Day 4 - Data Analysis

\* If you get stuck, feel free to ask your peers in your breakout group or refer to the corresponding sections in the lesson template code.

1. Creating Data Frames
   1. Open up the file named “Wave 5 RShiny Day 4 Workshop.R” in RStudio.
   2. Introduce yourself to the other students in your breakout group (First name, grade, favorite ice cream flavor). Write this information in the script as a comment.

You should collect 10 students’ information. If there are fewer than 10 students in your group, you can use a random number generator to create imaginary data.

* 1. Create three vectors, each containing information about the students in your breakout group: one consisting of all the students’ names, one consisting of all the students’ grades, and one consisting of all the students’ favorite ice cream flavors.
  2. Combine these vectors into a dataframe, where each row represents different students’ information. Rename the row and column headers to reflect the information stored inside this dataframe.
  3. Print a list of everyone’s favorite ice cream flavor by accessing the corresponding information in the data frame using two different methods.
  4. Print your ice cream flavor by accessing it in the data frame.
  5. Print a unique list of everyone’s favorite ice cream flavors.
  6. Create a new data frame of students who are in high school (9th grade or higher) and repeat questions e., f., and g. for this new subset of students.

1. Analyzing Covid Data
   1. Download “Wave5-RShiny-Covid-Dataset.csv” and import the dataset into R. Make sure to set your working directory as your downloads folder or wherever your dataset is located.
   2. Preview the data using different commands to get a better sense of the data you will be analyzing.
   3. Pick your favorite country and analyze the following results for that country.
      1. What is the total number of covid deaths in that country?
      2. On how many days has that country reported more than 100 cases?
      3. What is the average number of deaths reported for days on which that country has reported more than 100 cases?
      4. Replace the existing data under “dateRep” with numbers representing the number of days after January 1, 2020. You can assume a month has 30 days in this case.
      5. How many days after January 1, 2020, did that country first report over 100 cases in one day?
   4. Ask 3 questions about this data and try to come up with some interesting results!