**Homework 3**

**1.** How many students reported their “0.1 Extra Credit” bonus points?

-> 503

**2.** What are the meaning, the min, and the max of students reported “0.1 Extra Credit” bonus points?

-> **Mean:** 0.7976540755467197

-> **Min:** 0.1

-> **Max:** 5.0

**3.** Among students who reported their extra credits, what is the percentage of the students who reported no more than 1 point?

-> 12.72% students who reported 0.1 points.

**4.** What is the correlation coefficient between students’ extra credits and their final score? Is correlation positive or negative?

-> Positive correlation of **0.18706068470027787**

**5.** Please plot a scatter figure to show a student’s extra credits and their final score, where x-axis is a student’s extra credits and y-axis is their final score.

A graph of blue dots

Description automatically generated

**6.** For this task and remaining tasks, only consider students who reported “0.1 Extra Credit” bonus points. That is, remove the data points with missing data from the dataset. We want to use the student’s extra credits to predict their final score, by using a linear regression model and all data samples. What is the linear equation?

-> **y = 2.79x + 84.54**

**7.** What is the root mean square error (RMSE) of the predication of the linear regression model when the model is applied to all data samples?

-> Root Mean Square Error: 11.52

**8.** Please plot a figure to show a student’s extra credits and their final score in scatter points (i.e., the same as task 5) and the linear regression model predication line from task 6, where x-axis is a student’s extra credits and y-axis is their final score.

