S 3807109

2.(a)

X	Y	X^2	Y^2	XY
35.3	10.98	1246.09	120.56	387.59
29.7	11.13	882.09	123.87	330.56
30.8	12.51	948.64	156.50	385.30
58.8	8.4	3457.44	70.56	493.92
61.4	9.27	3769.96	85.93	569.17
71.3	8.73	5083.69	76.21	622.44
74.4	6.36	5535.36	40.45	473.18
76.7	8.5	5882.89	72.25	651.95
70.7	7.82	4998.49	61.15	552.87
57.5	9.14	3306.25	83.53	525.55
46.4	8.24	2152.96	67.89	382.33
28.9	12.19	835.21	148.59	352.29
28.1	11.88	789.61	141.13	333.82
39.1	9.57	1528.81	91.58	374.18
46.8	10.94	2190.24	119.68	511.99
48.5	9.58	2352.25	91.77	464.63
59.3	10.09	3516.49	101.80	598.33
70	8.11	4900	65.77	567.7
70	6.83	4900	46.64	478.1
74.5	8.88	5550.25	78.85	661.56
72.1	7.68	5198.41	59.98	553.72
58.1	8.47	3375.61	71.74	492.10
44.6	8.86	1989.16	78.49	395.15
33.4	10.36	1115.56	107.38	346.02
28.6	11.08	817.96	122.76	316.88
$\Sigma x_i = 1315$	$\Sigma y_i = 235.6$	$\Sigma x_i^2 = 76323.4$	$\Sigma y_i^2 = 2284.11$	$\Sigma xy = 11821.43$

$\Sigma x_i = 1315$ $\quad \Sigma x_i^2 = 76323.4$ $\quad \Sigma xy = 11821.43$

$\Sigma y_i = 235.6$ $\quad \Sigma y_i^2 = 2284.11$

S 3807109

$$S_{xx} = \sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}$$

$$= 76323.42 - \frac{(1315)^2}{25}$$

$$S_{xx} = 7154.42$$

$$S_{yy} = \sum_{i=1}^{n} y_i^2 - \frac{\left(\sum_{i=1}^{n} y_i\right)^2}{n}$$

$$= 2284.11 - \frac{(235.6)^2}{25}$$

$$S_{yy} = 63.82$$

$$S_{xy} = \sum_{i=1}^{n} x_i y_i - \left(\frac{\sum_{i=1}^{n} x_i \times \sum_{i=1}^{n} y_i}{n}\right)$$

$$= 11821.43 - \frac{1315 \times 235.6}{25}$$

$$S_{xy} = -571.13$$

$$\beta = \frac{S_{xy}}{S_{xx}} = \frac{-571.13}{7154.42} = -0.079$$

(slope) $\beta = -0.08$

$\alpha' = \bar{X}_{mean} = \sum_{i=1}^{n} \frac{x_i}{n} = \frac{1315}{25} = 52.6.$

$\bar{Y}_{mean} = \sum_{i=1}^{n} \frac{y_i}{n} = \frac{235.6}{25} = 9.42$

$$\alpha = \bar{Y} - \beta \cdot \bar{X}$$

$$= 9.4 - (-0.08) \times 52.6.$$

$$= 9.4 - (-4.20)$$

$\alpha = 13.6.$

(intercept)

Scanned with CamScanner

S 3807109

Equation of line of best fit

$$\hat{Y} = 13.6 - 0.08X$$

(b)

X	Y	$Y_i = 13.6 - 0.08X$	$Y_i - Y$ (Residual)
35.3	10.98	10.77	−0.20
29.7	11.13	11.22	0.09
30.8	12.51	11.13	−1.37
58.8	8.4.	8.89.	0.49
61.4.	9.27	8.68	−0.58
71.3	8.73	7.89	−0.83.
74.4.	6.36	7.64	1.28
76.7	8.5.	7.46	−1.03
70.7	7.82	7.94	0.12
57.5	9.14	9	−0.14.
46.4.	8.24	9.88	1.64
28.9	12.19	11.28	−0.90
28.1	11.88	11.35	−0.52
39.1	9.57	10.47	0.90
46.8	10.94	9.85	−1.08
48.5	9.58	9.72	0.14
59.3	10.09	8.85	−1.23
70	8.11	8	−0.11
70	6.83	8	1.17
74.5	8.88	7.64.	−1.24.
72.1	7.68	7.83	0.15
58.7	8.47	8.95	0.48
44.6.	8.86	10.03	1.17
33.4.	10.36	10.92	0.56.
28.6	11.08	11.31	0.23

$Y \rightarrow$ observed value

$(Y_i)\ \hat{Y} \rightarrow$ trend value

S 3807109

2.(c)

$(Y_i - \bar{Y})^2$	$(\hat{Y}_i - \bar{Y})^2$	$(Y_i - \hat{Y}_i)^2$
2.43 ~~1.56~~	1.83	0.04.
2.92	3.25	0.008
9.54.	2.94.	1.88
1.04	0.27	0.24.
0.02	0.53.	0.33
0.47	2.32	0.69
9.36	3.13	1.65
0.84	3.82	1.07
2.56	2.17	0.01
0.07	0.17	0.01
1.39	0.21	2.71
7.67	3.48	0.81
6.05	3.73	0.27
0.02	1.10	0.81
2.31	0.19.	1.17
0.02	0.09	0.01
0.44.	0.31	1.52
~~0~~ 1.71	2.01	0.01
6.70	2.01	1.36
0.29	3.16	1.53
3.02	2.52	0.02
0.90	0.21	0.23
0.31	0.37	1.37
0.88	2.27	0.32
2.75	3.57	0.05
63.81	45.80	18.24.

$\bar{Y} \rightarrow$ observed mean $= \frac{235.6}{25} = 9.42$

$F_{\alpha,1,23} = 4.28. \leftarrow F_c$ critical value

Scanned with CamScanner

S 3807109

$$SS_T = \sum_{i=1}^{n} (y_i - \bar{y})^2 = 63.81$$

$$SS_R = \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2 = 45.80$$

$$SS_{Res} = \sum_{i=1}^{n} (y_i - \hat{y})^2 = 18.24$$

$$MS_R = \frac{SS_R}{1} = 45.80$$

$$M.S_{Res} = \frac{SS_{Res}}{n-2} = \frac{18.24}{23} = 0.79$$

$$F_0 = \frac{M.S_R}{MS_{Res}} = \frac{45.8}{0.79} = 57.97$$

$$F_0 = 57.97$$

Since $57.97 > 4.28$ thus we can conclude Regression is significant
$F_0 > F_c$

Source of Variation	Sum of squares	Degree of Freedom	Mean Square	F_0
Regression	SS_R	1	$M.S_R$	MS_R / MS_{Res}
	45.80		45.80	57.97
Residual.	SS_{Res}	n-2	MS_{Res}	
	18.24.	23	0.79	—
Total.	SS_T	n-1		
	63.81	24	—	—

One way Anova is a stastical method that tests

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_K$$

by comparing the variability between groups to the variability within groups.

Scanned with CamScanner

S 3807109

(d) Coefficient of determination $\rightarrow r^2$

$$r^2 = \frac{S^2_{xy}}{S_{xx} \cdot S_{yy}}$$

$$= \frac{(-571.13)^2}{(7154.42)(63.82)}$$

$$= \frac{326189.47}{456595.08}$$

$r^2 = 0.714$

$r^2 = 71.4\%$

Correlation coefficient $\rightarrow r$.

$$r = \frac{S_{xy}}{\sqrt{S_{xx} \cdot S_{yy}}} = \frac{-571.13}{\sqrt{(7154.42)(63.82)}}$$

$r = -0.84$

(r^2) gives information about goodness of fit of a model. Coefficient of determination is statastical measure of how well the regression predictions approximate the real data points. In our case 0.71 indicates good regression prediction.

(r) Correlaton coefficient is a statastical measure of strength of relationship. between relative movements of 2 variables. -ve is indication that both variables move in opposite direction.

Scanned with CamScanner

S 3807109

(e) Standard deviation for the error (s)

$$s = \sqrt{\frac{S_{yy} - \beta \cdot S_{xy}}{n-2}}$$

$$= \sqrt{\frac{63.82 - (-0.08)(-571.13)}{23}}$$

$$= \sqrt{\frac{63.82 - (45.7)}{23}} = \sqrt{\frac{18.12}{23}}$$

$$s = 0.887$$

std for constant (S_α)

$$S_\alpha = s\sqrt{\frac{\Sigma x_i^2}{n \cdot S_{xx}}}$$

$$= 0.89\sqrt{\frac{76323.4}{25 \times 7154.4}} = 0.89\sqrt{\frac{76323.4}{178860}}$$

$$S_\alpha = 0.58$$

std for slope (S_β)

$$S_\beta = \frac{s}{\sqrt{S_{xx}}} = \frac{0.887}{\sqrt{7154.42}} = \frac{0.887}{84.5}$$

$$S_\beta = 0.010$$

S 3807109

(f) Significant test - Slope

$H_0 \rightarrow \beta_0 = 0$, $H_a : \beta_0 \neq 0$

$$se(\hat{\beta}) = \sqrt{\frac{M \cdot S_{Res}}{S_{xx}}} = \sqrt{\frac{0.79}{7154.42}}$$

$$se(\beta) = 0.010$$

$$t = \frac{\beta - \beta_0}{se(\beta)} = \frac{0.08 - 0}{0.010} = 8.$$

$t_{\alpha/2, n-2}$ $(\alpha = 0.05, n = 25)$

$t_{0.25, 23} = 2.069.$

$8 > 2.069$

$\therefore$ slope is significant.

Significant test - intercept

$H_0 : a = a_0$ $H_a : a \neq a_0$

$$se(\hat{a}) = \sqrt{M \cdot S_{Res}\left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right)} = \sqrt{0.79\left(\frac{1}{25} + \frac{(52.6)^2}{7154.42}\right)}$$

$$se(\hat{a}) = \sqrt{0.3318} = 0.57$$

$$t = \frac{a - a_0}{se(a)} = \frac{13.6}{0.57} = 23.8$$

$23.8 > 2.069.$

$t_0 > t_{0.025, 23}$

$\therefore$ Costant is significant

S3807109

(g) Confidence interval for slope and constant

significance level $a = 5\%$

$t_{\alpha/2, n-2} \rightarrow t_{0.025, 23} = 2.068$

CI for slope :

slope $\rightarrow$ $\beta \pm (t_{\alpha/2, n-2}) \times (S_\beta)$ $\leftarrow$ std for slope

$-0.08 \pm (2.068)(0.010)$

-0.08 ± 0.02

$[-0.1, -0.06]$

CI for constant :

$\alpha \pm (t_{\alpha/2, n-2}) \times (S_\alpha)$

$13.6 \pm (2.068)(0.58)$

$13.6 \pm 1.19.$

$[12.41, 14.79]$

Scanned with CamScanner