HOMEWORK #4

Read:

Chapter 2 in 'Regression Analysis by Example'.
R tutorial on "Correlation and Simple Linear Regression"

R Assignment:

Solve the following questions using R. Hand in your R code for Questions 1-3, as well as answers to all questions.

1. As concrete cures, it gains strength. The following data represents the 7-day and 28-day strength (in pounds per square inch) of a certain type of concrete:

7 day strength	28 day strength	7 day strength	28 day strength
2300	4070	2480	4120
3390	5220	3380	5020
2430	4640	2660	4890
2890	4620	2620	4190
3330	4850	3340	4630

(a) Compute the mean and variance of both 7-day and 28-day strength.
(b) Compute the correlation between 7-day and 28-day strength. Comment on the strength and direction of the linear relationship between the variables.
(c) Fit a simple linear regression using 7-day strength as the explanatory variable, and 28-day strength as the response variable.
(d) What is the estimated intercept and slope of the regression line?
(e) Write in words the interpretation of the slope.
(f) What is the standard deviation around the regression line, i.e. estimate σ?
(g) Make a residual plot and a QQ-plot. Comment on whether the assumptions of the regression appear to be valid?
(h) Test the hypothesis that there is no linear relationship between the 7-day strength and the 28-day strength. What is the p-value of the test?
(i) Construct a scatter plot of 28-day strength against 7-day strength. Superimpose the regression line.
(j) Construct a 95% confidence interval for the estimated mean strength at 28-days for concrete whose 7-day strength is 3000 lbs/in^2.
(k) Construct a 95% prediction interval for a specific concrete sample whose 7-day strength was 3000 lbs/in^2.

2. A researcher at NASA measured the right humerus and right tibia (both in mm) in 11 rats that had spent an extended time at a space station.

Right Humerus	Right Tibia	Right Humerus	Right Tibia
24.8	36.05	25.9	37.38
24.59	35.57	26.11	37.96
24.59	35.57	26.63	37.46
24.29	34.58	26.31	37.75
23.81	34.2	26.84	38.5
24.87	34.73		

(a) Compute the mean and variance of both the length of the humerus and the tibia.
(b) Compute the correlation between humerus and tibia. Comment on the strength and direction of the linear relationship between the variables.
(c) Fit a simple linear regression using the length of the right humerus as the explanatory variable, and the length of the right tibia as the response variable.
(d) What is the estimated intercept and slope of the regression line?
(e) Write in words the interpretation of the slope.
(f) What is the standard deviation around the regression line, i.e. estimate σ?
(g) Make a residual plot and a QQ-plot. Comment on whether the assumptions of the regression appear to be valid?
(h) Test the hypothesis that there is no linear relationship between the length of the right humerus and the right tibia. What is the p-value?
(i) Construct a scatter plot of the length of the tibia against the length of the humerus. Superimpose the regression line.
(j) Construct a 95% confidence interval for the estimated mean humerus length of rats whose tibia is 26mm.
(k) Construct a 95% prediction interval for a specific rat whose tibia is 26mm.

3. The director of admissions of 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.
(b) Construct a 95% confidence interval for the estimated mean GPA of students whose ACT test score was 28.
(c) Construct a 95% prediction interval for a particular student whose ACT test score was 20.