

Bivariate Linear Regression Exercises

Q1

A company report contains the average number of sick days per year (including time spent away from work on medical and dental appointments) for their employees with 1 to 20 years service with the company. These data are recorded as follows.

Number of Years of Service (X)	Average Number of Sick Days (Y)	Number of Years of Service (X)	Average Number of Sick Days (Y)
1	1.4	11	3.2
2	1.7	12	3.8
3	1.1	13	4.0
4	1.5	14	5.5
5	2.3	15	6.0
6	1.4	16	6.5
7	1.8	17	8.0
8	2.1	18	10.7
9	3.0	19	13.0
10	2.8	20	12.2

- Plot these data in a scatterplot.
- Find the least squares regression line and add it to the scatterplot.
- Use the least squares regression equation found to predict the average number of sick days for employees with 16.5 years of service.
- Convert the regression data to ranks and make a scatterplot of these ranks. How does the rank scatterplot compare with the raw data scatterplot.
- Find the rank regression equation.
- Use the rank regression equation found to predict the average number of sick days for employees with 16.5 years of service.

Q2

A psychology professor at Givens College is interested in the relation between hours spent studying and total points earned in the course. Data collected on 10 students who took the course last quarter are given below:

Hours Spent Studying	Total Points Earned
45	40
30	35
90	75
60	65
105	90
65	50
90	90
80	80
55	45
75	65

1. Compute the sample correlation coefficient for these data
2. Develop an estimated regression equation relating total points earned to hours spent studying
3. Test the significance of the model at the $\alpha = 0.05$ level
4. Predict the total points earned by Mark Sweeney. He spent 95 hours studying