8. Summary

Karl B Christensen
http://publicifsv.sund.ku.dk/~kach/SPSS

- Syntax documents all the steps
- Point-and-click and press 'Paste'

- keep original data sets without changing them

- Remember model control

- Results from wrong model are useless
- If you transform a variable (e.g. $y \mapsto \log(y)$) it changes the interpretation of effects (e.g. relative effect).

Remember

1. SPSS can make many (MANY!) calculations.
2. You have to decide what you to calculate
3. Sometimes this is easier to think about without sitting in front of a computer

SPSS syntax ends up a mixture:

- Things you planned (median score in two groups, test comparing groups, logistic regression model, ..)
- Things you have do do for technical reasons (`SPLIT FILE`, recode variables, change variable type, ..)

Edit syntax into coherent `*.sps` file. This is your documentation!

Example: Excel sheet data from SF36 items about physical functioning, information about gender and age.

1. read data
2. code scale score
3. look at histograms
   - normal distribution: compute mean (SD) in each group, compare using *t*-test
   - not normal distribution: compute median (IQR) in each group, compare using Wilcoxon test
4. find cut-point (lowest 20%)
5. use logistic regression to see if gender and age are associated with risk of being in the group with lowest score.

```
* Read excel data set.
GET DATA /TYPE=XLSX
  /FILE='P:\public_html\PF.xlsx'
  /SHEET=name 'PF'
  /CELLRANGE=full
  /READNAMES=on.
EXECUTE.

* Compute scale score. Define Variable Properties.
COMPUTE pf=SUM(pf01,pf02,pf03,pf04,pf05,pf06,pf07,pf08,pf09,pf10).
VARIABLE LEVEL  pf(SCALE).
EXECUTE.

* Visual inspection of distribution.
GRAPH
  /HISTOGRAM(NORMAL)=pf
  /PANEL COLVAR=gender COLOP=CROSS.
```

```
* Descriptive statistics.
SORT CASES  BY gender.
SPLIT FILE LAYERED BY gender.
EXECUTE.
FREQUENCIES VARIABLES=pf01
  /FORMAT=NOTABLE
  /NTILES=4
  /ORDER=ANALYSIS.
SPLIT FILE OFF.

*Nonparametric Tests: Independent Samples.
NPTESTS
  /INDEPENDENT TEST (pf)  GROUP (gender) MANN_WHITNEY
  /CRITERIA ALPHA=0.05.
```

```
* lowest value group.
FREQUENCIES VARIABLES=pf
  /FORMAT=NOTABLE
  /NTILES=5
  /ORDER=ANALYSIS.

* code indicator.
RECODE PF (Lowest thru 6=1) (ELSE=0) INTO low.
VARIABLE LABELS  low 'low PF value'.
EXECUTE.

* regression model.
LOGISTIC REGRESSION VARIABLES low
  /METHOD=ENTER age gender
  /CRITERIA=PIN(.05) POUT(.10) ITERATE(20) CUT(.5).
```