

# 1) DATA SETS

- > Dataset: variables are in columns, observations are in rows, and values are in cells.
- > Each variable must have its own **column** / Each observation must have its own **row**.  
Each value must have its own **cell**

country	year	cases	population
Afghanistan	1999	745	19987071
Afghanistan	2000	2666	20595360
Brazil	1999	37737	172006362
Brazil	2000	80488	174504898
China	1999	212258	1272915272
China	2000	213766	1280425583

variables

country	year	cases	population
Afghanistan	1999	745	19987071
Afghanistan	2000	2666	20595360
Brazil	1999	37737	172006362
Brazil	2000	80488	174504898
China	1999	212258	1272915272
China	2000	213766	1280425583

observations

country	year	cases	population
Afghanistan	1999	745	19987071
Afghanistan	2000	2666	20595360
Brazil	1999	37737	172006362
Brazil	2000	80488	174504898
China	1999	212258	1272915272
China	2000	213766	1280425583

values

# 1) DATA SETS

country	year	cases	population
Afghanistan	1999	745	19987071
Afghanistan	2000	2666	20595360
Brazil	1999	37737	172006362
Brazil	2000	80488	174504898
China	1999	212258	1272915272
China	2000	216766	1280425583

variables

country	year	cases	population
Afghanistan	1999	745	19987071
Afghanistan	2000	2666	20595360
Brazil	1999	37737	172006362
Brazil	2000	80488	174504898
China	1999	212258	1272915272
China	2000	216766	1280425583

observations

country	year	cases	population
Afghanistan	1999	745	19987071
Afghanistan	2000	2666	20595360
Brazil	1999	37737	172006362
Brazil	2000	80488	174504898
China	1999	212258	1272915272
China	2000	216766	1280425583

values

Source: Golemund & Wickham (2017)

Tabelle (Relation)

**Kundendaten**

Kd-Nr	Nname	Vname	Strasse	Plz	Ort
0001	Schneider	Thomas	Hauptstr. 2	50969	Köln
0002	Maier	Wolfgang	Gartenstr. 1	77933	Lahr

Primärschlüssel

Datenfeld

Attribut

Datensatz  
(Tupel)