70% of your grade is programming accuracy and elegance (appropriate utilization of SQL and other functions) and 30% of your grade is on professional presentation and clarity of your answers (no screenshots).

Consider the PERU2 and PERU3 Datasets

a. Combine files into a single dataframe.

b. Report the dimensions, structure and overall summary statistics from your new dataset.

c. Evaluate all of the variables in the dataframe against BMI and explain if they are potential predictors of BMI. (Consider potential transformations as appropriate).

d. Enter all of the relevant variables from part c into a regression model where BMI is the dependent variable. Report the results.

e. Please provide appropriate tables, graphics and formatted output in your paper. You are suggested to use MLA or APA for all formatting.

f. Provide all of your annotated code as an appendix. Your code must match your output.