Linear Regression

Data and calculation:

x	y	$x-\bar{x}$	$y-\bar{y}$	$(x-\bar{x})(y-\bar{y})$	$(x-\bar{x})^2$	$(y-\bar{y})^2$	y'	$y-y'$	$(y-y')^2$
17	94	1.4	14.3	20.02	1.96	204.49	84.061	9.94	98.72
13	73	-2.6	-6.7	17.42	6.76	44.89	71.372	1.63	2.65
12	59	-3.6	-20.7	74.52	12.96	428.49	68.20	-9.26	85.56
15	80	-0.6	0.3	-0.18	0.36	0.09	77.72	2.21	4.88
16	93	0.4	13.3	5.32	0.16	176.89	80.89	12.04	144.96
14	85	-1.6	5.3	-08.48	2.56	28.09	74.55	10.46	110.25
16	66	0.4	-13.7	-05.48	0.16	187.69	80.97	-14.97	223.80
16	79	0.4	-0.7	-02.28	0.16	0.49	80.97	-1.97	3.87
18	77	2.4	-2.7	-6.48	5.76	7.29	87.24	-10.24	104.80
19	91	3.4	11.3	38.42	11.56	127.69	90.41	0.6	0.35
$\bar{x}=$ 15.6	$\bar{y}=$ 79.7			$\Sigma = 134.8$	$\Sigma =$ 42.4	$\Sigma =$ 1206.1			$\Sigma =$ 779.79

Line Equation: $y = a + bx$

slope $\Rightarrow$ $b = r \cdot \frac{s_y}{s_x}$

Here, $s_y = \sqrt{\frac{\Sigma(y-\bar{y})^2}{n-1}}$, $r = \frac{\Sigma(x-\bar{x})(y-\bar{y})}{\sqrt{\Sigma(x-\bar{x})^2 \Sigma(y-\bar{y})^2}}$

$s_x = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}}$

Intercept $\Rightarrow$ $a = \bar{y} - b\bar{x}$

From the above table,

$r = \frac{134.8}{\sqrt{42.4 \times 1206.1}}$ (∵ Applying formula for r)

$$= \frac{134.8}{\sqrt{51,138.69}} = \frac{134.8}{226.139} = 0.596$$

$$S_y = \sqrt{\frac{1206.1}{10-1}} \quad (\because n = \text{number of samples} = 10)$$

$$= \frac{34.729}{3} = 11.576$$

$$S_x = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}} = \sqrt{\frac{42.4}{9}} = \frac{6.511}{3} = 2.171$$

$$b = r \cdot \frac{S_y}{S_x} \Rightarrow 0.596 \times \frac{11.576}{2.171} = 3.173$$

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

$$= 79.7 - 3.173 \times 15.6$$

$$= 79.7 - 49.49$$

$$= 30.123$$

$$y = a + bx = 30.123 + 3.173x \text{ //}$$

sum of squared error = 779.79

Data and calculation:

x	y	$x-\bar{x}$	$y-\bar{y}$	$(x-\bar{x})(y-\bar{y})$	$(x-\bar{x})^2$	$(y-\bar{y})^2$	y'	$y-y'$	$(y-y')^2$
1	1	-2	-1.06	2.12	4	1.12	1.2	-0.2	0.04
2	2	-1	-0.06	0.06	1	0.0036	1.63	0.37	0.14
3	1.3	0	-0.76	0	0	0.58	2.06	-0.76	0.58
4	3.75	1	1.69	1.69	1	2.86	2.49	1.26	1.59
5	2.25	2	0.19	0.38	4	0.04	2.92	-0.67	0.45
$\bar{x}=$ 3	$\bar{y}=$ 2.06			$\Sigma = 4.25$	$\Sigma = 10$	$\Sigma = 4.6$			$\Sigma = 2.8$

$$r = \frac{\Sigma (x-\bar{x})(y-\bar{y})}{\sqrt{\Sigma (x-\bar{x})^2 \Sigma (y-\bar{y})^2}} = \frac{4.25}{\sqrt{10 \times 4.60}} = \frac{4.25}{\sqrt{46}} = \frac{4.25}{6.78} = 0.63$$

$$s_y = \sqrt{\frac{\Sigma (y-\bar{y})^2}{n-1}} = \sqrt{\frac{4.6}{4}} = 1.07 \qquad (\because n = 5)$$

$$s_x = \sqrt{\frac{\Sigma (x-\bar{x})^2}{n-1}} = \sqrt{\frac{10}{4}} = 1.58$$

$$b = r \cdot \frac{s_y}{s_x} = 0.63 \times \frac{1.07}{1.58} = 0.43$$

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

$$= 2.06 - (0.43 \times 3) = 0.77$$

$$y = 0.77 + 0.43x$$

$\therefore$ sum of squared errors = 2.8