**Stat 250 ILP: Simple Linear Regression**  Our lab GSI Name is: \_\_\_\_\_\_\_\_\_\_\_\_\_\_

Group Members: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

The SPSS file “Stats250Data.sav” (available on CTools) includes data for a sample of 50 students in a previous term. The variables include **Exam 1 Score** (out of 75 points), **Exam 2 Score** (out of 75 points), **Average HW Score** (out of 30 points), and **Final Exam Score** (out of 100 points).

Your task is to develop a linear model to predict the final exam score, using one of the possible predictor variables: Exam 1 score, Exam 2 score, or Avg HW Score. First, you will decide which variable to use. Next, you will calculate your prediction equation (or least squares regression line).

1. Which variable will you use to predict the final exam score for a future student? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Why did you choose this variable? (Hint: The best answer will require some investigation via SPSS.)

2. Perform the regression analysis with SPSS. Write your prediction equation below. Be sure to use proper notation, and define variables, if appropriate.

|  |  |
| --- | --- |
| Variable: | Value: |
| Exam 1 | 58 |
| Exam 2 | 50 |
| Avg HW | 25.6 |

3. Use your equation to predict the final exam score for a student whose other scores are given in the table. Note: you will only need to use one of the values in this table.

Predicted final exam score =

These scores are for Student #11 who actually scored 92 out of 100 on the final exam.

Compute the corresponding residual for this student.

4. What are some issues you might consider before using this equation to predict the final exam score for yourself?

5. Use your regression analysis to complete the following statements:

Approximately \_\_\_\_\_\_\_\_ % of the variation in the final exam scores can be explained

by the linear relationship between final exam scores and \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

We estimate the predicted final exam score (using our model) to be off

from the observed final exam scores by about \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ points, on average.

*Hint: this quantity is sometimes referred to as approximately the average size of the residuals.*

Does there appear to be a significant positive linear relationship between your selected predictor variable and final exam score? Use a 5% significance level and provide support for your answer.