

STAT 362 R for Data Science W23

Assignment 2

Please follow the general instructions as in Assignment 1.

Due: Feb 10, 2023 (11:59pm)

Q1. Write a function called `A2_Q1` that takes a matrix as an argument and returns the same matrix with all negative values replaced with 0.

Example:

```
# example of a matrix with negative entry
A <- matrix(c(-2, 1, 1, 3), 2, 2)
A
##      [,1] [,2]
## [1,]  -2   1
## [2,]   1   3
A2_Q1(A)
##      [,1] [,2]
## [1,]   0   1
## [2,]   1   3
```

Q2: Write a function called `A2_Q2` that takes a square matrix as an argument and returns the sum of the diagonal elements.

Example:

```
A <- matrix(c(-2, 1, 1, 10), 2, 2)
A
##      [,1] [,2]
## [1,]  -2   1
## [2,]   1  10
A2_Q2(A)
## [1] 8
```

Q3: Write a function called `A2_Q3` that takes a data frame and a column name as arguments, and returns the number of missing values in the specified column.

Example:

```
data <- data.frame(x = 1:5, y = c(1, NA, NA, 4, 4))
data
##    x y
## 1 1 1
## 2 2 NA
## 3 3 NA
## 4 4 4
## 5 5 4
A2_Q3(data, "y")
## [1] 2
```

Q4(a): Let X and Y be two independent random variables, where $X \sim N(\text{mean} = 0, \text{sd} = 2)$ and $Y \sim \text{Exp}(\text{rate} = 3)$. Estimate $P(X > Y)$ using simulation.

Q4(b): Let X and Y be two independent random variables, where $X \sim N(\text{mean} = 2, \text{sd} = 1)$ and $Y \sim \text{Exp}(\text{rate} = 2)$. Estimate $E(\min(X, Y))$ using simulation.

Q5. Given that `x = 1:100`. Write R code to compute

$$S := 1^2 - 2^2 + 3^2 - \dots + 99^2 - 100^2.$$

Instruction for Q6-Q10: Load the package `tidyverse` by using `library(tidyverse)`. Install it if you haven't done so by typing `install.packages("tidyverse")`.

Recall that the dataset `mtcars` comes with base R.

Q6. Use `filter` to find the subset of `mtcars` such that the cars have exactly 4 forward gears.

Hint: type `?mtcars` to learn more about this dataset and find out which variable represents the number of forward gears.

Hint: use `"=="` instead `"="` for logical comparison in R.

Q7(a): Use `filter` to find the subset of `mtcars` such that the cars have 6 cylinders or 4 forward gears.

Note: Given two events A and B , when we say " A or B ", it always includes the possibility of " A and B ". In this case, you should include the cars with 6 cylinders and 4 forward gears.

Q7(b): How many cars in `mtcars` have both 6 cylinders and 4 forward gears? Write code to find the number. Do not count by hand.

Q8: Use `arrange` to sort `mtcars`. Then, find the heaviest car and the lightest car.

Now, load the library `nycflights13` (install it first if you haven't done so). In the following questions, flights refer to the flights in the dataset `flights` in the library `nycflights13`.

Q9: Use `filter` to select flights that departed in Jan or Nov such that the flights were not delayed on arrival by more than 10 minutes.

Q10: Write R code to find the number of flights that departed in May.