

Assignment 2: Coding Basics

Caroline Rowley

OVERVIEW

This exercise accompanies the lessons in Environmental Data Analytics on coding basics.

Directions

1. Rename this file `<FirstLast>_A02_CodingBasics.Rmd` (replacing `<FirstLast>` with your first and last name).
2. Change “Student Name” on line 3 (above) with your name.
3. Work through the steps, **creating code and output** that fulfill each instruction.
4. Be sure to **answer the questions** in this assignment document.
5. When you have completed the assignment, **Knit** the text and code into a single PDF file.
6. After Knitting, submit the completed exercise (PDF file) to Sakai.

Basics Day 1

1. Generate a sequence of numbers from one to 100, increasing by fours. Assign this sequence a name.
2. Compute the mean and median of this sequence.
3. Ask R to determine whether the mean is greater than the median.
4. Insert comments in your code to describe what you are doing.

```
#1.  
seq(1, 100, 4) #I am asking R to create a sequence from 1 to 100 by 4
```

```
## [1] 1 5 9 13 17 21 25 29 33 37 41 45 49 53 57 61 65 69 73 77 81 85 89 93 97
```

```
one_to_100_by_4 <- seq(1, 100, 4) #I am assigning the sequence a name
```

```
#2  
mean_one_to_100_by_4 <- mean(one_to_100_by_4) #I am asking R to calculate the mean of the data set, and  
median_one_to_100_by_4 <- median(one_to_100_by_4) #I am asking R to calculate the median of the data se
```

```
#3.  
mean_one_to_100_by_4 > median_one_to_100_by_4 #I am asking R to determine if the mean is greater than the
```

```
## [1] FALSE
```

Basics Day 2

5. Create a series of vectors, each with four components, consisting of (a) names of students, (b) test scores out of a total 100 points, and (c) whether or not they have passed the test (TRUE or FALSE) with a passing grade of 50.
6. Label each vector with a comment on what type of vector it is.
7. Combine each of the vectors into a data frame. Assign the data frame an informative name.
8. Label the columns of your data frame with informative titles.

```
#5
gibbon_names <- c("hoolock", "hylobates", "nomascus", "symphalangus") #the names of the students, each
scores <- c(100, 65, 87, 26) #scores that each gibbon got on the test
pass <- c(TRUE, TRUE, TRUE, FALSE) #whether or not they passed the test (received a score greater than 50)
gibbon_test_scores <- data.frame(gibbon_names, scores, pass) #this code creates a data frame, using the
```

9. QUESTION: How is this data frame different from a matrix?

Answer: The data frame allows heterogenous elements such as characters and numbers. A matrix only allows homogenous data to be stored- either all numerical or all characters. A matrix also has a fixed number of rows and columns, whereas a data frame can contain multiple data types in multiple columns called fields.

10. Create a function with an if/else statement. Your function should take a **vector** of test scores and print (not return) whether a given test score is a passing grade of 50 or above (TRUE or FALSE). You will need to choose either the **if** and **else** statements or the **ifelse** statement.
11. Apply your function to the vector with test scores that you created in number 5.

```
ifelse(scores>50, "pass", "fail") #this code will test with if the scores in the data frame are passing
```

```
## [1] "pass" "pass" "pass" "fail"
```

12. QUESTION: Which option of **if** and **else** vs. **ifelse** worked? Why?

Answer: 'ifelse' worked for me. I believe this is because R was given two clear options that worked with the data (pass or fail). When I tried to run the 'if' code I got errors over commas that I did not get when running the 'ifelse' code. It could be user error and my limited understanding of R.