PHW251 Problem Set 4

your name here

today

For this problem set you will tidy up a dataset of 500 individuals. We also want to calculate each individual's BMI and appropriately categorize them.

Load your data (500_Person_Gender_Height_Weight.csv):

Question 1

Clean the column headers to be all lower case, have no spaces, and rename "Location information" to location.

Question 2

Create a new variable that calculates BMI for each individual.

You will need to navigate the different system of measurements (metric vs imperial). Only the United States is using imperial.

- BMI calculation and conversions:
  - metric: $BMI = weight(kg)/[height(m)]^2$
  - imperial: $BMI = 703 * weight(lbs)/[height(in)]^2$
  - 1 foot = 12 inches
  - 1 cm = 0.01 meter

Although there's many ways you can accomplish this task, we want you to use an if_else() to calculate BMI with the appropriate formula based on each person's location.

Question 3

Create a new variable that categorizes BMI with case_when():

- Underweight: BMI below 18.5
- Normal: 18.5-24.9
- Overweight: 25.0-29.9
- Obese: 30.0 and Above

Could we have used if_else()?

YOUR ANSWER HERE

Question 4

Arrange your data first by location and then by descending order of BMI.

Question 5

Use a dplyr method to remove the height, weight, and BMI columns from your data.

Optional Challenge

Perform all the actions in this problem set with one dpylr call.