

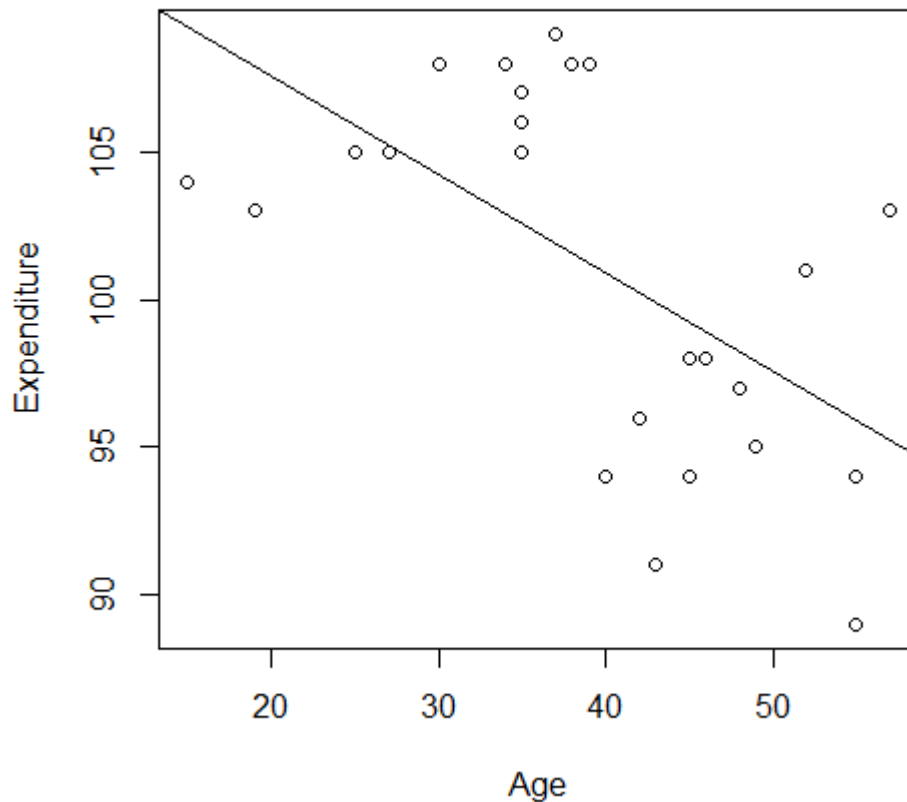
Question 1) Use all data to estimate the coefficients a and b in a simple regression model, where expenditures is the dependent variable and age is the explanatory factor. Also compute the standard error and the t -value of b .

Ans: standard Error and t Value are

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	141.6958	29.3112	4.834	6.33e-05
data1\$Expenditures	-1.0122	0.2894	-3.498	0.00185

Question 2) Make the scatter diagram of expenditures against age and add the regression line $y = a + bx$ of part (a) in this diagram. What conclusion do you draw from this diagram?



So according to scatterplot as age grows expenditure reduces

Question 3) It seems there are two sets of observations in the scatter diagram, one for clients aged 40 or higher and another for clients aged below 40. Divide the sample into these two clusters, and for each cluster estimate the coefficients a and b and determine the standard error and t-value of b.

Answer

Below 40 or equal to 40

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	13.8914	61.3444	0.226	0.825
one\$Expenditures	0.1689	0.5811	0.291	0.776

Above 40

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	25.1175	41.1544	0.610	0.555
two\$Expenditures	0.2410	0.4284	0.563	0.586

Question 4) Discuss and explain the main differences between the outcomes in parts (a) and (c). Describe in words what you have learned from these results.

Answer :

For every unit change in the age causes the expenditure decrease by -1.0122 i.e. age and expenditure are inversely proportional to each other. Where as in the part C expenditure and age are linearly proportional for the unit change in the age of observations having age less than or equal to 40 causes increase in the expenditure by 0.1689 and for the observations having age greater than 40 have value 0.2410. t value is less for the c part as compared to a part.