Coding Challenge 6

Madelyn Thompson

2025-03-27

Question 1 Regarding reproducibility, what is the main point of writing your own functions and iterations? To reduce the risk for copy and paste errors

Question 2 In your own words, describe how to write a function and a for loop in R and how they work. Give me specifics like syntax, where to write code, and how the results are returned.

A function takes inputs called arguments, runs a chuck of code using those arguments, and returns an output. It simplifies the code and reduces errors from repeatedly copying large chunks of code. A function is written using "function()" in which you list the names of arguments you'd like to input into the function. After that, you type a set of curly brackets {}. Within the curly brackets, you write the code that you'd like to run, taking the argument names you listed at the beginning. Finally, you can return the products using return().

A loop is a chunk of code that iterates through a set series of values that you input and runs code on all of those inputs. It starts by using the function "for()". Inside the parentheses, write the variable name for your iterations (usually called i) that you'll reference within the loop and what you'd like the loop to iterate through. This could be a string of numbers (1:10) or unique strings within a dataframe. Example: for (i in 1:100) Then, add a set a curly brackets. Inside those, you can have the code select the ith value of a dataframe or input the ith value into a function etc. Results are returned using the "print()" function.

Question 3 Read in the Cities.csv file from Canvas using a relative file path.

```
Cities <- read.csv("CodingChallenges/CodingChallenge6/Cities.csv")
# Cities <- read.csv("Cities.csv")
```

Question 4 Write a function to calculate the distance between two pairs of coordinates based on the Haversine formula (see below). The input into the function should be lat1, lon1, lat2, and lon2. The function should return the object distance_km. All the code below needs to go into the function

```
Distance <- function(lat1, lon1, lat2, lon2) {
  # convert to radians
  rad.lat1 <- lat1 * pi/180
  rad.lon1 <- lon1 * pi/180
  rad.lat2 <- lat2 * pi/180
  rad.lon2 <- lon2 * pi/180
  # Haversine formula
  delta_lat <- rad.lat2 - rad.lat1
  delta_lon <- rad.lon2 - rad.lon1
  a <- sin(delta_lat / 2)^2 + cos(rad.lat1) * cos(rad.lat2) * sin(delta_lon / 2)^2
  c <- 2 * asin(sqrt(a))
```

```
  # Earth's radius in kilometers
  earth_radius <- 6378137
  # Calculate the distance
  distance_km <- (earth_radius * c)/1000
  return(distance_km)
}
```

Question 5 Using your function, compute the distance between Auburn, AL and New York City a. Subset/filter the Cities.csv data to include only the latitude and longitude values you need and input as input to your function. b. The output of your function should be 1367.854 km

```
#Run function with inputs for the four arguments (filtered cities df for Auburn lat and long and New Yo
Distance(Cities$lat[Cities$city == "Auburn"], Cities$long[Cities$city == "Auburn"], Cities$lat[Cities$c
```

```
## [1] 1367.854
```

Question 6 Now, use your function within a for loop to calculate the distance between Auburn and all other cities in the data. The output of the first 9 iterations is shown below.

```
#create a null dataframe to pass into the loop
dist_from_AU <- NULL
#loop to calculate difference from Auburn and other cities in df
for(i in Cities$city) {
  distance <- data.frame(City1 = "Auburn",
                         City2 = i,
                         Distance_km = Distance(Cities$lat[Cities$city == "Auburn"],
                                                Cities$long[Cities$city == "Auburn"],
                                                Cities$lat[Cities$city == i],
                                                Cities$long[Cities$city == i])) #creates a one row df e
  dist_from_AU <- rbind.data.frame(dist_from_AU, distance) #adds the new row to a combined df
}
head(dist_from_AU)
```

```
##    City1       City2 Distance_km
## 1 Auburn    New York   1367.8540
## 2 Auburn Los Angeles   3051.8382
## 3 Auburn     Chicago   1045.5213
## 4 Auburn       Miami    916.4138
## 5 Auburn     Houston    993.0298
## 6 Auburn      Dallas   1056.0217
```

Question 7 Commit and push a gfm .md file to GitHub inside a directory called Coding Challenge 6. Provide me a link to your github written as a clickable link in your .pdf or .docx

Link to my GitHub