1 Construct the simple linear regression equation of Y on X if

$$n = 7, \sum_{i=1}^{n} x_i = 113, \sum_{i=1}^{n} x_i^2 = 1983,$$

$$\sum_{i=1}^{n} y_i = 182 \text{ and } \sum_{i=1}^{n} x_i y_i = 3186.$$

Solution:

The simple linear regression equation of Y on X to be fitted for given data is of the form

$\hat{Y} = a + bx$ ……..(1)

The values of 'a' and 'b' have to be estimated from the sample data solving the following normal equations.

$$na + b\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i \qquad (2)$$

$$a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i y_i \qquad (3)$$

Substituting the given sample information in (2) and (3), the above equations can be expressed as

$7a + 113b = 182$ (4)

$113a + 1983b = 3186$ (5)

(4) × 113 ⇒ $791a + 12769b = 20566$

(5) × 7 ⇒ $791a + 13881b = 22302$

$$7a + 113b = 182 \quad (4)$$

$$113a + 1983b = 3186 \quad (5)$$

$$(4) \times 113 \Rightarrow 791a + 12769b = 20566$$

$$(5) \quad \times 7 \Rightarrow 791a + 13881b = 22302$$

$$\quad (-) \quad (-) \quad (-)$$

$$-1112b = -1736$$

$$\Rightarrow b = \frac{1736}{1112} = 1.56$$

$$b = 1.56$$

Substituting this in (4) it follows that,

$7a + 113 \times 1.56 = 182$

$7a + 176.28 = 182$

$7a = 182 - 176.28$

$= 5.72$

Hence, $a = 0.82$

2 Number of man-hours and the corresponding productivity (in units) are furnished below. Fit a simple linear regression equation $\hat{Y} = a + bx$ applying the method of least squares.

Man-hours	3.6	4.8	7.2	6.9	10.7	6.1	7.9	9.5	5.4
Productivity (in units)	9.3	10.2	11.5	12	18.6	13.2	10.8	22.7	12.7

Solution:

The simple linear regression equation to be fitted for the given data is

$\hat{Y} = a + bx$

Here, the estimates of *a* and *b* can be calculated using their least squares estimates

$$\hat{a} = \bar{y} - \hat{b}\bar{x}$$

i.e.,

$$\hat{a} = \frac{1}{n}\sum_{i=1}^{n} y_i - \hat{b}\frac{1}{n}\sum_{i=1}^{n} x_i$$

$$\hat{b} = \frac{\frac{1}{n}\sum_{i=1}^{n} x_i y_i - (\bar{x} \times \bar{y})}{\frac{1}{n}\sum_{i=1}^{n} x_i^2 - \bar{x}^2}$$

or equivalently $$\hat{b} = \frac{n\sum_{i=1}^{n} x_i y_i - \left(\sum_{i=1}^{n} x_i \times \sum_{i=1}^{n} y_i\right)}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}$$

From the given data, the following calculations are made with *n*=9

Man-hours x_i	Productivity y_i	x_i^2	$x_i y_i$
3.6	9.3	12.96	33.48
4.8	10.2	23.04	48.96
7.2	11.5	51.84	82.8
6.9	12	47.61	82.8
10.7	18.6	114.49	199.02
6.1	13.2	37.21	80.52
7.9	10.8	62.41	85.32
9.5	22.7	90.25	215.65
5.4	12.7	29.16	66.42
$\sum_{i=1}^{9} x_i = 62.1$	$\sum_{i=1}^{9} y_i = 121$	$\sum_{i=1}^{9} x_i^2 = 468.97$	$\sum_{i=1}^{9} x_i y_i = 894.97$

Substituting the column totals in the respective places in the of the estimates $\hat{a}$ and $\hat{b}$, their values can be calculated as follows:

$$\hat{b} = \frac{(9\times 894.97)-(62.1\times 121)}{(9\times 468.97)-(62.1)^2}$$

$$= \frac{8054.73-7514}{4220.73-3856.41}$$

$$= \frac{540.73}{364.32}$$

Thus, $\hat{b} = 1.48$.

Now $\hat{a}$ can be calculated using $\hat{b}$ as

$\hat{a} = 121/9 - (1.48\times 62.1/9)$

$= 13.40 - 10.21$

Hence, $\hat{a} = 3.19$

Therefore, the required simple linear regression equation fitted to the given data is

$\hat{Y} = 3.19 + 1.48x$

It should be noted that the value of Y can be estimated using the above fitted equation for the values of x in its range *i.e.*, 3.6 to 10.7.