No.: ASSIGNMENT 2 Date:

Question 1.

a) Mean

$$\text{mean} = \frac{\sum (f \times x)}{\sum f}$$

Class Interval	Frequency	Midpoint (x)	$f \times x$
150-160	12	155	1860
160-170	20	165	3300
170-180	5	175	875
180-190	3	185	555
Total	40		6590

$$\text{mean} = \frac{6590}{40}$$
$$= 164.75 \text{ cm}$$

b) Median

$$\text{median} = L + \left(\frac{\frac{N}{2} - F}{f}\right) \times w$$

$L = 160$ (lower boundary of median class)
$N = 40$ (total frequency)
$F = 12$ (cumulative frequency before median class)
$f = 20$ (frequency of median class)
$w = 10$ (class width)

$$\text{median} = 160 + \left(\frac{20-12}{20}\right) \times 10$$
$$= 160 + \left(\frac{8}{20}\right) \times 10$$
$$= 160 + (0.4) \times 10$$
$$= 160 + 4$$
$$= 164$$

No.: Date:

c) Mode

$$\text{mode} = L + \left(\frac{f_1 - f_0}{2f_1 - f_0 - f_2}\right) \times w$$

$L = 160$
$f_1 = 20$ (modal class frequency)
$f_0 = 12$ (frequency before modal class)
$f_2 = 5$ (frequency after modal class)
$w = 10$

$$\text{Mode} = 160 \left(\frac{20-12}{(20-12)+(20-5)}\right) \times 10$$
$$= 160 + \left(\frac{8}{8+15}\right) \times 10$$
$$= 160 + 0.3478 \times 10$$
$$= 160 + 3.478$$
$$= 163.48$$

d) Modal class

$= 160 < x < 170$

No.: Date:

Assignment 2.

Question 2.

a. Calculate the mean, median and mode of the participation scores.

Original scores:

85, 90, 75, 88, 92, 80, 85, 82, 90, 85

$$\text{Mean} := \frac{85+90+75+88+92+80+85+82+90+85}{10} = \frac{852}{10}$$

$= 85.2$ #

median :

75, 80, 82, 85, 85, 85, 88, 90, 90, 92.

Since there are 10 numbers in the list, the median is the two values in the middle (even), 5th and 6th.

$$\text{median} = \frac{85+85}{2} = 85 \#$$

mode : value that occurs most often.

75, 80, 82, 85, 85, 85, 88, 90, 90, 92.

85 appears 3 times.

mode = 85 #

b. Interpret what these values (mean, median, mode) indicate about the participation performance of the science club members. Based on these values, determine which statistic that more appropriate represent the overall summary of the participation scores.

- Mean (85.2) is slightly higher than the median, showing a balanced participation.
- median (85) half of the students scores below 85, equal to 85 and above 85.
- mode (85) shows most common scores and some of the students participated at this level

All mean, median and mode are close around 85 and this reflects a consistent performance. Therefore, the median and the mode are strong candidate to represent the performance.

No.: Date:

c. Corrected logs, 55, 65, 65, 70, 85, 95, 95, 95, 100, 100

1. Calculate new mean, median and mode

$$\text{mean} = \frac{55+65+65+70+85+95+95+95+100+100}{100}$$

$$= \frac{825}{100} = 82.5 \text{ new mean}_{\#}$$

median =

55, 65, 65, 70, 85, 95, 95, 95, 100, 100

Since there are 10 numbers and its even the median will be in the middle 5th & 6th numbers.

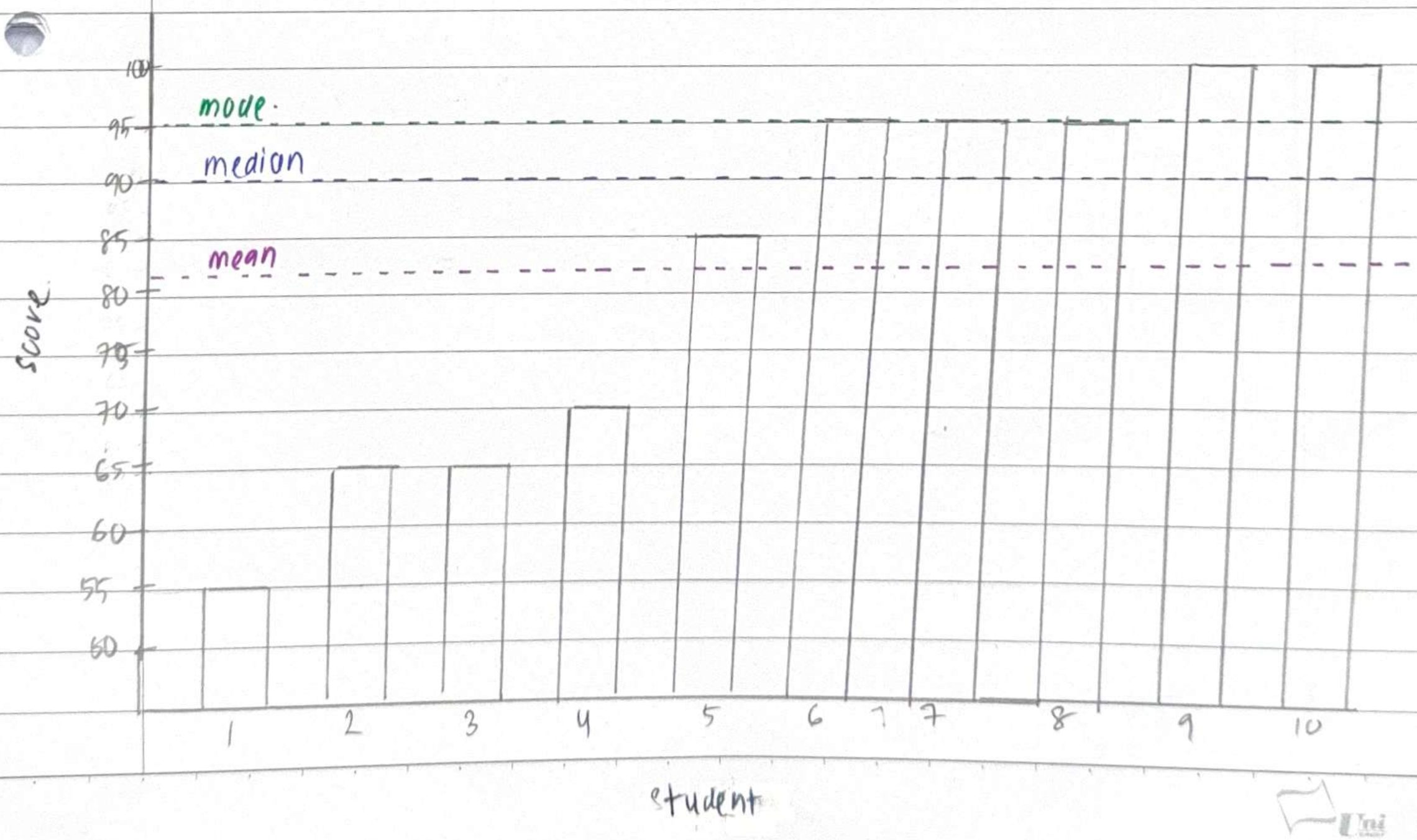
$$\frac{85+95}{2}$$

$$= 90_{\#}$$

mode, the value that most often occurs is 95.

hence the mode is 95 #

11. central tendency graph.



No.: Date:

III. Evaluate the overall participation performance of the science club members based on the corrected scores, and compare it with the original data.

original : mean : 82.5
median : 85
mode : 85

corrected : mean : 82.5
median : 90
mode : 95

The correct data shows extreme variation as some low performers (50-70) and some high performers (85-100).
Hense, median is the best summary for central tendency because it is less affected by extreme values.

No.
Date. / /

Question 3

a) Calculate

i- Range

Range = max - min

= 40 - 25

= $\boxed{15}$

ii- variance

Find mean $\bar{x}$

$= \frac{25 + 30 + 28 + 35 + 32 + 27 + 40 + 38 + 33 + 36 + 31 + 29}{12}$

$= \frac{384}{12}$

$\bar{x} = 32$ variance $= \frac{\sum (x - \bar{x})^2}{n}$

month	x	$x-32$	$(x-32)^2$
January	25	-7	49
February	30	-2	4
March	28	-4	16
April	35	3	9
May	32	0	0
June	27	-5	25
July	40	8	64
August	38	6	36
September	33	1	1
October	36	4	16
November	31	-1	1
December	29	-3	9

No.
Date. / /

$\sum (x - \bar{x})^2 = 49 + 4 + 16 + 9 + 0 + 25 + 64 + 36 + 1 + 16 + 1 + 9$

$= 230$

variance $= \frac{230}{12}$

$= 19.167$

iii) standard deviation

$\sqrt{19.167} \approx 4.378$

b) Interpretation of validity

i. Range (15): Sales fluctuate moderately (25 to 40), indicating seasonal or operational variations.

ii. Variance (19.167) & SD (4.378): Data points deviate from the mean (32) by ~4.38 on average, suggesting moderate consistency.

iii. Key insight: Higher sales in mid-year (July: 40) and lower in January (25) may reflect holiday trends or marketing impacts.

c) Strategic Decisions

i. Inventory Management: Stock more during peak months (July).

ii. Cash flow Planning: Reserve funds for low-sales months (January).

iii. Marketing: Investigate causes of dips (December) and replicate high-performance strategies.

No.: Date:

Assignment 2

(a) - Any scores above 50 indicates an increase.

- Percentage of employees with a score > 50

$P(X>50) = P(Z>50-50/10) = P(Z>0)$ $P(X>50) = P(Z>\frac{50-50}{10}) = P(Z>0)$

Standard normal table:

$P(Z-0) = 0.5$ $P(Z-0) = 0.5$

Answer:- 50% of employees showed an increase in productivity.

(b) Z-scores

For 37

$Z = 37-50/10 = -1.32 = \frac{37-50}{10} = -1.3$

For 65

$Z = 65-50/10 = 1.52 = \frac{65-50}{10} = 1.5$

using 2 tables:-

- $P(Z<1.5) \approx 0.9332$ $P(Z<1.5) \approx 0.9332$
- $P(Z-1.3) \approx 0.0968$ $P(Z<-1.3) \approx 0.0968$

$P(37<X<65) = P(-1.3<Z<1.5) = 0.9332-0.0968 = 0.8364$ $P(37<X<65) = P(-1.3<Z<1.5) =$

$0.9332-0.0968 = 0.8364$

Answer:- The probability is 0.8364 or 83.64%.

Date: .....

No.: .....

(c) Employee scoring < 20:

$z = 20 - 5010 = -3.0z = \frac{20-50}{10} = -3.0$

From z-tables

$P(Z<-3.0) = 0.0013P(Z<-3.0) = 0.0013$

Multiply by total employees

$1000 \times 0.0013 = 1.3 \approx 1$ employee (rounding down) $1000 \times 0.0013 = 1.3 \approx \text{text}$

{employee (rounding down)}

Budget per person = RM200

Estimated cost = 1×200 = RM200 \text{ & Estimated cost } = 1 \times 200 = \text{text } \{200\}

The company should estimate RM200 for team building.

(d) Cut-off scores for the top 5%.

From z-tables

- $P(Z<z) = 0.95 \Rightarrow z \approx 1.645P(Z<z) = 0.95 \Rightarrow z \approx 1.645$

Convert back to x:

$x = \mu + z\sigma = 50 + 1.645 \times 10 = 66.45x = \mu + z\sigma = 50 + 1.645 \times 10 = 66.45$

Answer: An employee needs a minimum score of approximately 66.45 to be in the top 5%.

No: ..... Date: .....

QUESTION 5 (15 MARKS)

In the Information Technology course, students take an exam with 6 mutliple-choice questions. Each question has 4 option, with only one correct answer. The number of correct answers a students gets on the exam represents their perfomance in the cource.

a) Identify the random variable X.

= x is a discreate random variable because it counts the numbers of successes in a fixed numbers of trials. And follows a Binomial Distribution.

b) Construct the table for the probability of x

X	$P(X = x)$	Value
0	$\binom{6}{0}(0.25)^0(0.75)^6$	0.178
1	$\binom{6}{1}(0.25)^1(0.75)^5$	0.356
2	$\binom{6}{2}(0.25)^2(0.75)^4$	0.296
3	$\binom{6}{3}(0.25)^3(0.75)^3$	0.132
4	$\binom{6}{4}(0.25)^4(0.75)^2$	0.033
5	$\binom{6}{5}(0.25)^5(0.75)^1$	0.004
6	$\binom{6}{6}(0.25)^6(0.75)^0$	0.0002

c) What is the mean of this distribution?

$(\mu = np;\ \sigma^2 = npq)$

$n = 6$ (number of questions)

$p = \frac{1}{4} = 0.25$ (probility of correctly)

$q = 1 - p = 0.75$

$\mu = n \times p = 6 \times 0.25 = 1.5$

Mean = 1.5 #

No ..... Date .....

d) Determine the probability of a student getting at least 3 correct answers

At least 3 $= X \geq 3$

$P(X \geq 3) = P(3) + P(4) + P(5) + P(6)$

$P(X \geq 3) = 0.132 + 0.033 + 0.004 + 0.00024 = 0.16924$ #

e) What is the probbility that the first incorrect answer occurs in the fourth question?

- probability of failure $(q) = 1 - p = 0.75$
- probability of success $(p) = 0.25$

$P^3 \cdot q = (0.25)^3 \times 0.75 = 0.0117$

The probability is 0.0117

No: ..... Date: .....

Question 6 (10 MARKS)

In a coffee shop, customers arrive at random intervals throughout the day. The shop owner is interested in understanding the arrival patterns to better manage staffing levels. The probability of a customer ordering a cappuccino is 0.70. X represent the numbers of customers that will arrive until the 4th customer who orders a cappuccino.

a) what is probability distribution of X?

Negative Binomial distribution with parameters.

- Number of success (k) : 4
- Probability (p) : 0.70

b) What is the standard derivation of this distribution?

Mean (μ) : $\frac{k}{p} = \frac{4}{0.70} = 5.714$.

Variance (σ^2) : $\frac{k \cdot q}{p^2} = \frac{4 \times 0.30}{(0.70)^2} = \frac{1.2}{0.49} = 2.449$

Standard Derivation (σ) : $\sqrt{2.449} = 1.565$ #

No: ..... Date: .....

c Determine the probability that exactly 6 customers will arrive until the 4th customer who orders cappuccino.

Use the Negative Binomial PMF

$x = 6$ $\quad P(X=6) = \binom{6-1}{4-1}(0.70)^4(0.30)^{6-4} = \binom{5}{3}(0.70)^4(0.30)^2$

$k = 4$

$= 10 \times 0.2401 \times 0.09 = 0.216$ #

d) Determine the probability that exactly 7 cappuccinos will be ordered within the next 12 customer arrivals.

Use the Binomial Distribution.

number of trials ($n = 12$ customers)

Probability of success ($p = 0.70$)

$P(Y=7) = \binom{12}{7}(0.70)^7(0.30)^5$

$= 792 \times 0.082354 \times 0.00243 = 0.158$ #