**Loops and iterations – 25 pts**

PLEASE READ THIS BEFORE CONTINUING

This assignment will help you practice writing and executing for loops and writing functions. It will also involve some more practice with GitHub. You may collaborate with a partner to enhance your learning experience. Please ensure the following:

* **Collaboration**: If you work with a partner, include both names on the final submission by editing the YAML header.
* **Submission**: Only one person should submit the assignment to Canvas *in a Word document or .pdf file generated through R markdown*. Additionally, you should provide a link to your GitHub, where the assignment should be viewable by rendering it as a GitHub-flavored markdown file.
* **Setup**: It is also assumed you already have a GitHub repository for this class.
* **Time**: This should take you no longer than the class period to complete.

1. 2 pts. Regarding reproducibility, what is the main point of writing your own functions and iterations?
2. 2 pts. 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.

This dataset contains the population and coordinates (latitude and longitude) of the 40 most populous cities in the US, along with Auburn, AL. Your task is to create a function that calculates the distance between Auburn and each other city using the Haversine formula. To do this, you'll write a for loop that goes through each city in the dataset and computes the distance from Auburn. Detailed steps are provided below.

1. 2 pts. Read in the Cities.csv file from Canvas using a relative file path.
2. 6 pts. 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.

*# 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

1. 5 pts. Using your function, compute the distance between Auburn, AL and New York City
   1. Subset/filter the Cities.csv data to include only the latitude and longitude values you need and input as input to your function.
   2. The output of your function should be 1367.854 km
2. 6 pts. Now, use your function within a for loop to calculate the distance between all other cities in the data. The output of the first 9 iterations is shown below.

## [1] 1367.854

## [1] 3051.838

## [1] 1045.521

## [1] 916.4138

## [1] 993.0298

## [1] 1056.022

## [1] 1239.973

## [1] 162.5121

## [1] 1036.99

Bonus point if you can have the output of each iteration append a new row to a dataframe, generating a new column of data. In other words, the loop should create a dataframe with three columns called city1, city2, and distance\_km, as shown below. The first six rows of the dataframe are shown below.

## City1 City2 Distance\_km

## 1 New York Auburn 1367.8540

## 2 Los Angeles Auburn 3051.8382

## 3 Chicago Auburn 1045.5213

## 4 Miami Auburn 916.4138

## 5 Houston Auburn 993.0298

## 6 Dallas Auburn 1056.0217

1. 2 pts. 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