Multivariate 652

6.9)

From the depression data set described in Table 3.4 create a data set containing only the variables AGE and INCOME.
(a) Find the regression of income on age.
(b) Successively add and then delete each of the following points:

AGE	INCOME
42	120
80	150
180	15

and repeat the regression each time with the single extra point. How does the regression equation change? Which of the new points are outliers? Which are influential?

Ans:

a) The regression equation is given by **Income = 27.77274 - 0.16206* AGE**

b) After adding the values mentioned above the regression equation is given as follows:

Income = 26.02329-0.10441*AGE
We can see a decrease in the y-Intercept(β0) from 27.77274 to 26.02329 and an increase in slope.