

Business Analytics (110-1)

Assignment 1

Due: 9:00 am, Tue 19-Oct-2021

1.

Ontario high school students must complete a minimum of **six Ontario Academic Credits (OACs)** to gain admission to a university in the province. Most students take more than six OACs because universities take the advantage of the best six in deciding which students to admit. Most programs at universities require high school students to select certain courses. For example, science programs require two of chemistry, biology, and physics. Students applying to engineering must complete at least two mathematics OACs as well as physics. In recent years, **one business program** began an examination of all aspects of its program, including the criteria used to admit students. Students are required to take **English and calculus OACs**, and **the minimum high school average is about 85%**. Strangely enough, even though students are required to complete English and calculus, the marks in these subjects are not included in the average unless they are in the top six courses in a student's transcript. To examine the issue, the registrar took a random sample of students who recently graduated with the BBA (bachelor of business administration degree). He recorded the university GPA (range 0 to 12), the high school average based on the best six courses, and the high school average using English and calculus and the four next best marks in the data file "OCAs.txt".

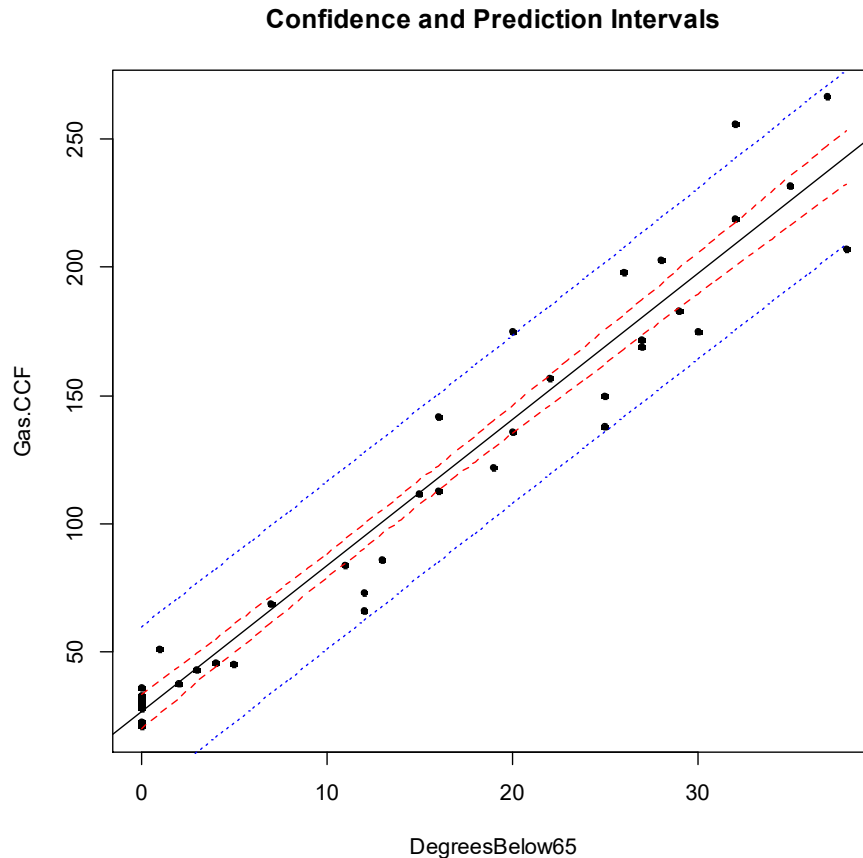
- (a) Is there a relationship between university grades and high school average using the best six OACs?
- (b) Is there a relationship between university grades and high school average using the best four OACs?
- (c) Write a comment to the university's academic vice president describing your statistical analysis and your recommendations.

2.

Utility companies in many older communities still rely on "meter readers" who visit homes to read meters that measure consumption of electricity and gas. Unless someone is home to let the meter reader inside, the utility company has to estimate the amount of energy used. The utility company in this example sells natural gas to homes in the Philadelphia area. Many of these are older homes that have the gas meter in the basement. We can estimate the use of gas in these homes with a simple linear model.

The explanatory variable is the average number of degrees below 65 during the billing period, and the response is the number of hundred cubic feet of natural gas (CCF) consumed during the billing period (about a month). The explanatory variable is set to 0 if the average temperature is above 65 (assuming a homeowner won't need heating in this case).

- (a) Use R to fit a simple linear model with the data “gas_consumption.txt”. Do the analysis, make the plot, and summarize the results.
- (b) Modify the script provided for this lecture to create the “Confidence and Prediction Intervals” plot shown on lecture note p. 2-32 and re-printed below.



3.

The number of car accidents on a particular stretch of highway seems to be related to the number of vehicles that travel over it and the speed at which they are traveling. A city alderman decided to ask the county sheriff to provide him with statistics covering the last few years, with the intention of examining these data statistically so that he can (if possible) introduce new speed laws that will reduce traffic accidents. Using the number of accidents as the dependent variable, he obtained estimates of the number of cars passing along a stretch of road and their average speeds (in miles per hour). The observations for 60 randomly selected days were recorded in file “car_accident.txt”.

Find a linear model that appropriately reveals the relationship among these variables. Clearly state your model building process and interpret the results of your final model.