Course "Empirical Evaluation in Informatics" SoSe 2017

Freie Universität Berlin, Software Engineering Research Group
Prof. Dr. Lutz Prechelt, Franz Zieris

Practice sheet 1 **Get to know R** due on 2017-04-23

Important: Due to unfortunate positions of floating holidays this year, we are three Mondays short, but still need to cover all content. The first practice sheet is therefore *due before the first lecture*, which will take place on 2017-04-24. But don't worry: You can do this without much prior knowledge.

Task 1-1: Organization

Make sure you're enrolled in this course in the KVV, as all of the tutorial's organization will be handled there. The practice sheets will be uploaded there along with supplementary material.

Through the semester, there will be different types of tasks. Task you are supposed to work alone on are marked with . Many tasks, however, are meant for pair work – if you don't have a partner already, we will find one for you in the first tutorial.

After you finished working on the three tasks below, make sure to upload your solutions in the KVV.

Task 1-2: Research definition

Learning goal: Getting an understanding of the concept of "research".

Find several (at least two) definitions of "research" (from dictionaries, books on research methods, and from friends, family, or colleagues. What are the recurring themes in the definitions you found? Can you now provide a good definition on your own?

Task 1-3: Learn basic terminology in R

Learning goals: Getting to know the most important operators of R and how to use them; gaining a general understanding of the use of vectors, matrices, and lists in R; being able to use the `help` function and documentation of R.

a) **Install R** (reference version for this course is **3.3.3**) and start it.
- Download from https://www.r-project.org (or https://sourceforge.net/projects/rportable for a portable version)
- Start the program by using the commands `Rterm` or `Rgui` (Windows) or `R` (Linux).
- (There might be suitable installations on the public computer pools as well. Contact the tutor if you want to work there, but encounter problems.)

b) Download the file `R_intro_session.r` from the KVV and open it with a text editor. Repeat the individual commands in your R environment and **understand** what happens.
- Use the `help` function to get more detailed information on the functions and operators used.
- Be prepared to **explain** for each command what happens and why.
- If necessary, read up background information in the documents "An Introduction to R" or "The R language definition" which you can find via `help.start()`.

c) **List three parts** in the *official R documentation* (in three different places) which explain the indexing in R.

Above all: Feel free to play around with R! This might be the best way to get a feeling of how R works.

Task 1-4: Small R task

Learning goal: Being able to define and use your own functions.

Simulate **100 000 throws with a dice** and count how often the sequence $3, 4, 5$ came up.

Use a formulation being suited to R as much as possible (i.e. without explicit loop); not much longer than 2 lines. For the solution you merely need the functions `sample` and `sum` as well as indexing and operators (there are other solutions as well).

To give you an idea of what the result should look like, the precise expected value can be calculated as follows:

```
(100000-(3-1))/6^3

## [1] 462.9537
```