Let's practice – 1

1) Inside Rstudio: Project set-up

In RStudio, create a new project in a new directory named Rcourse2022.

2) Outside Rstudio: Prepare course data for exercises

- Download course data sets (course_datasets.zip) from the course web site:
  https://edu.sib.swiss/course/view.php?id=594

- Unzip into the newly created course directory.

- Make sure the data files appear inside a sub-directory: course_datasets

Let's practice – 2

1) Prepare your first script

- Open a script file and save it with file name "ex1.R"
- Comment it (# symbol at the beginning of the line).
- Type or paste the following code:

```r
# First Steps, ex 1
w <- 3
h <- 0.5
area <- w * h
area
```

2) Look at the script (before trying to run it)

- Can you understand each line? What do you expect it to print to the console?

3) Run the script and explore RStudio features

- Run the script line by line. Try both the "Run" button and the keyboard shortcut. Watch variables appear in the Environment panel (top right).

- Watch what is printed to the console (bottom left). Does it match your expectation?

Let's practice – 2bis

4) Look at the project options (RStudio's *Tools* menu). If needed, modify them to save your workspace and history and to restore them at startup.

Check if this works:

- Close RStudio.
- In your course folder, (double-)click the .rproj file.
- Does your project open? Are your variables still in the Environment?

Check if this works, too:

- Close RStudio.
- Open RStudio again.
- Verify that your project is currently closed. How do you see this?
- From inside RStudio, open your project. Are your variables now in the Environment?

Let's practice – 3

For all exercises, feel free to use

      - cheat sheets provided

      - R help (? at command prompt)

Open a new script file and save it as ex3.R

1) Assign the values 6.7 and 56.3 to variables **a** and **b**, respectively.

2) Calculate (2*a)/b + (a*b) and assign the result to variable **x**. Display the content of **x**.

3) Find out how to compute the square root of variables. Compute the square roots of **a** and **b** and of the ratio **a/b**.

4) a) Calculate the logarithm to the base 2 of **x** (i.e., $\log_2 x$).
   b) Calculate the natural logarithm of **x** (i.e., $\log_e x$).

Let's practice – 4

1) Create two vectors, **vector_a** and **vector_b**, containing the values from −5 to 5 and from 10 down to 0, respectively.

2) Calculate the (element-wise) sum, difference and product between the elements of **vector_a** and **vector_b**.

3) a) Calculate the sum of elements in **vector_a**.
   b) Calculate the overall sum of elements in both **vector_a** and **vector_b**.

4) Identify the smallest and the largest value among both **vector_a** and **vector_b**.

5) Compute the overall mean of the values among both **vector_a** and **vector_b**.

Hint: Each task in exercises 6-9 can be performed in a single statement per vector

Let's practice – 5

Open a new script and save it as "Ex5.R". Comment it.

1) In your script, write commands to install and load the package "MASS".

2) Write a command to load the bacteria data set from the package MASS:

```r
data(bacteria) # loads the bacteria data set (from MASS)
```

Execute the command. Check: You should have a variable named "bacteria" in your Environment.

3) What are the names of the columns of the **bacteria** data.frame ?

4) Use **[]** to select rows 100 to 119 of the column "ap" .

5) Use $ to get the column "week" and check how many missing values it has.

Optional : 6) Count how many rows correspond to a "placebo" treatment ("trt" column) using the comparison operator "==".

Let's practice – 6

A dataset from mouse experiments at 18 weeks is available in the file
mice_data.csv *(courtesy of F Schutz and F. Preitner)*. Let's explore the dataset to see what it contains.

1) Open a new script file in R studio, comment it and save it.

2) Have look at the csv file in R studio's file explorer. What do you need to check in order to be able to read in the file correctly?

3) Read the file into R, assign its content to object "mice_data". Examine the object.

4) How many observations and variables does the dataset have?

5) What is the structure of the dataset? What are the names and classes of the variables?

6) Which variables appear to be categorical? Convert them to factors.

7) Get the summary statistics of "mice_data"

Let's practice – 6bis

8) Use the function table() to compute the number of observations in different mouse groups. a) How many mice are included of each genotype (WT, KO)? b) How many mice are included per diet (HFD, CHOW)? c) Make a 2x2 table by genotype and diet crossed.

9) Isolate the observations for the mice on high fat diet (HFD) using subset(). Compute a summary statistics just for the weights of the subset. Then do the same for the mice on regular chow diet (CHOW). Export the data of each subgroup to a csv file.

10) Look at the results from the two previous exercises. What does this initial exploration of the data suggest about mouse weights?

11) *Optional:* Compute the means and standard deviations for WT and KO mouse weights using tapply(). Then do the same for CHOW and HFD groups.