

Data Import

Jan-Philipp Kolb

Wed Oct 14 07:54:24 2015

Import csv-files

You can download a statistic on dead refugees from
unitedagainstracism.org

Link datahub.io

```
data.path <- "C:/Work/data/"  
setwd(data.path)  
lod <- read.csv("listofdeath.csv")
```

- ▶ use read.csv2 for German data

Read csv-files directly from the web

```
link <- "https://raw.githubusercontent.com/  
mihi-tr/list-of-deaths-dpkg/master/data/  
listofdeath.csv"  
lod <- read.csv(link)
```

Show the data

date	count	name	origin
25/10/12	54	N.N.	unknown
23/10/12	1	N.N. ($\hat{A} \pm 25$, man)	Sub-Saharan Africa
09/10/12	1	N.N. ($\hat{A} \pm 5$, girl)	Sub-Saharan Africa
09/10/12	1	N.N. ($\hat{A} \pm 5$, girl)	Sub-Saharan Africa
08/10/12	6	N.N.	Africa
08/10/12	10	N.N.	Africa

```
colnames(lod)
```

```
## [1] "date"          "count"         "name"
## [5] "cause.of.death" "source"        "Column.7"
## [9] "Column.9"      "Column.10"    "Column.11"
```

Import Excel-files

Unesco world heritage sites: [Link datahub.io](https://datahub.io)

- ▶ You have to provide a sheet index as second argument
- ▶ For German data use `read.xlsx2` because of the comma separation

```
library(xlsx)
whc <- read.xlsx("whc-sites-2015.xls",1)
```

Overview of World Heritage Dataset

```
table(whc$category)
```

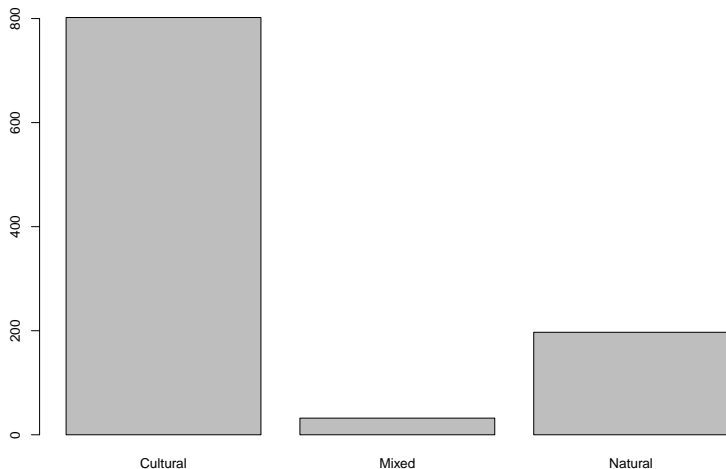
```
##  
## Cultural      Mixed    Natural  
##      802         32      197
```

```
colnames(whc)
```

```
## [1] "unique_number"      "id_no"      "rev_1"  
## [4] "name_en"            "name_fr"    "short_name"  
## [7] "short_description_fr" "justification_en" "justification_fr"  
## [10] "date_inscribed"     "secondary_dates" "danger_status"  
## [13] "longitude"          "latitude"    "area_sqkm"  
## [16] "C1"                 "C2"          "C3"  
## [19] "C4"                 "C5"          "C6"  
## [22] "N7"                 "N8"          "N9"  
## [25] "N10"                "criteria_txt" "category"  
## [28] "category short"     "states name en" "states name fr"
```

Create a barplot

```
barplot(table(whc$category))
```



Import sav-files

```
library(foreign)
mz10 <- read.spss("mz2010_cf.sav",to.data.frame=T)
```


Command read.spss

read.spss {foreign}

R Documentation

Read an SPSS Data File

Description

`read.spss` reads a file stored by the SPSS `save` or `export` commands.

This was originally written in 2000 and has limited support for changes in SPSS formats since (which have not been many).

Usage

```
read.spss(file, use.value.labels = TRUE, to.data.frame = FALSE,  
          max.value.labels = Inf, trim.factor.names = FALSE,  
          trim_values = TRUE, reencode = NA, use.missings = to.data.frame)
```

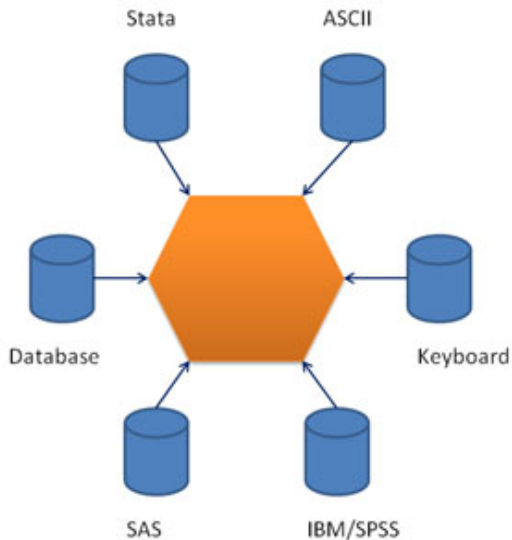
Read stata-files

- ▶ Datset about students (2000) from Forschungsdatenzentrum
- ▶ Datset has 206.867 entries

```
studis <- read.dta("studenten_cf_2000.dta")
```

- ▶ Blog post about reading and writing stata files on is.R

Quick R on data import



<http://www.statmethods.net/input/index.html>

Data management a bit similar to SPSS

```
install.packages("Rz")  
library(Rz)
```