1. In a seven-day study on the effect of ozone, a group of 23 rats was kept in an ozone-free environment (Group-1) and a group of 22 rats in an ozone-rich environment (Group-2). From each member in both groups the increase in weight (in grams) was recorded.

   Ozone-free: $\bar{X}_1 = 22.40, \quad s_1 = 4.2$

   Ozone-rich: $\bar{X}_2 = 12.2, \quad s_2 = 4.7.$

   (a) Do the data provide sufficient evidence to indicate that there is an ozone affect to the increase of weight? Test using $\alpha = 0.05$.

   (b) Find the p-value for the test.

2. Shown below are the number of galleys for a manuscript (X) and the dollar cost of correcting typographical errors (Y) in a random sample of recent orders handled by a firm specializing in technical manuscripts. Assume that the simple linear regression is appropriate with normally distributed independent error terms, answer the following questions:

   $$n = 6, \quad \sum X = 92, \quad \sum Y = 1652, \quad \sum X^2 = 1930, \quad \sum Y^2 = 620394, \quad \sum XY = 34602$$

   (1) Find the estimated regression line.

   (2) Test for positive correlation at 3% level.

   (3) Find a 90% confidence interval for the cost when the number of galleys is 10.

3. A criminologist studying the relationship between level of education and crime rate in medium-sized U.S. counties collected the following data for a random sample of 84 counties. X is the percentage of individuals in the county having at least a high-school diploma, and Y is the crime rate (crimes reported per 100,000 residents) last year. Use a simple linear regression model to fit the data given in the file (Test1_Appl_Reg_V5_Data.xlsx)

   a Obtain the estimated regression function. Does the linear regression function appear to give a good fit here at 3% level?

   b. Obtain point estimate of the mean crime rate last year in counties with 80% high school graduation.

   EXTRA CREDIT (2 points)

   c. Obtain point estimates of the difference in the mean crime rate for two counties whose high-school graduation rates differ by 4.5%

4. (Extra Credit 5 points): In a simple linear data set X and Y variables are both standardized first (i.e X_i is replaced by $X_i^\star = \frac{X_i - \bar{X}}{\sqrt{SS_x}}$ and Y_i is replaced by $Y_i^\star = \frac{Y_i - \bar{Y}}{\sqrt{SS_y}}$) and then SLR model is executed using $X_i^\star$ and $Y_i^\star$. Then show that (algebraically) for this new SLR model the estimated slope and the sample correlation coefficient are equal (i.e $\hat{\beta}_1 = r$).