**dLab 5: Tidyverse**

**Instructions**

* Create a Quarto file called “Lab 5: Tidyverse”
* Copy the questions/prompts with the numbers/letters into the markdown file as text (i.e., in between code chunks, without any #). Use a header for each question #.
* Provide the code responses into code chunks directly beneath the questions (or beneath the text if the question requires both verbal and code answers).
* Submit both a knitted HTML file and your .qmd file to ELMS before 11:59pm.
* *See ‘lab assignment demo’ file (.qmd) on ELMS or Jupyter for an example.* *Do not directly edit this file, instead create your own quarto file, copy the content from the demo and edit that.*

**Question 1**

For this question, we will use the *diamonds* dataset from base R.

1a) Select the bottom 6 rows of the original dataframe

1b) Order the dataframe by price and then select the top 9 rows

1c) Select the 9 rows containing the diamonds with the lowest price (use a different action than you used in #3b)

1d) Select a random 6 rows

1e) Select a random 30% of the rows and output the number of rows this is (using built-in R function)

1f) Select only the carat, cut, color, and clarity variables

1g) Select only the carat, cut, color, and clarity variables (use a different method than you used in #1f)

1h) Select only the carat, color, and clarity variables (use a different action than you used in #1f and #1g)

1i) Select the top 100 most expensive diamonds, only for diamonds with a premium cut and over 2 carats

**Question 2**

For this question, we will use the *starwars* dataset. (Note: You will need dplyr loaded before seeing this in the data() window.)

Create a new variable that calculates the BMI of each character and name the variable bmi. Output the 3 male characters from Tatooine with the highest BMI. The equation for BMI is below.

BMI = mass / (height/100)2

**Question 3**

For this question, we will use the *character.ratings.csv* dataframe. This is on Jupyter in the Data folder.

Convert the data to long format. All the columns showing a trait rating should now be arranged in two columns, one called “trait” with the trait description and another called “rating” with the rating value.