Project 3.2

Math 206, BCC

In general, which is the best predictor of a student's score on the final exam, the score on the first midterm or the second midterm? What is the predicted score on the final exam for a student who scores a 77 on the first midterm and an 88 on the second midterm?

Introduction

In a course called *Introduction to Statistics*, an instructor wants to gain insight into his students' performance by analyzing gradebook data at the end of the semester. The data set describes 105 students. You will have to decide whether scores on the first midterm or scores on the second midterm are a better way to predict a student's performance on the final exam.

Description of the Variables

- `Midterm1`: student's score on the first midterm (0-100 scale)
- `Midterm2`: student's score on the second midterm (0-100 scale)
- `Extra_Credit`: Did the student turn in the extra credit assignment? (0 = NO, 1 = YES)
- `Final`: Student's score on the final (0-100 scale)
- `Class`: Student's class (1 = Freshman, 2 = Sophomore, 3 = Junior, 4 = Senior)

Requirements

The goal of this project is to answer the research question *AND* demonstrate your understanding of concepts in Module 3. So show your work where appropriate. To tie your observations to back your thesis include explanations of general concepts to show what you know about these ideas.

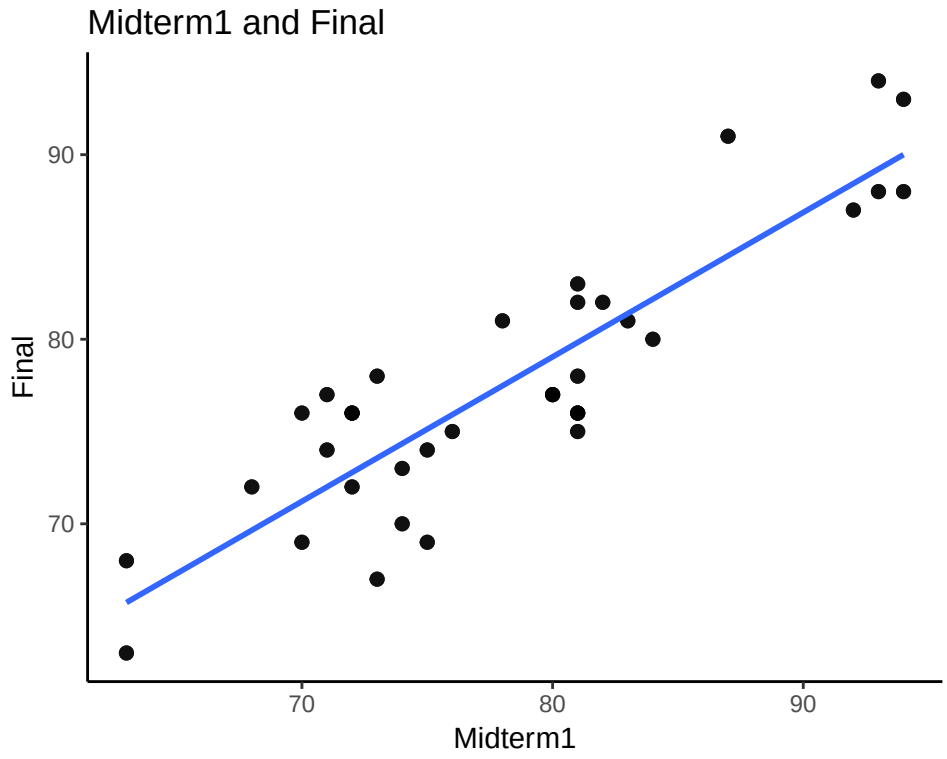
- Identify the explanatory and response variables.
- Describe the form, direction and strength of the relationship between variables. Note any outliers.
- Discuss the concept of correlation,
- Interpret the slopes and y-intercepts of the regression lines in the context of this data.
- Use the regression lines to make a prediction for the student who scored a 77 on the first midterm and an 88 on the second midterm. Show your calculations.

Relation between `Midterm1` and `Final`

Correlation between `Midterm1` and `Final` = 0.8855

Regression line:

`Final` = 16.418 + 0.783 `Midterm1`



Relation between `Midterm2` and `Final`

Correlation between `Midterm2` and `Final` = 0.7251

Regression Line:

`Final` = 35.507 + 0.542 `Midterm2`