### Daten laden und speichern

Dirk Seidensticker/Clemens Schmid

6. Februar 2016

```
# nicht lauffähig
df <- read.csv("../data/AtlantPottery.csv")</pre>
```

#### Fehlermeldung:

```
Error in read.table(
  file = file,
  header = header,
  sep = sep,
  quote = quote, :
  doppelte 'row.names' sind nicht zulässig
```

```
df <- read.csv("../data/AtlantPottery.csv", sep = '\t')</pre>
```

### head(df)

	site	featur	re ol	oject c	lass	sherd	qty	wt	size	wall	muendungsD
1	C	surfac	ce -1	:3,13	K	G	1	1448	500	7	17.5
2	C	surfac	ce	-1:1	K	R	1	56	120	5	17.0
3	C	surfac	ce -:	1:2,5	K	R	1	92	120	9	20.0
4	C	surfac	ce	2	K	R	1	45	70	11	22.5
5	E	surfac	ce	1	K	G	1	1298	500	8	18.5
6	E	surfac	ce		K	G	1	667	200	NA	6.0
	muend	dungsH	${\tt minD}$	minD_H	maxD	maxD_	H b	odenD	tempe	erSize	${\tt vesselShape}$
1		19	16.0	16.5	21.0	11.	0	7.5		M	3
2		NA	14.5	NA	16.0	1	IA	NA		М	7
3		NA	18.5	NA	24.5		IA	NA		VF	7
4		NA	NA	NA	27.0	1	IA	NA		VF	8
5		14	22.5	9.5	22.5	9.	5	0.0		VC	8
6		14	5.5	11.5	10.5	5 5	5	6.0		C	1
	1 2 3 4 5 6 1 2 3 4 5 6	1 C 2 C 3 C 4 C 5 E 6 E muenc 1 2 3 4 5	1 C surface 2 C surface 3 C surface 4 C surface 5 E surface 6 E surface muendungsH 1 19 2 NA 3 NA 4 NA 5 14	1 C surface -1 2 C surface 3 C surface 4 C surface 5 E surface 6 E surface muendungsH minD 1 19 16.0 2 NA 14.5 3 NA 18.5 4 NA NA 5 14 22.5	1 C surface -1:3,13 2 C surface -1:1 3 C surface -1:2,5 4 C surface 2 5 E surface 1 6 E surface	1 C surface -1:3,13 K 2 C surface -1:1 K 3 C surface -1:2,5 K 4 C surface 2 K 5 E surface 1 K 6 E surface K muendungsH minD minD_H maxD 1 19 16.0 16.5 21.0 2 NA 14.5 NA 16.0 3 NA 18.5 NA 24.5 4 NA NA NA NA 27.0 5 14 22.5 9.5 22.5	1 C surface -1:3,13 K G 2 C surface -1:1 K R 3 C surface -1:2,5 K R 4 C surface 2 K R 5 E surface 1 K G 6 E surface K G muendungsH minD minD_H maxD maxD_1 1 19 16.0 16.5 21.0 11.2 2 NA 14.5 NA 16.0 M 3 NA 18.5 NA 24.5 M 4 NA NA NA 27.0 M 5 14 22.5 9.5 22.5 9.5	1 C surface -1:3,13 K G 1 2 C surface -1:1 K R 1 3 C surface -1:2,5 K R 1 4 C surface 2 K R 1 5 E surface 1 K G 1 6 E surface K G 1 muendungsH minD minD_H maxD maxD_H b 1 19 16.0 16.5 21.0 11.0 2 NA 14.5 NA 16.0 NA 3 NA 18.5 NA 24.5 NA 4 NA NA NA 27.0 NA 5 14 22.5 9.5 22.5 9.5	1 C surface -1:3,13 K G 1 1448 2 C surface -1:1 K R 1 56 3 C surface -1:2,5 K R 1 92 4 C surface 2 K R 1 45 5 E surface 1 K G 1 1298 6 E surface K G 1 667 muendungsH minD minD_H maxD maxD_H bodenD 1 19 16.0 16.5 21.0 11.0 7.5 2 NA 14.5 NA 16.0 NA NA 3 NA 18.5 NA 24.5 NA NA 4 NA NA NA 27.0 NA NA 5 14 22.5 9.5 22.5 9.5 0.0	1 C surface -1:3,13 K G 1 1448 500 2 C surface -1:1 K R 1 56 120 3 C surface -1:2,5 K R 1 92 120 4 C surface 2 K R 1 45 70 5 E surface 1 K G 1 1298 500 6 E surface K G 1 667 200 6 muendungsH minD minD_H maxD maxD_H bodenD temper 1 19 16.0 16.5 21.0 11.0 7.5 2 NA 14.5 NA 16.0 NA NA 3 NA 18.5 NA 24.5 NA NA 4 NA NA NA 27.0 NA NA 5 14 22.5 9.5 22.5 9.5 0.0	2 C surface -1:1 K R 1 56 120 5 3 C surface -1:2,5 K R 1 92 120 9 4 C surface 2 K R 1 45 70 11 5 E surface 1 K G 1 1298 500 8 6 E surface K G 1 667 200 NA muendungsH minD minD_H maxD maxD_H bodenD temperSize 1 19 16.0 16.5 21.0 11.0 7.5 M 2 NA 14.5 NA 16.0 NA NA MA 3 NA 18.5 NA 24.5 NA NA VF 4 NA NA NA 27.0 NA NA VF 5 14 22.5 9.5 22.5 9.5 0.0 VC

# **Andere Datenquellen**

#### Datenbanken und XML

- SQLite: library(RSQLite)
- PostGreSQL: library(RPostgreSQL)
- MySQL, Oracle, ODBC (library(RODBC)), MongoDB, . . .
- XML, JSON

```
# nicht lauffähig
library(RSQLite)
drv <- dbDriver("SQLite")
con <- dbConnect(drv, "../data/DB.sqlite")

df = dbGetQuery(con, "SELECT * FROM ...")
head(df)</pre>
```