MATH 2240 – Problem Set #9 Due Wednesday, November 14

1. Read the intro to Problem 11.36 on p. 603. We would like to see if "tweet rate" is a good linear predictor for revenue. (Use the TWEETS dataset.)
   (a) Create a simple scatterplot for the data (no regression line or anything else). Describe what you see. Does it look like a linear relationship? What would you guess is the direction and strength of the relationship? Are there any outliers?
   (b) Find the equation of the regression line and the value of the correlation coefficient. What does the correlation coefficient tell you? What does the slope of the regression line tell you? Be specific!
   (c) Use the regression line to predict what the revenue will be for a movie that has a tweet rate of 500.
   (d) The latest blockbuster has a tweet rate of 1500. It is wise to use our model to predict this movie's revenue? Why or why not?

2. Read the intro to Problem 11.125 on pp. 639-640. (Use the LEGAL dataset.)
   (a) Use Minitab to run an ANOVA for this linear regression.
   (b) What are the null and alternative hypotheses being tested by the ANOVA? What assumptions are needed for this test?
   (c) What, specifically, does the p-value tell us here?
   (d) What, specifically, does the coefficient of determination (r-squared) tell us here?
   (e) Use Minitab to generate a fitted line plot including confidence and prediction bands. Comment on what you see. In particular, explain the difference between the two types of bands. Would you say there are any outliers? Explain.
   (f) Another law firm has a six month cumulative advertising expenditure of $100,000. Use Minitab to find the 95% CI and the 95% PI for this value of x based on this model. Interpret both of these intervals. Be specific!
   (g) What are two or three other factors (besides advertising expenditure) that might influence the number of new personal injury cases?