Due Date: April 4th at midnight

Here, this is data from my last class. It includes the sum of their quizzes as points, the final exam grade as points, their final grade in the class as a percentage, whether they attended class or not, and their name (note: names were randomly generated). We are interested in the effect that attending class has on one’s grade.

There should be 3 (or 4 if you do the extra credit) blocks in this code. The first:

Our libraries that we are importing

The second block should contain the following code:

Load the data using the following code: data =

data = pd.read\_csv('problem\_set\_1\_data.csv')

Rename the columns to be “quiz\_total”, “attended”, “final\_grade”, “final\_exam”, and “name”

Print a “head” and “tail” of the dataframe and a summary of the dataframe. Then, in a text block, describe what the averages mean from the dataframe

The next block of code should:

Subset the data by whether someone attended class

What is the average “quiz\_total”, “final\_grade”, “final\_exam” for those who attended class versus did not attend? Describe what you find from these results in a text box below

For 1 bonus point of extra credit: based on the data “quiz\_total”, “final\_grade”, “final\_exam” you have, was anyone working together? If so, spit out their names.