

Errata BIBB-Übergangsstudie 2011 September 2023

September 2023: Korrektur für eine Variable der BIBB-Übergangsstudie 2011:

Die Variable *hbild* "Höchster allgemeinbildender Schulabschluss" im Querschnittsdatensatz ist fehlerhaft. Die Variable wurde mit Hilfe des Stata-Befehls `egen max(v2113)` aus den Angaben zum Schulabschluss (v2113) des Schulspells der Ereignisdaten gebildet. Dies führte zu einem Fehler da die Kategorien 9 "Kein Abschluss", 97 „verweigert“ und 98 „weiss nicht“ höher codiert sind als die gültigen Angaben. Somit wurde für manche Personen nicht die richtige Information in die Variable *hbild* übernommen.

Der SUF wurde nicht geändert. Die Korrektur muss von den Nutzenden mit Hilfe der folgenden Befehle im Ereignisdatensatz selbst vorgenommen und ggf. an den Querschnittsdatensatz angehängt werden.

```
Stata: bysort dataser: egen hbild=max(v2113) if v2113<9
      replace hbild=9 if v2113==9 & hbild==.
      lab val v2113 V2113
```

```
SPSS: SORT CASES dataser v2113 (d).
      COMPUTE hbild=v2113.
      if v2113 gt 96 hbild=$Sysmis.
      if (lag(hbild) lt 9 and lag(hbild) gt -1 and lag(hbild) gt v2113 and dataser eq
      lag(dataser)) hbild=lag(hbild).
      EXECUTE.
      SORT CASES dataser episode.
      apply dictionary from * / source var = v2113 / target var = hbild/ varinfo
      varlabels vallabels.
      EXECUTE.
```

R: Option A: Factor Vector (Empfehlung)

```
#Benötigte Pakete öffnen bzw. Installieren
if (!require("pacman")) install.packages("pacman")
library(pacman)
p_load(char=c("haven", "labelled", "sjlabelled", "forcats"), install = T)

#Datensatz im Objekt dat speichern
```

```

dat <- read_dta("BIBB_Uebergangstudie2011_Ereignisdaten_SUF.dta")
dat <- labelled::to_factor(dat, levels="1")
dat1 <- read_dta("BIBB_Uebergangstudie2011_Querschnitt_SUF.dta")
dat1 <- labelled::to_factor(dat1, levels="1")

# Variable hbuild bilden
dat$zv2113_ <- fct_shift(dat$zv2113, n=-3)
dat$zv2113_ <- fct_relevel(dat$zv2113_, "9 ohne Abschluss", after = 3)
hbuild <- tapply(as.numeric(dat$zv2113_), dat$datsaser, FUN= function(x){max(x,
na.rm=T)})
hbuild <- as_factor(hbuild)
levels(hbuild) <- levels(dat$zv2113_)
hbuild<- fct_shift(hbuild, n=2)
hbuild <- fct_relevel(hbuild, "9 ohne Abschluss", after = 9)
#Variable dem Datensatz zuspielden
dat1$hbuild <- hbuild

```

Option B: Labelled Data

```

#Benötigte Pakete öffnen bzw. Installieren
if (!require("pacman")) install.packages("pacman")
library(pacman)
p_load(char=c("haven", "labelled", "sjlabelled"), install = T)

#Datensätze einlesen
dat <- read_dta("BIBB_Uebergangstudie2011_Ereignisdaten_SUF.dta")
dat1 <- read_dta("BIBB_Uebergangstudie2011_Querschnitt_SUF.dta")

#Variable hbuild bilden
dat$zv2113_ <- dat$zv2113
dat$zv2113_ [dat$zv2113_ ==9] <- 0
dat$zv2113_ [dat$zv2113_ ==98] <- -98
dat$zv2113_ [dat$zv2113_ ==97] <- -97
hbuild <- labelled(tapply(dat$zv2113_ , dat$datsaser, FUN= function(x){max(x,
na.rm=T)}))
hbuild[hbuild==0] <- 9
hbuild[hbuild==-98] <- 98
hbuild[hbuild==-97] <- 97
hbuild <- set_labels(hbuild, labels =get_labels(dat$zv2113, values="n"))
#Variable dem Querschnittedatensatz zuspielden
dat1$hbuild <- hbuild

```