# Requirements on the final Submission

*Steffen Wagner u. Sören Pannier*

*16 January 2018*

## Objective

In your final submission you demonstrate that you can present and implement a statistical method. You do this by formally introducing and discussing one(!) of the suggested statistical methods. Then you implement this method using R. The implementation is contained in an R package and uses a substantial amount of techniques from the lecture. This includes functions, S3-Object-Orientation, visualisations, and control structures. Hence your final submission contains two parts: First, a vignette containing the statistical methodology and a guide on how to use your package and second the R code implementing the method. The requirements for both are presented below.

## Structure

### Vignette (40 %)

- The aim of the vignette is
    - to introduce the methodology you implement, (10 %)
    - to show *how* (10 %)
    - and *that* your implementations works (10 %)
        * if you do not compare your results with a reference from cran we will assume your code is returning wrong results.
- Here you also demonstrate that you are acustome to the relevant literature and are able to formally discuss and introduce a statistical method
- Stick to the standard structure of a research paper (10 %)
    - as an example combining software and statistical theory see https://www.jstatsoft.org/article/view/v055i13
    - for formal requirements see http://www.wiwiss.fu-berlin.de/studium-lehre/bachelor/vwl/bachelorarbeit/VWL_Leitfaden_2012.pdf
    - spelling
    - language

### Code (60 %)

- Every exported function has a complete help page, explaining all arguments, their effects and the return values
- Important functions should also contain an example in the help file
- The package ist checked using `R CMD check`. Any notes, warnings and errors need to be discussed in your vignette. If any exist they may affect your final grade
- Comment your functions if (1) they do not have a documentation or (2) their purpose is not self explanatory
- The implementation of standard methods like summary, plot, and print is mandatory
- Grading of your code:
    - implementation of the method (20 %)
    - use of S3 System (20 %)
    - R cmd check (10 %)
    - usability of the exported functions (5 %)
    - form (5 %)

- Solutions not containing an own implementation of the methodology are not sufficient to pass
  - if you realize you struggle implementing the methodology contact us and we will try to find an individual solution

# Formal requirements

- You submit your work before tuesday 8am the 02.04.2019 to soeren.pannier@fu-berlin.de
- Your submission is written in englisch or german (and R)
- You follow the style guide throughout your code (see lecture slides)
- The vignette contains at maximum 3000 words
- You follow the author-year (Harvard) citation style
- You submit 1 package with 1 vignette implementing 1 method of your choice

# Possible methods

- Logistic regression Models
- k-means clustering
  - if you choose k-means clustering it must be possible to define starting points in the function interface.