Midterm Assessment
Sta201

section : 09

Tasnim Sakib Apon
20241068

Midterm Assesment
Sta201
Sec: 09
Tasnim Sakib Apon : 20241068

1.

Class	Frequency (f_i)	Mid value (x_i)	Cumulative Frequency (CF)	$f_i x_i$
51-60	3	55.5	3	166.5
61-70	13	65.5	16	851.5
71-80	5	75.5	21	377.5
81-90	8	85.5	29	684
91-100	29	95.5	58	2769.5
101-110	11	105.5	69	1160.5
111-120	10	115.5	79	1155
121-130	4	125.5	83	502
131-140	14	135.5	97	1897
140-150	5	145.5	102	727.5
	$\Sigma f_i = 102$			$\Sigma f_i x_i \Rightarrow 10291$

~~Mean $\bar{x} = \frac{\Sigma f_i x_i}{n}$~~

$$\text{Mean} \Rightarrow \frac{10291}{102}$$

$$= 100.892.$$

20291068

Median: *

Median class $= \frac{n}{2} = \frac{102}{2}$

$= 51$

So median class => 91 – 100

$L = 91$

$n = 102$

$cf = 29$

$f = 29$

$c = 10$

$\therefore$ Median $M = L + \frac{\frac{n}{2} - cf}{f} \cdot c$

$= 91 + \frac{51-29}{29} \cdot 10$

$= 91 + \frac{22}{29} \cdot 10$

$= 91 + 7.5862$

$= 98.5862.$

20241068

Mode:

Mode class $\Rightarrow$ 91-100

$L = 91$

$f_1 = 29$

$f_0 = 8$

$f_2 = 11$

$c = 10$

so $\text{Mode} = L + \left(\frac{f_1 - f_0}{2 f_1 - f_0 - f_2}\right) c$

$= 91 + \left(\frac{29 - 8}{2.29 - 8 - 11}\right) 10$

$\cancel{= 9 + \left(\frac{21}{39}\right) 10}$

$= 91 + \left(\frac{21}{39}\right) 10$

$= 91 + 5.3846$

$= 96.3846$

[Ans]

20241068

9

	Precious	Semi precious	Unverified
Surface level	0.12	0,26	0,23
Underground	0.35	0.39	0.54
Ocean floor	0.53	0.35	0.23

So, $P(\text{Semi Precious} \mid \text{Ocean}) = \dfrac{P(\text{Semi Precious} \cap \text{Ocean})}{P(\text{Ocean})}$

$[3+9+28) = 40$

$$= \frac{\frac{9}{40} \times 0.35}{\frac{2}{6}}$$

$= 0.23625$

[Ans].

20241068

8 Total countries = 44

Number of ways sushant can choose = $^{44}C_6$

Let,
X = Number of countries choosen by sushant.

$\therefore P(\text{sushant getting treat}) = P(X>=2)$

$= 1 - P(X<2)$

$= 1 - [P(X=0) + P(X=1)]$

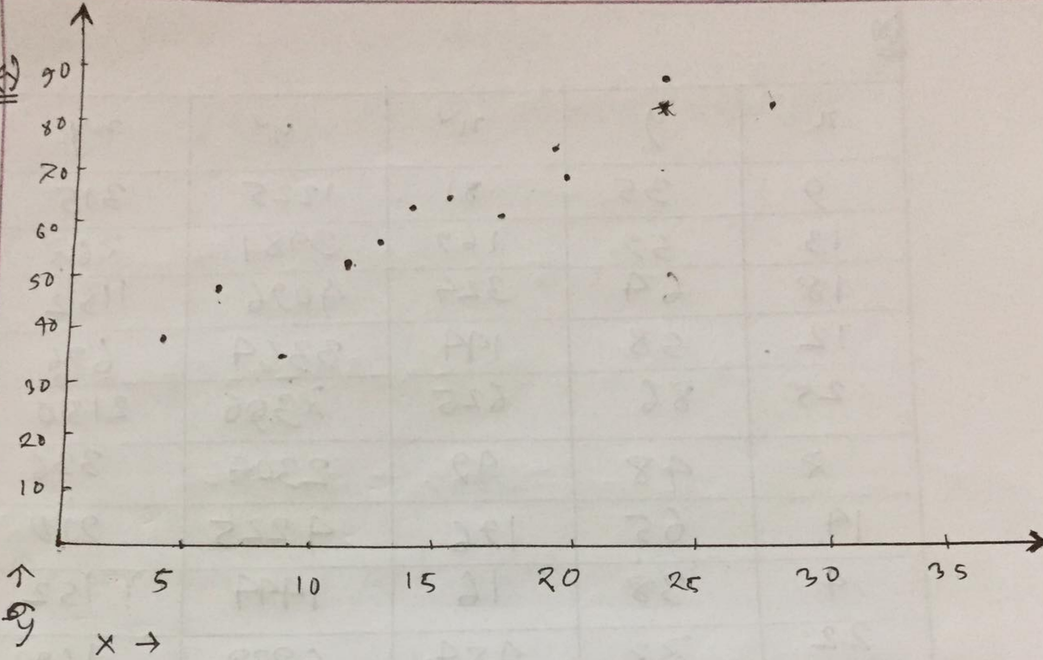
$= 1 - \left(\frac{^{12}C_0 \times {^{32}C_6}}{^{44}C_6} + \frac{^{12}C_1 \times {^{32}C_5}}{^{44}C_6}\right)$

$= 1 - 0.4707$

$= 0.5293$

[Ans]

5(a)



x ⇒ Daily fertilizer use

y ⇒ Plant Length.

As per the scatter plot, the points are upward direction from left to right. Therefore it is a positive correlation.

20241068

(b)

x	y	x^2	y^2	xy
9	35	81	1225	315
13	59	169	3481	767
18	64	324	4096	1152
12	58	144	3364	696
25	86	625	7396	2150
7	48	49	2304	336
14	65	196	4225	910
4	38	16	1444	152
22	77	484	5929	1694
16	68	256	4624	1088
29	80	841	6400	2320
19	72	361	5184	1368
$\Sigma x = 188$	$\Sigma y = 750$	$\Sigma x^2 = 3546$	$\Sigma y^2 = 49672$	$\Sigma xy = 12948$

$$\therefore\ r = \frac{n\,\Sigma xy - \Sigma x.\Sigma y}{\sqrt{n\{\Sigma x^2 - (\Sigma x)^2\}}.\sqrt{n\{\Sigma y^2 - (\Sigma y)^2\}}}$$

$$= \frac{(12\times 12948) - (188\times 750)}{\sqrt{12\times 3546 - (188)^2}\times\sqrt{12\times 49672 - (750)^2}}$$

20291068

$$= \frac{155326 - 141000}{\sqrt{42552 - 35344} \times \sqrt{596064 - 562500}}$$

$$= \frac{14326}{15554.022}$$

$$\therefore r = 0.9243$$

There is a high positive correlation between daily ~~fre~~ fertilizer use and plant length. [Ans]

(c) Mean $\bar{x} = \frac{\sum x}{n} = \frac{188}{12} = 15.6667$

Mean $\bar{y} = \frac{\sum y}{n} = \frac{750}{12} = 62.5$

$$b_{yx} = \frac{n \sum xy - (\sum x)(\sum y)}{n \sum x^2 - (\sum x)^2}$$

$$= \frac{(12 \times 12948) - (188 \times 750)}{(12 \times 3546) - (188)^2}$$

$$= \frac{14326}{7208}$$

$$= 1.9945$$

So, Regression Line:-

$$y-\bar{y}=b_{yx}(x-\bar{x})$$

$$y-62.5=1.9945(x-15.6662)$$

$$\Rightarrow y=1.9945x-31.2464+62.5$$

$$\Rightarrow y=1.9945x+31.2536.$$

[Ans].

(d) For 17gm daily fertilizer

$$\Rightarrow y=1.9945\times 17+31.2536$$

$$=33.9052+31.2536$$

$$\therefore y=65.1593$$

[Ans]

(e) In the table, for 22gm ferlizer use plant length is 72cm.

According to our regression line

$\Rightarrow 1.9945 \times 22 + 31.2536$

$\Rightarrow 75.1326$

~~So~~

Our predicted value is 75.1326 where actual value is 72.

Thus we can say that the regression equation is the best fit for the given data.

4 (a) Putting the numbers in ascending order:

Store 1:

300	400	500	520	545
550	560	590	600	610
620	680	690	700	750
780	790	800	840	860
870	900	920	990	1000

25 data

So, Median will be 13th value which is 690

$\therefore Q_2 = 690$

$\therefore Q_1 = \frac{550+560}{2}$

$= 555$

$\therefore Q_3 = \frac{840+860}{2}$

$= 850$

So,

$Q_1 = 555$

$Q_2 = 690$

$Q_3 = 850$

Minimum value = 300

Maximum value = 1000

Maximum value is 1000

Minimum value is 300.

20241068

For store 2:

400	410	445	445	455	470	490
500	500	510	510	515	540	545
555	570	580	590	600	615	620
655	670	680	700			

25 data.

$\therefore$ Median = 13th value = 540

$\therefore Q_2 = 540$

$$Q_1 = \frac{470 + 490}{2} = 480$$

$$Q_3 = \frac{600 + 615}{2} = 607.5$$

Minimum value = 400

Maximum value = 700

So,

$Q_1 = 480$

$Q_2 = 540$

$Q_3 = 607.5$

Minimum = 400

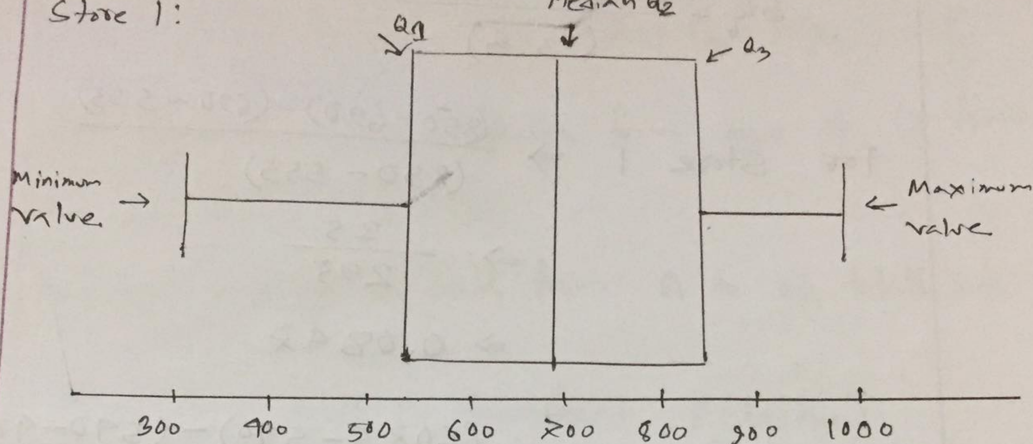
Maximum = 700

[Ans]

20241063.

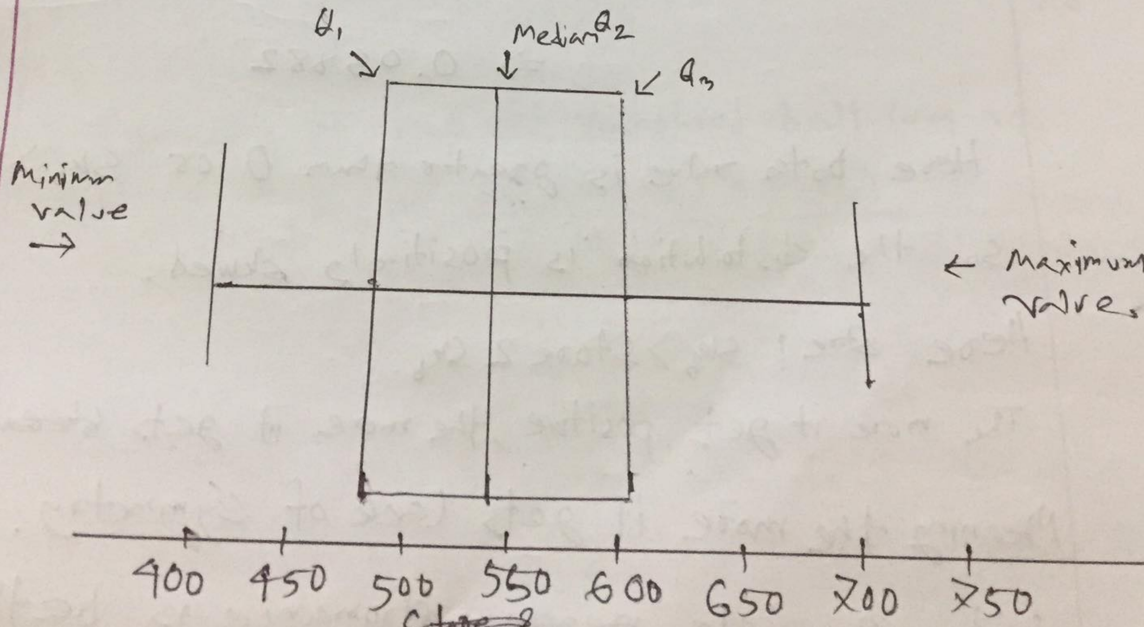
(b) Box and whisker plot.

Store 1:



Store 1

Store 2:



~~Store 2~~

Store 2

We know,

$$Sk_b = \frac{(Q_3-Q_2)-(Q_2-Q_1)}{(Q_3-Q_1)}$$

For Store 1 $\Rightarrow \frac{(850-690)-(690-555)}{(850-555)}$

$\Rightarrow \frac{25}{295}$

$\Rightarrow 0.0847$

Store 2 $\Rightarrow \frac{(602.5-540)-(540-480)}{602.5-480}$

$\Rightarrow \frac{2.5}{122.5}$

$\Rightarrow 0.05882$

Here both value is greater than 0 or $Sk_b > 0$.

So the distribution is positively skwed.

Here Store 1 Sk_b > Store 2 Sk_b.

The more it gets positive, the more it gets ~~skew~~ skwed.

Meaning the more it gets lack of symmetry.

∴ Store 2's sale ~~perof~~ perfromance is better.

20241068

2. Mr parvez invested 2850.

4% by first year = 2850×1.04
$= 2964$

2% decrease by 2nd year = 2964×0.98
$= 2904.72$

7% increase by 3rd year = 2904.72×1.07
$= 3108.0504$

(a) let,
r = growth rate

So, Growth rate of startup after 3 years

$$\Rightarrow 2850\left(1+\frac{r}{100}\right)^3 = 3108.0504$$

$$\Rightarrow 1+\frac{r}{100} = \left(\frac{3108.0504}{2850}\right)^{\frac{1}{3}}$$

$$\Rightarrow 1+\frac{r}{100} = 1.02931$$

$$\Rightarrow \frac{r}{100} = 0.02931$$

$$\therefore r = 2.931\%$$

20241068

So growth rate = 2.931%. [Ans]

(b) 7% increase by 6th year = 3108.0509 × 1.07 × 1.07 × 1.07

= 3807.495386.

[Ans]

3.

Points	Frequency fi	
20-30	1	
30-40	1	
40-50	2	
60-70	4	
70-80	4	
80-90	11	
90-100	24	
	10	
	$\sum fi/n = 57$	

20241068

7. a Given data:

20,000 eligible voters:

(E) Employed => 13200

(Un) Unemployed => 5000

(SE) Self-employed => 1800.

Voted in last presidential election:

(E) employed = 34% = 0.34

(Un) unemployed = 57% = 0.57

(SE) self-employed = 73% = 0.73

~~If a person~~

voted in the last election = (L).

$$\therefore P(L|SE) = \frac{1800 \times 0.73}{(13200 \times 0.34) + (5000 \times 0.57) + (1800 \times 0.73)}$$

$$= \frac{1314}{4488 + 2850 + 1314}$$

$$= \frac{1314}{8652}$$

$$\therefore P(L|SE) = 0.15187$$ [Ans]

20241068

2(b) ~~Number of people who voted in last in last~~

Number of people who voted in last election.

$\Rightarrow \cancel{(1300}$

$\Rightarrow (13200 \times 0.34) + (5000 \times 0.57) + (1800 \times 0.73)$

$\Rightarrow 4488 + 2850 + 1314$

$\Rightarrow 8652.$

Number of people who did not vote $= (20000 - 8652)$

$= 11348.$

Total voters who did not vote and who were employed or self-employed ~~$= 13200 \times 0.66 + 1800 \times 27$~~

$= 13200 \times (1 - .0.33) + (1800 * (1 - .73)$

$= 13200 \times .66 + 1800 \times .27$

$= 9198$

$\therefore P(\text{Eligible voter not being unemployed} \mid \text{did not vote})$

$= \dfrac{9198}{11348} = 0.8105$ [Ans]

20291068

6 we have sample space as follows that is sum of two numbers is 27 from 1 to 27

$S = \{(1,26), (2,25), (3,24), (4,23), (5,22), (6,21), (7,20), (8,19),$
$(9,18), (10,17), (11,16), (12,15), (13,14), (14,13), (15,12),$
$(16,11), (17,10), (18,9), (19,8), (20,7), (21,6), (22,5)$
$(23,4), (24,3), (25,2), (26,1)\}$

Total element = 26

We have $n = 15$, $p = \frac{1}{26} = 0.038$

$q = (1 - 0.038) = 0.962.$

To find $P(x \geq 2) = 1 - [P(x=0) + P(x=1)]$

$P(x=0) = 15C_0 (0.038)^0 (0.962)^{15}$

$= 0.55927586 8$

$P(x=1) = 15C_1 (0.038)^1 (0.962)^{14}$

$= 0.331379672$

Thus $P(x \geq 2) = 1 - [0.559275868 + 0.331379672]$

$= 0.109344$

(Ans)

20241068

3 Section A:

Points	Frequency (f)	Mid value (x)	fx	fx^2
20-30	1	25	25	625
30-40	1	35	35	1225
40-50	2	45	90	4050
50-60	4	55	220	12100
60-70	4	65	260	16900
70-80	11	75	825	61875
80-90	24	85	2040	173400
90-100	10	95	950	90250
	$n = 57$		$\Sigma = 4445$	$\Sigma fx^2 = 360425$

$\therefore$ Mean $\bar{x} = \dfrac{\Sigma x_i}{n}$

$= \dfrac{4445}{57}$

$= 77.9825$

variance $= \dfrac{1}{n}\left(\Sigma fx^2 - \dfrac{(\Sigma fx)^2}{n}\right)$

$= \dfrac{1}{57}\left(360425 - \dfrac{(4445)^2}{57}\right)$

20241068

$= 241.9821$

$\therefore$ standard deviation $= \sqrt{\text{variance}}$

$= \sqrt{241.9821}$

$= 15.5552$

We know,

$$Cv = \frac{S.D}{mean} \times 100$$

$$= \frac{15.5552}{77.9825} \times 100$$

$$= 19.9476\%$$

So, For A, $CV = 19.9476\%$.

20291068

For section B:

Points	Frevency(f)	Mid (x)	f.x	f.x^2
15-25	1	20	20	400
25-35	1	30	30	900
35-45	2	40	80	3200
45-55	2	50	100	5000
55-65	5	60	300	18000
65-75	9	70	630	44100
75-85	17	80	1360	108800
85-95	19	90	1710	153900
	n = 56		$\Sigma fx = 4230$	$\Sigma fx^2 = 334300$

$$\therefore \text{Mean}, \bar{x} = \frac{\Sigma fx}{n} = \frac{4230}{56} = 75.5357$$

$$\text{so Varriance} = \frac{1}{n}\left(\Sigma fx^2 - \frac{(\Sigma fx)^2}{n}\right)$$

$$= \frac{1}{56}\left(334300 - \frac{(4230)^2}{56}\right)$$

$$= 263.9987$$

$\therefore$ Standard deviation $= \sqrt{\text{variance}}$

$= \sqrt{263.9982}$

$\doteq 16.248$

So

$$CV = \frac{SD}{mean} \times 100$$

$$= \frac{16.248}{75.5352} \times 100$$

$$= 21.5103\%.$$

So $CV_A = 19.9426\%$

and $CV_B = 21.5103\%$.

So $CV_B > CV_A$.

Higher coefficient of variation means greater level of depression around the mean. lower CV means more precise and efficient.

$\therefore$ So, section A is more precise and efficient

[Ans]

20241068

(b) We have,

~~$n_A = 57$~~

~~$n_A\, n_1 = 57$~~

For section A,

$$n_1 = 57\ ;\ \bar{x}_1 = 72.9825\ ;\ \sigma_1^2 = 241.9821$$

For B:

$$n_2 = 56\ ;\ \bar{x}_2 = 75.5357\ ;\ \sigma_2^2 = 263.9987$$

So, combined mean $\bar{X} = \dfrac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1 + n_2}$

$$= \frac{(57 \times 72.9825) + (56 \times 75.5357)}{57 + 56}$$

$$= 76.7699$$

Combined variance $= n_1\sigma_1^2 + n_2\sigma_2^2 + n_1|\bar{x}_1 - \bar{x}|^2 + n_2|\bar{x}_2 - \bar{x}|^2$

$= 57 \times 241.9821 + 56 \times 263.9987 + 57\,|72.9825 - 76.7699|^2$

$+ 56\,|75.5352 - 76.7699|^2$

20241068

$$= 254.3845$$

We know,

$$\text{Standard deviation} = \sqrt{\text{variance}}$$

$$= \sqrt{254.3845}$$

$$= 15.9494$$

[Ans]