

ST203: R for Data Science and Statistics

Rafael Moral

Assignment 1 – 2021

- Do all questions: only one randomly chosen question will be marked.
- Upload your script file via Moodle before 23:59 on Friday 22 October.
- You may include your code and your **commented** answers in the same script file.
- You may submit either an R script (‘.R’) or an R Markdown file (‘.Rmd’).
- Place your name and student number on the first line of your R script or in the YAML header in your R Markdown file.

Question 1

Explain the result of the following.

```
x <- 1:10
x * (2/3) * (3/2) - x

## [1] 0.000000e+00 0.000000e+00 0.000000e+00 0.000000e+00 0.000000e+00
## [6] 0.000000e+00 -8.881784e-16 0.000000e+00 0.000000e+00 0.000000e+00
```

Question 2

Explain the result of the following.

```
2 + 2 == 4

## [1] TRUE

sqrt(2) * sqrt(2) == 2

## [1] FALSE

all.equal(sqrt(2) * sqrt(2), 2)

## [1] TRUE
```

You might want to check the help page for `all.equal`.

Question 3

Look at the help page for the function `rep()`. Using `rep()` and `seq()`, create the vector

```
1 2 3 4 5 2 3 4 5 6 3 4 5 6 7 4 5 6 7 8 5 6 7 8 9
```

Question 4

Calculate the sum $\sum_{j=1}^n j$ by constructing a vector and using the `sum` function. Compare with $n(n+1)/2$ for $n = 100, 200, 400, 800$.

Question 5

Assign `x` to $1/3$ and `n` to 10. Calculate the sum $\sum_{j=0}^n x^j$ by constructing a vector and using the `sum` function. Compare with $\frac{1-x^{n+1}}{1-x}$.

Question 6

- a) Using the `cars` data frame, write a logical expression which extracts cars whose `dist` is between 20 and 25 inclusive. (Do not use `subset`.)
- b) Write an expression which computes the average speed of cars whose `dist` is between 20 and 25.
- c) Use `subset` to extract cars whose `dist` is between 20 and 25.

Question 7

The following are a sample of observations on incoming solar radiation (W/m^{-2}) at a greenhouse.

11.1 10.6 6.3 8.8 10.7 11.2 8.9 12.2

- a) Assign the object to a vector called `solar_radiation`.
 - b) Find the mean, median, and variance.
 - c) Add 10 to each observation and assign the resulting vector to `sr10`. Find the mean, median, and variance of `sr10`. Which statistics change and by how much?
 - d) Multiply each observation by -2 and assign the resulting vector to `sr2`. Find the mean, median, and variance of `sr2`. Which statistics change and by how much?
1. Plot a histogram of each of `solar_radiation`, `sr10`, and `sr2`.