Assignment 2

Elements of Statistics and Probability (STA 201)

Department of Mathematics and Natural Sciences

BRAC University

Fall 2024

30 Marks

Here, A= Last digit of your student ID

B= Second last digit of your student ID

1. Problem: The following table provides data on the number of hours spent studying (in hours per week) and the corresponding scores achieved in a standardized test (out of 100):

Hours Studied	5	10	15	20	25	30	35	40	45	50+A
Test Scores	50-B	55	60	65	70	75	80	85	90	95

a. Construct a scatter plot for the data. Do the data appear to have a positive correlation, a negative correlation, or no correlation? Explain. **[3.0]**

b. Calculate the Pearson correlation coefficient for the data points. Is there a correlation between variables study hours and test score? If so, is it a positive or negative correlation? How strong the relationship between them. **[4.0]**

c. Using simple linear regression, find the equation of the regression line to predict test score based on study hours. **[4.0]**

d. Interpret the intercept and the slope of the regression coefficients equation in the context of this scenario. **[2.0]**

e. Predict the test score for the study hours of 7 and 37. **[1.0]**

2. Imagine you are conducting a study on the number of pages a group of students reads daily for leisure.

The data represents the pages read by each group: 12, 8, 20, 25, 15, 18, 22, 10, 30, 28, 16, 7, 11, 13, 40+A, 50+B.

a. Create a box plot and determine if there exist any outliers. **[4.0]**
b. Determine Pearson's coefficient of skewness an interpret your result. **[3.0]**

3. You are analyzing the monthly foot traffic (in hundreds of people) at two shopping malls, Mall X and Mall Y:

Month	Mall X	Mall Y
1	18-A	27
2	22	25
3	20	23
4	24	20-A
5	26	29
6	31+B	24
7	27	34+A
8	25	28
9	28	32
10	30	33
11	29	31
12	23	30

a. Calculate standard deviation and mean deviation for both sets of data. **[6.0]**
b. Determine which Mall has relatively more consistent monthly foot traffic. **[3.0]**