Session 1 Problem Sheet

Welcome to the First Problem Sheet! Wohooo!

It will be very different from the upcoming problem sheets because its main purpose is to help you get comfortable with R and RStudio. Please keep in mind that the tasks will vary strongly in their difficulty and some tasks are meant to make you (1) struggle, (2) seek out help online, (3) discuss with your colleagues, and (4) ask questions during the live sessions or during meetings with demonstrators. These steps should in fact be your essential approach to each of the tasks in the upcoming term. First, try it yourself. If that does not work – ask a friend! If the friend has no idea – ask *us*! This makes sure that you really learn the most out of each exercise.

There are a few things you should complete before getting started with this exercise sheet.

First, you need to install R and RStudio. To do so, please complete steps 4.1.1. and 4.1.2 in https://smobsc.readthedocs.io/en/latest/chapter_ana/Introduction%20to%20Programming%20with%20R.html with additional support from the first steps of http://milton-the-cat.rocks/learnr/r/r_getting_started/#section-overview. Here, you should probably watch all videos until "*R Markdown*", but ideally you should actually watch all the provided content.

In case you have existing R and RStudio installations on your machine, make sure that you have recent versions of the software and not some super old versions.

During the live session we can troubleshoot any issues that have come up with installation, but really try your best to have completed these steps.

Next, check out the resources in the *MT Session 0: General Resources* section in Canvas. There you will find further useful materials and references that will likely proof useful throughout the course.

Now, you are ready for the first problem sheet! Please open RStudio and record your solutions to each of the questions in an R Script. Make sure to save your script regularly to avoid any loss of progress.

Hey advanced R users! You are bored by all of this? You crushed this problem sheet in 10 minutes and now you are wondering if this entire course will be just a big waste of time? Not to worry, I promise you that you will learn a lot during the next two terms. In the following weeks, there will be some extra content for advanced users AND problem sheets will have *beginner* and *expert* tracks. So, stick out the first couple of session if they feel slow, help novices (no better way of learning) and you will get your taste soon enough. I strongly encourage you to keep coming or at least completing the exercise sheets on your own time – last year we had advanced users who skipped the first term and had troubles keeping up with content of the second term because of it. And here is something fun to keep you busy and constructive now: https://www.dropbox.com/sh/xs3usyp3q9o88wn/AADgsjeMzQnS_IZSIkyzA9lda?dl=0. Also, you should make a RMarkdown out of both scripts and export as a nice html file.

Dr. Dejan Draschkow, October 2020

Session 1 Problem Sheet

Tipp: Use the **#** sign in front of a line in order to make it a comment. Comments are not read by R and are only used as helpful notes in your script. The comments might be helpful reminders when you look at the code again in a couple of weeks.

1. R as a calculator.

a) Assign 344 to a using **<-**

b) Compute $25 * n - a$ for $n = 13.76$

c) Compute x^3 for $x = 14$ using **^**

2. Basic vector operations. Define ***x*** and ***y*** with:

x <- c(2, 4, 6, 8, 10)

y <- c(1, 3, 5, 7, 9, 11)

Try to guess the outcome of the following R commands and explain in a few words what they do (**#** by commenting in the script):

a) ***x+1;*** ***y*2;*** ***x + y*** (why is there a warning?)***;*** ***x*y***

b) ***length(x)*** and ***length(y)***

c) ***sum(x>5)*** and ***sum(x[x>5])*** (the difference between ***()*** and ***[]*** brackets?)

d) ***y[-3];*** ***y[x];*** ***y[y >= 7]***

e) ***c(x,y);*** ***?plot***

3. The revolving door at the entrance of the Math building is a departmental favorite. Recently, 10 Nobel laureates were invited to determine the maximum number of students that can walk through it in 8 hours. A listing of the occurrences that each scientist counted has been made public: 159; 214; 199; 219; 200; 165; 179; 175; 144; 201.

a) What is the smallest amount? What is the largest amount?

b) What is the average amount?

c) Find the differences of the largest and smallest amounts from the mean.

Use R commands ***min( )***, ***max( )***, and ***mean( )*** for your answer.

Dr. Dejan Draschkow, October 2020

4. Prof. Vo likes to give chocolate, candy and other sweets to her PhD students in order to motivate them to work harder. Her PostDocs have noticed that since she started doing that, there are significantly less sweets at lab meetings. In order to increase the amount of candy in the lab they randomly sampled 3 PhD students and assessed their *weight* and *height*, comparing it with their own.

a) Type in the following data into the vectors *weight* and *height* using ***c( )***:

person	*1*	*2*	*3*	*4*	*5*	*6*
weight	*73*	*67*	*61*	*62*	*75*	*79*
height	*1.87*	*1.78*	*1.75*	*1.70*	*1.77*	*1.81*

b) Determine the BMI for each person using the formula: $BMI = \frac{weight}{height^2}$

c) To prove to Prof. Vo that the PhD students are better nourished, and it is in fact the PostDocs who need more sweets – compare the average BMIs of the two groups (1 to 3 = PostDocs; 4 to 6 = PhD Students).

Use the functions and basic vector operators from the previous exercises.

d) *Use the function ***data.frame()*** to create a matrix like this one

person	weight	height	BMI
1	73	1.87	20.87563
2	67	1.78	21.14632
3	61	1.75	19.91837
4	62	1.70	21.45329
5	75	1.77	23.93948
6	79	1.81	24.11404

and assign it to ***data***

e) *Try to access the data of the person with the highest BMI with the bracket notation: ***data[***insert row number here***,*** insert column number here***]***

f) *Interpret what ***data$person[data$BMI>mean(BMI)]*** is returning. Formulate it in a sentence.