-12.296	0.4
ASM	AQ	AS	ASM	16	American Samoa	20	64051	9	61	-170.730	-14.318	NA
ARG	AR	AR	ARC	32	Argentina	273669	38747148	19	5	-65.167	-35.377	3.5
AUS	AS	AU	AUS	36	Australia	768230	20310208	9	53	136.189	-24.973	17.2

Der wrld_simpl Datensatz

Holen Sie sich die Ländernamen:

```
CNames <- wrld_simpl@data$NAME  
head(CNames)
```

```
## [1] Antigua and Barbuda Algeria Azerbaijan  
## [4] Albania Armenia Angola  
## 246 Levels: Aaland Islands Afghanistan Albania Algeria
```

```
CNames <- as.character(CNames)  
head(CNames)
```

```
## [1] "Antigua and Barbuda" "Algeria" "Azerbaijan"  
## [4] "Albania" "Armenia" "Angola"
```

Substring eines Zeichen-Vektors

```
CNames1 <- substr(CNames,1,1)  
head(CNames1)
```

```
## [1] "A" "A" "A" "A" "A" "A"
```

```
CNames2 <- substr(CNames,1,2)  
head(CNames2)
```

```
## [1] "An" "Al" "Az" "Al" "Ar" "An"
```

Auswahl vornehmen

```
CNames[CNames2=="An"]
```

```
## [1] "Antigua and Barbuda" "Angola" "Anguilla"  
## [4] "Andorra" "Antarctica"
```

CO2 Emissionen

```
(load("data/CO2emissions.RData"))
```

```
## [1] "co2"
```

Rank	Country	j1990	j1991	j1992	j1993	j1994	j1995
1.	Qatar	25.2	36.7	54.3	60.9	58.7	55.1
2.	Trinidad and Tobago	13.9	17.1	17.0	13.5	15.8	16.1
3.	Netherlands Antilles	32.6	26.9	22.6	35.0	34.3	33.1
4.	Kuwait	19.0	5.1	10.0	16.9	20.8	21.1

Wir müssen Länder in diesem Datensatz und Ländernamen in wrld_simpl-Datensatz zusammenbringen

Vektoren zum Matching

Wie bringt man zwei Vektoren zusammen:

```
A <- c(1,2,3,4)
B <- c(4,3)
match(A,B)
```

```
## [1] NA NA 2 1
```

```
match(B,A)
```

```
## [1] 4 3
```


Vektoren zum Matching

```
D <- c(1,3,5,6,7)
E <- c("A",1,98,4)
match(D,E)
```

```
## [1]  2 NA NA NA NA
```

- ▶ Matching mit Ländernamen, um eine Karte mit CO2 Indikatoren zu produzieren

Matching

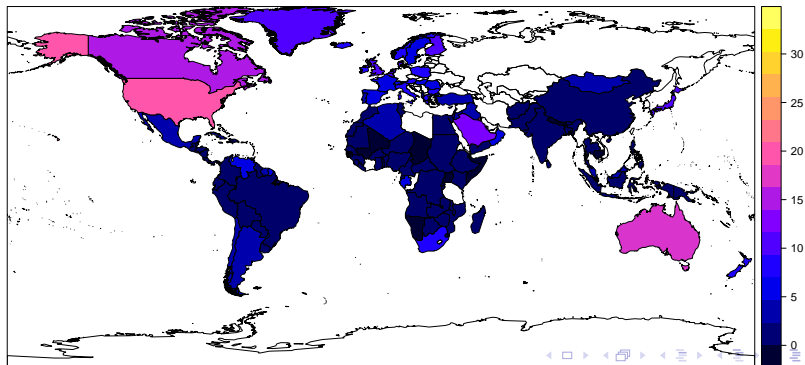
```
ind <- match(wrld_simpl@data$NAME,co2$Country)
ind
```

[illegible]

Daten anspielen

```
wrld_simpl@data$co2_90 <- co2$j1990[ind]
```

```
library(sp)  
spplot(wrld_simpl,"co2_90")
```



Zusätzliche Länder matchen

```
ind2 <- match(co2$Country, wrld_simpl@data$NAME)
fehlt <- co2$Country[is.na(ind2)]
fehlt
```

```
## [1] Brunei
## [2] United Arab Emirates
## [3] Falkland Islands
## [4] South Korea
## [5] Taiwan[4] [5]
## [6] Libya
## [7] European Union
## [8] Iran
## [9] Macedonia
## [10] World
## [11] Réunion
## [12] Syria
## [13] North Korea
## [14] Saint Helena, Ascension and Tristan da Cunha
```

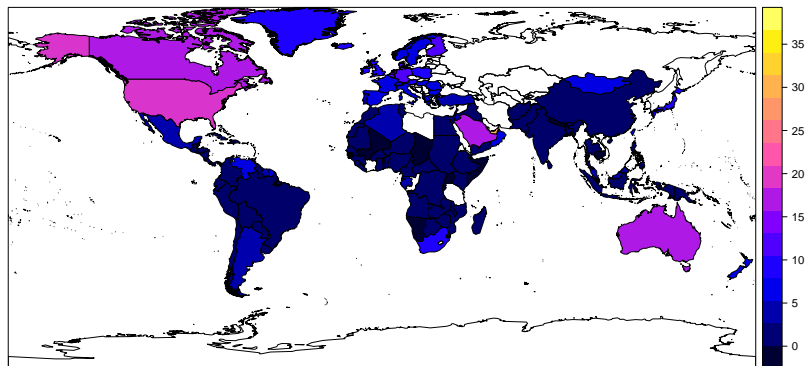
Matching mit agrep

```
Namen_ws <- as.character(wrld_simpl@data$NAME)
Namen_co2 <- as.character(co2$Country)
for (i in 1:length(ind)){
  if(is.na(ind[i])){
    ind4 <- agrep(Namen_ws[i],Namen_co2)
    if(length(ind4)==1){
      ind[i] <- ind4
    }
  }
}
```

Daten anspielen

```
wrld_simpl@data$co2_91 <- co2$j1991[ind]
```

```
spplot(wrld_simpl, "co2_91")
```



Matching mit Eurostat Daten

Sie können eine Statistik der Sparquote bei Eurostat downloaden.
link

Der wrld_simpl Datensatz

Wir brauchen nur den europäischen Kontinent

```
EUR <- wrld_simpl[wrld_simpl$REGION==150,]  
plot(EUR)
```



Europa ohne Russland

- Russland ist zu groß für unsere Karte :

```
EUR <- EUR[-which(EUR@data$NAME=="Russia"),]  
plot(EUR)
```



Matching Staaten

```
as.character(HHsr[5:10,1])
```

```
EUR_Names <- as.character(EUR@data$NAME)  
head(EUR_Names)
```

## [1]	"Albania"	"Bosnia and Herzegovina"
## [3]	"Bulgaria"	"Denmark"
## [5]	"Ireland"	"Estonia"

```
ind <- match(EUR_Names,HHsr[,1])  
head(ind)  
table(is.na(ind))
```

Die Daten an Polygondaten anfügen

```
EUR@data$HHsr <- as.numeric(as.character(HHsr[ind,2]))
```

```
# Plot the map:  
library(sp)  
spplot(EUR, "HHsr")
```