R Final Review

2022-11-29

We would like to get a file named with your first name. The file should be an R-script with the file extension ".R". Send the file by Email to yunrui.liu@unibas.ch

Duration: 60 minutes

Material: Everything that is helpful

NO COMMUNICATION!

Elegant R syntax: +1pt

Clearly structured script file with comments: +1pt

Create data

Below, we create an imaginary (random) dataset named *measures_data* representing 2 successive measures (percentages ranging from 0 to 100), taken on 10 different "patients":

```
set.seed(1234) # this allows to reproduce the same random sampling every time we run the command
measures_data <- matrix(sample(1:100, size = 100, replace = F),
                        ncol=10, nrow=2)
colnames(measures_data) <- LETTERS[1:10]
rownames(measures_data) <- c("measure_1", "measure_2")
measures_data
```

```
##            A  B  C  D  E  F  G  H  I  J
## measure_1 28 22  5 16 86 70 78 56 93 21
## measure_2 80  9 38  4 90 79 14 62 84 40
```

Exercise 1

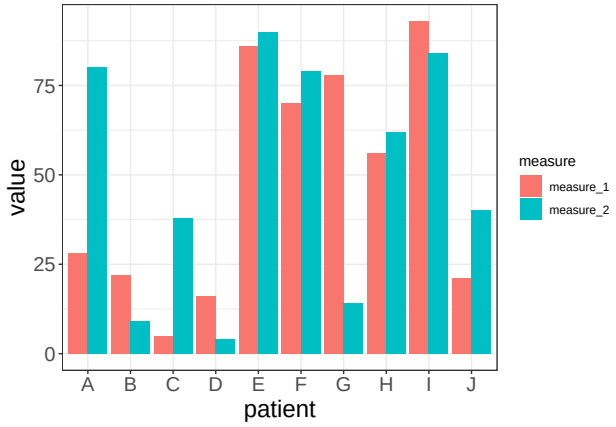
Please reshape the *measures_data* dataset into tidy (long) format, with one variable in each column. New data should looks like that:

```
## # A tibble: 20 x 3
##    measure   patient value
##    <chr>     <chr>   <int>
##  1 measure_1 A          28
##  2 measure_1 B          22
##  3 measure_1 C           5
##  4 measure_1 D          16
##  5 measure_1 E          86
##  6 measure_1 F          70
##  7 measure_1 G          78
```

```
##  8 measure_1 H          56
##  9 measure_1 I          93
## 10 measure_1 J          21
## 11 measure_2 A          80
## 12 measure_2 B           9
## 13 measure_2 C          38
## 14 measure_2 D           4
## 15 measure_2 E          90
## 16 measure_2 F          79
## 17 measure_2 G          14
## 18 measure_2 H          62
## 19 measure_2 I          84
## 20 measure_2 J          40
```

Exercise 2

Please create a bar plot (following figure) to represent the data, using different color to represent different measure methods



Exercise 3

Please create a boxplot with jittered points (following figure) to represent the data