East West University

Assignment 3

Submitted To:

M H M Imrul Kabir Senior Lecturer

Department of Mathematical and Physical Sciences

Submitted By:

Name: Sheikh Sarafat Hossain

Student ID: 2022-3-60-109

Course Code: STA102

Course Name: Statistics and Probability

Section:04

Semester: Summer - 23

Ans to the Q: No:- 1

Number of customers who visited over the last 23 days in ascending order:

32 32 32 37 37 40 40 40 40 40 42 46 46 46 46 46 46 52 52 52 52 52 52

i) We know,

Coefficient of skewness,

$$= \frac{3 \times (\text{mean} - \text{median})}{\text{standard deviation}}$$

Mean, $\bar{x} = \frac{\Sigma x}{n} = \frac{1000}{23} = 43.478$

Median Position $= \frac{N+1}{2} = \frac{24}{2} = 12^{th}$ value

$\therefore$ Median Value $= 46$ customers

x	$\bar{x}$	$(x-\bar{x})^2$
32		131.74
37		41.96
40		12.096
42	43.478	2.18
46		6.36
52		72.6

$$\sigma = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}} = \sqrt{\frac{(131.74\times3)+(41.96\times2)+(12.096\times5)+2.18+6.36\times6+(72.6\times6)}{23-1}}$$

$$= \sqrt{\frac{1015.56}{22}}$$

$$= 6.79$$

coefficient of skewness $= \dfrac{3 \times (43.478 - 46)}{6.79}$

$= -1.1$

<u>Comment:</u> The shape of the data is left skewed.

ii) Box-Whisker plot:

$Q_1 = \dfrac{(N+1)\times 25}{100} = \dfrac{24 \times 25}{100} = 6^{th}$ value

$= 40$ customers

$Q_3 = \dfrac{(N+1)\times 75}{100} = \dfrac{24 \times 75}{100} = 18^{th}$ value

$= 52$ customers

min $= 32$

max $= 52$

upper limit $= Q_3 + 1.5 \times IQR$

$= 52 + 1.5 \times (52 - 40) = 70$

lower limit $= Q_1 - 1.5 \times IQR$

$= 40 - 1.5 \times (52 - 40) = 22$

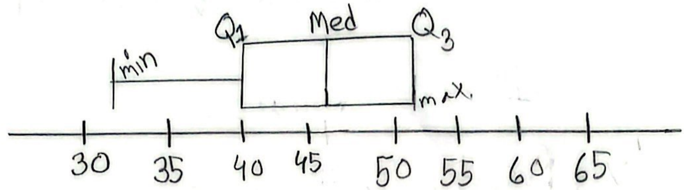

fig: Box-plot diagram of number of customers.

No, there are no outliers. The lower limit of the number of customer is 22.

Ans to the Q: No: 2

i) Estimated median = 450 dollars

ii) Estimated Q_1 = 300 dollars
Estimated Q_3 = 700 dollars

iii) Estimating $IQR = Q_3 - Q_1 = (700 - 300) = 400$ dollars

iv) Beyond upperlimit and lower limit, any value in the data set is considered an outlier.

v) The estimated value of the outlier is 1,500 dollars.

vi) Positively skewed.

Ans to the Q: NO:-3

x	y	xy	$(x-\bar{x})^2$	$(y-\bar{y})^2$
4	4	16	2.56	3.24
5	6	30	0.36	0.04
3	5	15	6.76	0.64
6	7	42	0.16	1.44
10	7	70	19.36	1.44

$\bar{x} = 5.6$

$\bar{y} = 5.8$

$\sum x = 28$

$\sum y = 29$

$\sum xy = 173$

$$S_x = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}} = \sqrt{\frac{29.2}{4}} = 2.702$$

$$S_y = \sqrt{\frac{\Sigma(y-\bar{y})^2}{n-1}} = \sqrt{\frac{6.8}{4}} = 1.304$$

$$S_{xy} = \frac{\Sigma xy - \frac{(\Sigma x)(\Sigma y)}{n}}{n-1} = \frac{173 - \frac{(28)(29)}{5}}{5-1}$$

$$= 2.65$$

Correlation coefficient $= \dfrac{S_{xy}}{S_x S_y}$

$$= \frac{2.85}{(2.702)(1.304)}$$

$$= 0.75$$

Comment: There is a strong positive linear relationship between x and y

Interpretation: Strong positive This is a correlation, which means that higher the x the higher the y's value goes (and vice versa).

Ans to the Q! NO:- 4

x: Sales and y: Earnings.

i) We Know,

estimated regression equation, $\hat{y} = a + bx$

So, $b = \frac{S_{xy}}{S_x^2}$ and $a = \bar{y} - b\bar{x}$

x	y	xy	$(x-\bar{x})^2$	$(y-\bar{y})^2$	$\hat{y}$	$(y-\hat{y})^2$
89.2	4.9	437.08	2169.3	0.49	7.08	4.75
18.6	4.4	81.84	522.28	0.04	2.81	2.53
18.2	1.3	23.66	545.86	8.41	2.79	2.22
71.7	8	573.6	908.27	14.44	6.02	3.92
58.6	6.6	386.76	290.29	5.76	5.23	1.88
46.8	4.1	191.88	27.43	0.01	4.52	0.176
17.5	2.6	45.5	579.9	2.56	2.75	0.0225
11.9	1.7	20.23	879.9	6.25	2.408	0.5013

$\Sigma x = 332.5$, $\Sigma xy = 1760.55$

$\Sigma y = 33.6$, $\Sigma (x-\bar{x})^2 = 6027.27$

$$\Sigma(y-\bar{y})^2 = 37.96 \quad , \Sigma \hat{y} = 33.61$$

$$\Sigma(y-\hat{y})^2 = 15.9998$$

Mean x, $\bar{x} = \frac{\Sigma x}{n} = \frac{332.5}{8} = 41.56$ million dollars.

Mean y, $\bar{y} = \frac{\Sigma y}{n} = \frac{33.6}{8} = 4.2$ million dollars

$$S_{xy} = \frac{\Sigma xy - \frac{(\Sigma x)(\Sigma y)}{n}}{n-1}$$

$$= \frac{1760.55 - \frac{(332.5)(33.6)}{8}}{8-1}$$

$$= 52.007$$

$$S_x^2 = \left(\sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}}\right)^2 = \left(\sqrt{\frac{6027.27}{8-1}}\right)^2 = 861.04$$

Now,

$$b = \frac{52.007}{861.04} = 0.06$$

$a = 4.2 - (0.0604)(41.56) = 1.6896$

Estimated regression equation,

$\hat{y} = 1.6896 + 0.06x$

<u>Comment:</u> On an average, for one earnings increase the estimated sales volume is \$0.06 (ratio).

ii) Give, $x = 50.0$ dollars

Estimated earnings, $\hat{y} = 1.6896 + 0.06(50.0)$

$= 4.71$ million dollars

iii) Standard error, $s_e = \sqrt{\frac{\Sigma(y-\hat{y})^2}{n-2}}$

$= \sqrt{\frac{15.9998}{8-2}}$

$= 1.633$

Comment: On an average the gap between the estimated and actual sales volume is \$1.633 million.

iv) We know,

$$SSE = \sum (y-\hat{y})^2 = 15.9998$$

$$SST = \sum (y-\bar{y})^2 = 37.96$$

$$\text{Coefficient of Determination} = 1 - \frac{SSE}{SST}$$

$$= 1 - \frac{15.9998}{37.96}$$

$$= 0.5785$$

$$= 0.5785 \times 100$$

$$= 57.85\%$$

Comment: The variation in earnings volume is 57.85% explained by the variation in number of sales volume.

Interperetation: The co-efficient of determination indicates how much the linear

relationship can explained. It shows that
the live is located on 0.579.