HOMEWORK #5

Read:

Chapter 2 in 'Regression Analysis by Example'.
R tutorial on "Additional Topics in Simple Linear Regression"

R Assignment:

Solve the following questions using R. Hand in your R code for Questions 1-2, as well as answers to all questions.

1. The director of admissions at a small college selected 120 students at random from the new freshman class in a study to determine whether a student's grade point average (GPA) at the end of the freshman year can be predicted from their ACT test score. The data can be found at

    www.stat.columbia.edu/~martin/W2024/Data/GPA.txt

    (a) Fit a simple linear regression using ACT score as the explanatory variable, and GPA as the response variable. Verify all necessary model assumptions and include all necessary plots.
    (b) Estimate the parameter σ.
    (c) Use a t-test to determine whether or not there is a linear relationship between ACT score and GPA.
    (d) Compute the ANOVA table corresponding to the model.
    (e) From the table determine the mean square error (MSE).
    (f) Use the ANOVA F-test to determine whether or not there is a linear relationship between ACT score and GPA.
    (g) How do the results in (e) compare to those in (b)?
    (h) How do the results in (f) compare to those in (c)?
    (i) What proportion of the variation in GPA is explained by the regression model?

2. The Tri-City Office Equipment Corporation sells an imported copier on a franchise basis and performs preventive maintenance and repair service on this copier. Data was collected from 45 recent calls to perform routine preventive maintenance service; for each call, X is the number of copiers serviced and Y is the total number of minutes spent by the service person. The data can be found at

www.stat.columbia.edu/~martin/W2024/Data/Copiers.txt

(a) Fit a linear regression through the origin model to the data.
(b) Plot the fitted regression line and the data. Does the linear regression model through the origin appear to be a good fit here?
(c) Obtain the residuals. Do they sum to zero?
(d) Plot the residuals against the explanatory variable. Comment on the plot.
(e) Construct a 90% confidence interval for the estimated mean service time for calls when 5 copiers are to be serviced.
(f) Predict the service time on a new call in which six copiers is to be serviced. Use a 95% prediction interval.