Prof. Vera Demberg

Statistics with R

WS 2020/2021

Exercise Sheet №0

Deadline: 09.11.2020 23:55

Submission via Moodle: https://lms.sulb.uni-saarland.de/moodle/course/view.php?id=4018
(Statistics with R (2020/2021; for CS and Coli students))

It is required to work in groups of three students, It is possible that group members participate in different tutorial sessions.
Please fill in the excel sheet stating your group members (name, matriculation #, email).
https://docs.google.com/spreadsheets/d/1T2OC3uY09nRCUGdqLwRJDMFDjJDVQ9fJgwy8Ii6qu0I/edit?usp=sharing

If you are not able to find a group till the submission of the first sheet, we will assign you one.
If you are in a group, all team members need to submit the results individually and write name and matriculation number of all students you have worked with.

Preliminaries

• Install R on your laptop (for instructions see e.g. chapter 3.1 from the Navarro book).

• Make yourself familiar with the basics of the programming language R (the topics covered in chapters 3 and 4 of the Navarro book). Additional resources such as QuickR and a book on R programming are linked from the course website.

• If you have little programming experience, please work through the datacamp course "Introduction to R".

Measurement Scales, Populations and Samples

1. Give an example of a continuous and a discrete measurement variable from your interest area (e.g. HCI, Bioinformatics, etc.). What are their scales?

2. Give an example of a population and of a sample (in the context of some specific research question). Do you know (from studies you have read or heard about), how the samples are typically chosen in your field?

3. In your example, is it really a random sample of the population or not? Why?

4. Identify the independent and dependent variable in the following example and mention its scale

- Comparing the time taken by three different modes of transportations namely car, bus and train.

- Calculating the number of “YES” or “NO” votes of population measures for “**green revolution**”.
- Whether caffeine affects the appetite of a person by measuring the hunger level in a 5-point Likert scale after the caffeine intake.

5. Please find a recent research paper form an area you're interested in, which includes a study that reports statistical significance. Write down:

a) the research

b) the population

c) the sample and whether you think it was random (and why you think that)

d) do you feel there was a bias in the collection of data?

e) how was the design of the study? Was it experimental or corpus?

f) the dependent variable

g) the independent variable

h) for all variables, whether they're continuous or discrete

i) the measurement scale of each variable

j) What statistical test did the authors use?

k) Can you find out why they used that specific statistical test?

Please also provide a reference for the paper you chose.

Submission in groups of three students .