Midterm

Instructions

- Read all of these instructions closely.
- The midterm is due Monday, March 6, 2023 at 4:15pm EST.
- Submit files electronically to me via:
  - Github or
  - Slack or
  - Email
- You may use class resources and online resources during the exam. You may not use ChatGPT or any other software the write the code for you. You also may not message each other during the exam.

Question 1

For question 1, we'll use the model output from a simple regression.

```
X <- c(4.17,5.58,5.18,6.11,4.50,4.61,5.17,4.53,5.33,5.14)
Y <- c(4.81,4.17,4.41,3.59,5.87,3.83,6.03,4.89,4.32,4.69)
lm_output <- lm(Y ~ X)
```

1a

`lm_output` is a list, containing regression model output. Print only the first element of this list.

Answer:

```
#answer here
lm_output[[1]]
```

```
## (Intercept)           X
##   7.7957139  -0.6229559
```

```
lm_output$coefficients
```

```
## (Intercept)           X
##   7.7957139  -0.6229559
```

1b

In words, what is the **class,** or type of data, that you printed in 1a? It is not required, but you can also use code to justify your answer.

Answer:

It is of class numeric.

```
class(lm_output[[1]])
```

```
## [1] "numeric"
```

1c

In words, what is the **data structure** that you printed in 1a? It is not required, but you can also use code to justify your answer.

Answer:

It is a vector.

```
is.vector(lm_output[[1]])
```

```
## [1] TRUE
```

Question 2

In words, describe what the following code is doing. Note the `cars` data has 50 observations and two columns: `speed` which is the stopping speed in mph, and `dist` which is the stopping distance in feet. The data were recorded in the 1920s so the cars were pretty slow!

```
data(cars)
for(i in 1:nrow(cars)){
  if(cars$speed[i] == min(cars$speed)){
    print("Stopping distance (in feet) for slowest car(s) in the dataset")
    print(cars$dist[i])
  }
}
```

```
## [1] "Stopping distance (in feet) for slowest car(s) in the dataset"
## [1] 2
## [1] "Stopping distance (in feet) for slowest car(s) in the dataset"
## [1] 10
```

Answer:

The for loop iterates over 1, 2, 3, ... 50, each time checking the `speed` variable in the ith position to assess if it equal to the minimum speed in the dataset. If so, it prints the stopping distance in the ith position as well.

Question 3

3a

Write a function that will take a vector `vec` as the only input and returns the number of elements that are `NA`in `vec`.

Answer:

```
#answer here
na_function <- function(vec){
  sum(is.na(vec))
}
```

3b

Use your function to print the number of elements that are NA in the following vector.

Answer:

```r
vec <- c(1, 10, NA, NA, 12)
#answer here
na_function(vec)
```

```
## [1] 2
```