Correlation and Regression Analysis

1. An auto manufacturing company wanted to investigate how the price of one of its car models depreciates with age. The research department at the company took a sample of eight cars of this model and collected the following information on the ages (in years) and prices (in hundreds of dollars) of these cars

Age	8	3	6	9	2	5	6	3
Price	45	210	100	33	267	134	109	235

   a) Construct a scatter diagram for these data. Does the scatter diagram exhibit a linear relationship between ages and the prices of cars?
   b) Calculate the correlation coefficient between age and price.
   c) Compute the coefficient of determination and interpret its meaning.
   d) Find the estimated regression equation of price on age.
   e) Give a brief interpretation of the values of a and b calculated in part c.
   f) Predict the price of a 7-year-old car of this model. Comment on the finding.
   g) Estimate the price of an 18-year-old car of this model. Comment on the finding.
   h) Compute standard error estimate.

2. While browsing through the magazine rack at a bookstore, a statistician decides to examine if the price of a magazine depends on the percentage of the magazine space that contains advertisements. The data collected for eight magazines are given in the following table.

Percentage containing ads	37	43	58	49	70	28	65	32
Price ($)	5.5	7	4.9	5.6	3.9	8.3	5.5	6.8

   a) Construct a scatter diagram for these data. Does the scatter diagram exhibit a linear relationship between the percentage of a magazine's space containing ads and the price of the magazine?
   b) Find the estimated regression equation of price on the percentage of space containing ads.
   c) Calculate the correlation coefficient between the percentage of containing ads and the price.
   d) Give a brief interpretation of the values of a and b calculated in part c.
   e) Predict the price of a magazine with 50% of its space containing ads.
   f) Compute standard error estimate.
   g) Compute the coefficient of determination and interpret its meaning.

3. The following table contains information on the amount of time spent each day (on average) on social networks and the Internet for social or entertainment purposes and the grade point average for a random sample of 9 college students. Here GPA is the dependent variable

Time (hours per day)	1.4	5.0	5.2	3.4	2.5	6.1	1.3	4.2	3.9
GPA	3.91	3.20	3.15	3.3	3.7	3.0	3.95	3.4	3.5

   a) Construct a scatter diagram for these data. Does the scatter diagram exhibit a linear relationship between given variables?
   b) Find the estimated regression equation of GPA on the time (hours per day).
   c) Give a brief interpretation of the values of a and b calculated in part b.
   d) Predict the GPA with time is 5 hours per day.
   e) Compute standard error estimate and interpret its meaning.
   f) Compute the coefficient of determination and interpret its meaning.

4. The city council of Bowie, Maryland, has gathered data on the no. of minor traffic accidents and the no. of youth football games that occur in town over a weekend.

Football Games	20	30	10	12	15	25	34
Minor Accidents	6	9	4	5	7	8	9

   a) Plot these data.
   b) Develop the equation that best describes these data.
   c) Predict the no. of minor traffic accidents that will occur on a weekend during which 17 football games take place in Bowie.
   d) Compute standard error estimate and interpret its meaning.
   e) Compute the coefficient of determination and interpret its meaning.

5. A household survey of monthly expenditure on food yield the following data:

Monthly expenditure on food (Rs. 1000)	10	15	20	25	30	35	40	45
Size of the family (No.)	3	4	5	6	8	7	5	9

   a) Estimate the line of best fit.
   b) Estimate the expenditure on the food of a family having 10 family members
   c) Obtain the value of correlation coefficient and coefficient of determination.
   d) Compute standard error estimate and interpret its meaning.
   e) Compute the coefficient of determination and interpret its meaning.