

# Module 1 R Exercises

Dr. Foti

If you are reading this, then you successfully downloaded, saved, and opened the RMarkdown file! Nice Job!

Please complete the following exercises by typing your code under each question. When finished, click Knit and include your name in the filename. Upload the file you saved back to Canvas as your submission.

As you write your code in the code chunks below, you can click run or use ctrl + enter to run/test your code as you go. When you are finished, click Knit to create the final document for submission. If you are getting errors that you cannot figure out, ask for help!

## Question 1

We are going to use the `runif()` function to complete some tasks. Using the code chunk below, open the help documentation for the function named `runif` (hint: use the `?` symbol). Underneath the code chunk, explain what the function does.

```
?runif
```

```
## starting httpd help server ... done
```

`runif()` generates random deviates for the purpose of providing information on an interval's uniform distribution.

## Question 2

Draw 20 random numbers ( $n = 20$ ) from the Uniform distribution with the arguments `min = 0` and `max = 10`.

```
runif(20, min = 0, max = 10)
```

```
## [1] 7.9978192 6.5781334 0.8674984 4.9980546 7.4436896 8.3878056 6.7551182
## [8] 0.7257023 4.9918772 9.0111731 4.8897993 6.7081606 5.0343159 3.6616082
## [15] 1.4144319 0.8221175 4.2522581 3.3388541 9.6563665 7.1331993
```

## Question 3

Repeat question 2, except this time store your 20 numbers in a variable called: `numbers`.

```
numbers <- runif(20, min = 0, max = 10)
```

## Question 4

Use what you know or use Google to complete the following on your numbers variable: - Find the mean of the 20 numbers, find the median of the 20 numbers - Find the minimum value in your set, and the maximum value in your set.

```
mean(numbers)
```

```
## [1] 5.409157
```

```
median(numbers)
```

```
## [1] 5.155979
```

```
min(numbers)
```

```
## [1] 1.536836
```

```
max(numbers)
```

```
## [1] 9.546184
```

## Question 5

Try to use the function called `summary()` on your numbers variable. What does it do?

```
summary(numbers)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   1.537   2.811   5.156   5.409   7.773   9.546
```

It finds all the previous values of question 4 (min, max, mean, median) as well as the 1st. and 3rd. quartile

## Question 6

Run the following code: `numbers < 10` What did this do? What is the result telling you?

```
numbers < 10
```

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

The code determines if each number in the object is less than 10. It returns all TRUE because our max possible value was set to 10.

## Question 7

Instead, try running the following: `numbers < 5` What did this do? What is the result telling you?

```
numbers < 5
```

```
## [1] FALSE TRUE TRUE TRUE TRUE FALSE FALSE FALSE FALSE TRUE TRUE TRUE
## [13] FALSE TRUE FALSE FALSE TRUE FALSE TRUE FALSE
```

Similarly to question 8, the result is determining if each number is less than 5. We get some FALSE results as expected as many of the numbers in the object are less than 5.

## Question 8

The data type “logical” is equivalent to the binary 0 (FALSE) and 1 (TRUE). Try to take the `sum()` of your result from step 7, what is the result? What does it mean?

```
sum(numbers<5)
```

```
## [1] 10
```

The result is the total amount of true values, i.e. how many numbers are less than 5. In my first instance it was 12, but the result will be different after this is knitted as the code will be rerun with a new list of numbers.

## Question 9

Install the package named `palmerpenguins` by using the `install.packages()` function. Alternatively, use Tools -> Install Packages. Then, load the package by running the following code: `library(palmerpenguins)`. Open the help documentation for the dataset called `penguins` by using the `?` symbol. Run the summary function on the dataset called `penguins`.

```
library(palmerpenguins)
```

```
## Warning: package 'palmerpenguins' was built under R version 4.4.2
```

```
?penguins
summary(penguins)
```

```
##      species      island bill_length_mm bill_depth_mm
## Adelie   :152  Biscoe   :168   Min.    :32.10   Min.    :13.10
## Chinstrap: 68  Dream    :124   1st Qu.:39.23   1st Qu.:15.60
## Gentoo   :124  Torgersen: 52   Median :44.45   Median :17.30
##                                     Mean    :43.92   Mean    :17.15
##                                     3rd Qu.:48.50   3rd Qu.:18.70
##                                     Max.    :59.60   Max.    :21.50
##                                     NA's    :2      NA's    :2
## flipper_length_mm body_mass_g      sex      year
## Min.    :172.0     Min.    :2700  female:165  Min.    :2007
## 1st Qu.:190.0     1st Qu.:3550  male  :168  1st Qu.:2007
## Median :197.0     Median :4050  NA's   : 11  Median :2008
## Mean    :200.9     Mean    :4202                Mean    :2008
## 3rd Qu.:213.0     3rd Qu.:4750                3rd Qu.:2009
## Max.    :231.0     Max.    :6300                Max.    :2009
## NA's    :2        NA's    :2
```

## Question 10

Install the package named tidyverse. This package will be useful throughout the course for visualizations and data manipulation. Removing the “comments” (# symbols) from the following code (check the “Code” menu, next to File and Edit; highlight your code and then click comment/uncomment lines) to get a sneak preview at some of what we will learn this semester. Once you see the graph, answer the following question: Are we able to differentiate penguin species by their body mass and flipper length?

```
library(tidyverse)
```

```
## Warning: package 'tidyverse' was built under R version 4.4.2
```

```
## Warning: package 'tidyr' was built under R version 4.4.2
```

```
## Warning: package 'readr' was built under R version 4.4.2
```

```
## Warning: package 'purrr' was built under R version 4.4.2
```

```
## Warning: package 'forcats' was built under R version 4.4.2
```

```
## Warning: package 'lubridate' was built under R version 4.4.2
```

```
## -- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
```

```
## v dplyr      1.1.4      v readr      2.1.5
```

```
## v forcats   1.0.0      v stringr   1.5.1
```

```
## v ggplot2   3.5.1      v tibble    3.2.1
```

```
## v lubridate 1.9.3      v tidyr     1.3.1
```

```
## v purrr     1.0.2
```

```
## -- Conflicts ----- tidyverse_conflicts() --
```

```
## x dplyr::filter() masks stats::filter()
```

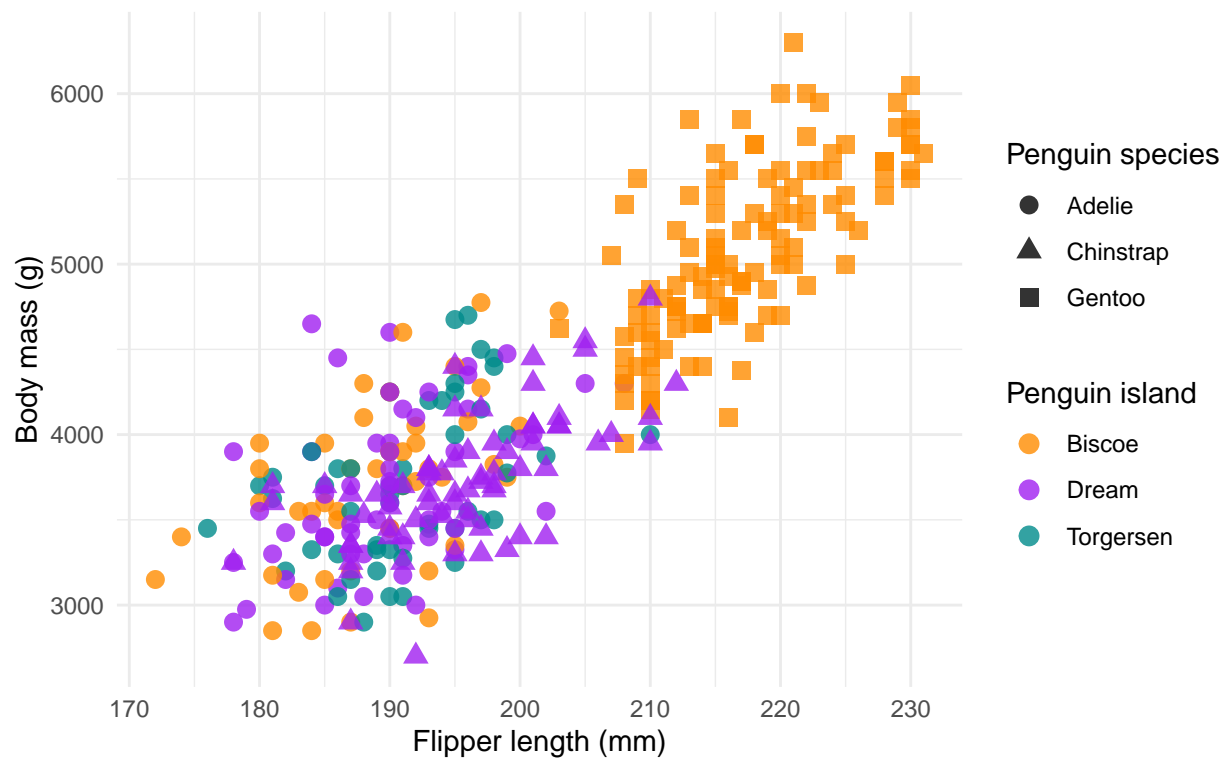
```
## x dplyr::lag()     masks stats::lag()
```

```
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
ggplot(data = penguins,
       aes(x = flipper_length_mm,
           y = body_mass_g)) +
  geom_point(aes(color = island,
                 shape = species),
            size = 3,
            alpha = 0.8) +
  scale_color_manual(values = c("darkorange", "purple", "cyan4")) +
  labs(title = "Penguin size, Palmer Station LTER",
       subtitle = "Flipper length and body mass for each island",
       x = "Flipper length (mm)",
       y = "Body mass (g)",
       color = "Penguin island",
       shape = "Penguin species") +
  theme_minimal()
```

```
## Warning: Removed 2 rows containing missing values or values outside the scale range
## ('geom_point()').
```

# Penguin size, Palmer Station LTER Flipper length and body mass for each island



Based on the graph, it seems we are able to differentiate Gentoo penguins from Chinstrap and Adelie, but not necessarily Chinstrap from Adelie as their flipper lengths are roughly in the same range.