Recap & Exercises

Applied Data Science using R, Session 5

Prof. Dr. Claudius Gräbner-Radkowitsch
Europa-Universität Flensburg & Institut für die Gesamtanalyse der Wirtschaft (JKU Linz)
www.claudius-graebner.com | @ClaudiusGraebner | claudius@claudius-graebner.com

Europa-Universität Flensburg

Institute for Comprehensive Analysis of the Economy

Exercises on basic operations and data types

1. Create a vector with the numbers from `-10` to `22`, evenly spaced by `0.5`-steps

2. Compute the logarithm for each element.

3. Use R to complete the following calculation:

$$(50 - 22)^2 + \frac{22}{3}$$

4. Create a vector that contains ten times the character “good”, 15 times the character “bad”, as well as the numbers from 1 to 5

5. Transform this vector into a factor that only allows for the values “good” and “bad” as levels

6. Transform this vector into an ordered factor with “good” being better than “bad”

Exercises on function definitions

- **Z-score normalization** refers to the process of normalizing every value in a vector such that the mean of all of the values is 0 and the standard deviation is 1.

- The formula for z-normalization is as follows:

$$x_Z = \frac{x_i - \mu}{\sigma}$$, where μ is the mean and σ the standard deviation

- Your task is to write an R function that takes a vector and z-normalizes it!
  - Hint: the functions `mean()` and `sd()` will be useful!

- Also: test whether your function actually works as intended!

Exercises on advanced data types

- Create a factor with the levels `"still"`, `"medium", "sparkling"`, and arbitrary instances of the three levels

- Get the relative frequencies for `"medium"` of this factor → check out the function `table()`

- Create a data frame with two columns, one called `"nb"` containing the numbers `1` to `5` as double, the other called `"char"` containing the numbers `6` to `10` as character

- Transform this data frame into a `tibble`! Use the adequate test functions to verify the transformation was successful!

- Extract the second column of the `tibble` such that you get an atomic vector