Unit 1 - Exercise 1

Repetition - Create report for data set struma

Data set - struma

A goitre, or goiter, is a swelling in the neck resulting from an enlarged thyroid gland.
[in German: goiter = Struma; goitre = Kropf].

In patients with goiter, a clinical study was conducted to investigate whether medication in addition to standard therapy (MED; yes/no) has an effect on thyroid hormones FT3, FT4 and TSH.

Import data set

Aim: Create Rdata file with prepared struma data set

- Open *0_struma_import_vYYYYMMDD.R* from folder 'R/Rfiles'
- Execute the code and try to understand what happens
  - at the end there should be *struma_prepared_vYYYYMMDD.Rdata* in folder 'R/Rdata'
- Look at the created data set 'dt_struma' typing the following commands in the console

```
dt_struma
view(dt_struma)
str(dt_struma)
```

Create report

Aim: Create a report with descriptive statistics for struma data set

Task 0

- Open *UNIT1_ex1_struma_vYYYYMMDD.Rmd* from folder 'R/Rmarkdown'
  - adapt author statement in the Rmd file
- The chunks underneath the headings 'List of parameters' and 'Information about parameters' are just there to give you an idea how to get an overview of the parameters

Task 1 - Descriptive statistics

a) Generate a table with descriptive statistics for all parameters separated by medication (MED) except for patient ID

- look at homepage of gtsummary: https://www.danieldsjoberg.com/gtsummary/
- tbl_summary()
  - look at arguments: 'included', 'by'
  - hint for 'included': use negative sign ,e.g., '-c(test)'
- Optional

  - make the parameter labels bold
  - add the spanning header ‘Medication’
  - change for continuous parameters the statistic from ‘median (IQR)’ to ‘mean (SD)’

b) Generate for parameters ‘gender’ and ‘age’ descriptive statistics and corresponding plot(s)

- a table with descriptive statistics
- plot for ‘gender’
  - bar charts (hint: *geom_bar()*)
- plots for ‘age’
  - box plot (hint: *geom_boxplot()*)
  - histogram (hint: *geom_histogram()*)
  - qq-plot (hint: *geom_qq()*, *geom_qq_line()*)
- general hints
  - *facet_grid()*
  - use *plot_grid()* from package ‘cowplot’ to combine all three plots of ‘age’ into one
  - use following code to adapt colours in all plots

```
scale_fill_manual(values = c("#00AFBB", "#E7B800"))
```