Essential STATA commands for SOEPcampus Workshops:

To benefit as much as possible from the upcoming SOEPcampus workshop please make sure you are familiar with the following STATA commands:

Working with the „Do-file Editor"

All your work should be script-based, using STATA's "do-Files", such that all your work can easily be replicated by running the respective script.
- Access the editor, access do-Files.
- Run entire do-Files and run parts of do-Files.

Localizing the data on your computer or the network respectively

STATA data-files (.dta) which you find in your Explorer or Finder can also be accessed directly from a STATA script using the "use" command.

```
use D:/temp/apausl, clear
use "\\hume\rdc-gen\apausl", clear
save D:/temp/apausl_test, replace
```
- What do you need to consider when your Folder Names contain spaces or special characters?
- What is the difference between / and \ in pathnames?

Macros

You should be familiar with the use of "locals" and "globals" in STATA.

```
global pfad "D:/temp/"
use "${pfad}apausl", clear

local uncpfad \\hume\rdc-gen\
use `uncpfad'apausl, clear
```

- What is the difference between locals and globals?
- How can you create the signs that surround locals?

If conditions and logical operators in STATA

You sould be familiar with conditioning commands with the if extension and know STATA's logical operators

```
sum income if year == 2007 & age >= 45 & age <= 46 & health != .
tab xnetto xpop if (xnetto==10 | xnetto==12) & (xpop==1 | xpop==2)
```

- How are "and" and "or" operators combines correctly?

list

```
list sex gebjahr todjahr if persnr==2102 | persnr==19202
```
- What does this command do? Would you use it?
- With which other command should list always be combined?

keep/drop

You should know the effects oft he commends "keep" and "drop".

```
keep if age >= 18 & age <=65
drop wp01742
```

rename

You should know how to rename variables.
```
rename yp10601 smoke2008
```

recode

You should know how to recode variables and how to chance label values.
```
recode age_3 (17/29=1) (30/64=2) (65/120=3)
label define age_3 1 "17-29" 2 "30-64" 3 "65+"
label values age_3 age_3
```

gen / egen / replace

You should be able to generate new variables using the "gen" and "egen" commands.
```
gen alt2003_15 = 1 if  2003-gebjahr == 15
```
- What is the difference between the gen and the egen command? When should which command be used?
- What does the "replace" command do?

tab / tabstat / sum

You should know the "tab", "tabstat" and "sum" commands.
```
tabstat yp0101,  by(sex)
tab yp0101 sex
tab diff if year==2007
sum health if smokestop==1
```
- What is the difference between tab and tabstat?

merge

You should know how the "merge" command works and how to interpret the "1:1" part of the command. You sould also be able to interpret the merge-results.
```
merge 1:1 syear pid using "${MY_IN_PATH_long}\pgen.dta", keepus(pgemplst) nogen
```
- What does _merge == 2 mean?
- What is the effect of specifying the options "keep" and "nogen"?
- Which command should follow when using "m:n"?

foreach/forvalues

It would also be useful to be familiar with STATA's different looping options "foreach" and "forvalues".
```
local y=2006
    foreach w in w x y {
        rename `w'p0101 health`y'
        rename `w'phrf phrf`y'
        local y=`y'+1
}
```
- When do you use foreach, and when forvalues?
- How is the loop-interval determined in forvalues?

mvdecode

This command is very useful for recoding missing values.
```
mvdecode _all, mv(-1=. \ -2=.t \ -3=.x)
```

Log-Files

Some people like working with log-files, which record the outcomes of the analyses.
```
capture log close
log using "${MY_LOG_FILE}", text replace
```
- Where in your do-File would youl place the command "capture log close"?
- What are log-files useful for?