



UTM
UNIVERSITI TEKNOLOGI MALAYSIA

PROBABILITY & STATISTICAL DATA ANALYSIS

SECTION-02

SECI2143

20202021/2

GROUP/ASSIGNMENT :

GROUP 4(HIV)/ASSIGNMENT 1

LECTURER'S NAME :

DR CHAN WENG HOWE

GROUP MEMBERS :

NAME	MATIRC NO.
HONG PEI GEOK	A20EC0044
LOW JUNYI	A20EC0071
TAN YONG SHENG	A20EC0157
VINCENT BOO EE KHAI	A20EC0231

Question 1 :

1. a) Population : Football club in Malaysia
b) Sample : Malaysia Super League
c) Discrete variable: total number of goals scored
d) Continuous variable: Amount of time
e) Primary data: 2.97 (number of goals scored recorded)
f) Secondary data: 3.08 (old newspaper articles showed)

2. a) Population : Student at the school in JayBie
b) Sample : Second year pupils
c) Discrete variable: Place of Birth, gender, age (in year)
d) Continuous variable: height, weight
e) Primary data : Weight of second year pupils
f) Secondary data: Data of first year students

Question 2

1.)

Scales Factors	Frequency				Total
	1	2	3	4	
Price	14	6	0	0	20
Condition of Car	3	6	11	0	20
Fuel Efficiency	1	11	7	1	20
Car Depreciation	6	11	2	1	20

10

2.) Price

Scale	Frequency	Relative Frequency	Cumulative Frequency	Relative Cumulative Frequency
1	14	0.7	14	0.7
2	6	0.3	20	1.0
3	0	0.0	20	1.0
4	0	0.0	20	1.0
Total	20	1.0		

Condition of car

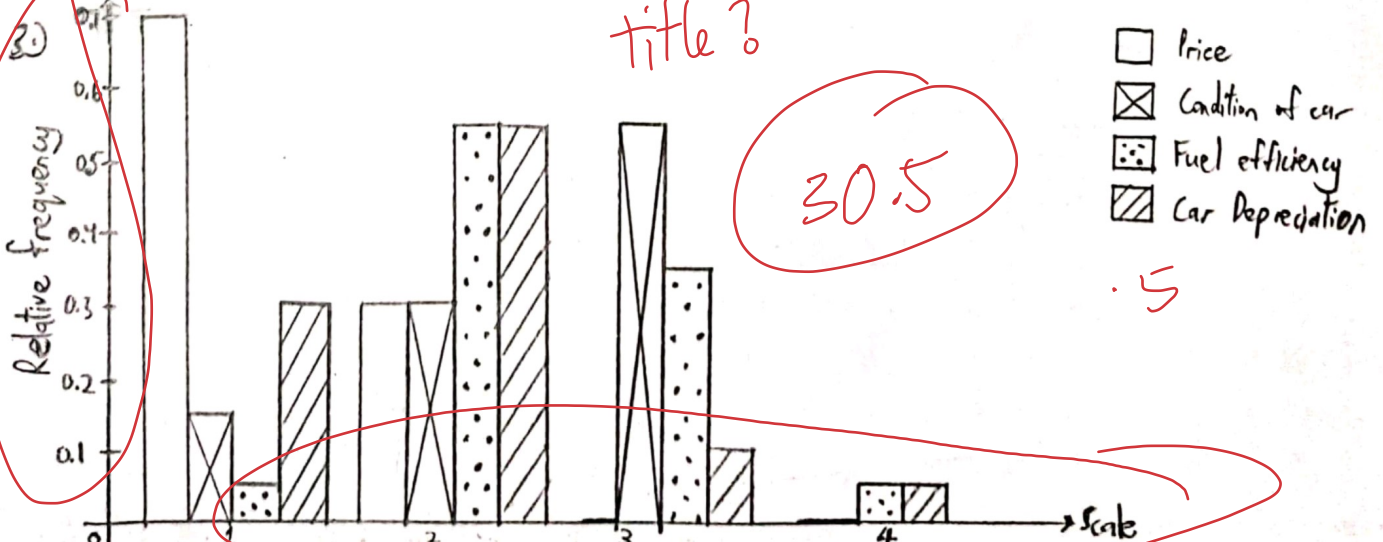
Scale	Frequency	Relative Frequency	Cumulative Frequency	Relative Cumulative Frequency
1	3	0.15	3	0.15
2	6	0.30	9	0.45
3	11	0.55	20	1.00
4	0	0.00	20	1.00
Total	20	1.00		

Fuel Efficiency

Scale	Frequency	Relative Frequency	Cumulative Frequency	Relative Cumulative Frequency
1	1	0.05	1	0.05
2	11	0.55	12	0.60
3	7	0.35	19	0.95
4	1	0.05	20	1.00
Total	20	1.00		

Car Depreciation

Scale	Frequency	Relative Frequency	Cumulative Frequency	Relative Cumulative Frequency
1	6	0.30	6	0.30
2	11	0.55	17	0.85
3	2	0.10	19	0.95
4	1	0.05	20	1.00
Total	20	1.00		



Question 3:

4 8 9 16 18 21 23 24 30 32 33 38 42 43 44 55 65 81

1.

Stem	Leaf
0	4, 8, 9
1	6, 8
2	1, 3, 4
3	0, 2, 3, 8
4	2, 3, 4
5	5
6	5
7	
8	1

Key: 014 means 4 days

2. a) Mean = $\frac{4+8+9+16+18+21+23+24+30+32+33+38+42+44+55+65+81}{18}$
 $= 32.56$

b) Mode = no mode

c) Median = $\frac{30+32}{2}$
 $= 31$

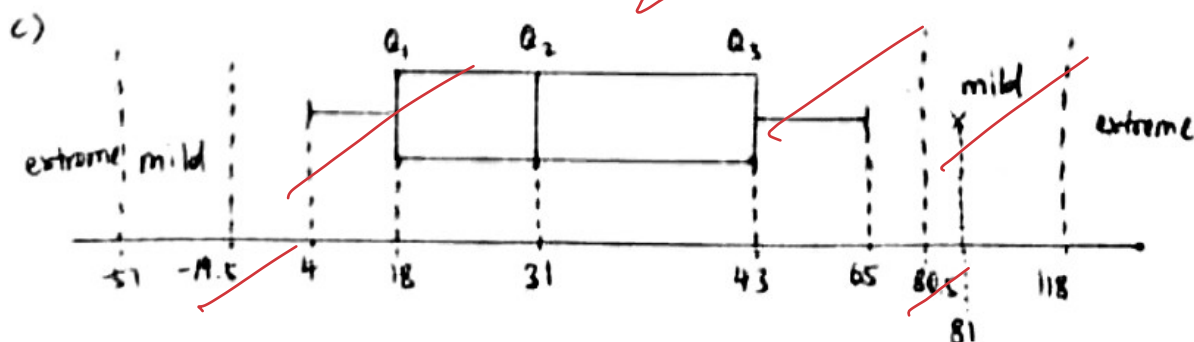
3. a) $Q_1 = \frac{18}{4}$
 ≈ 5
 $Q_1: 18$

$Q_2 = 31$

$Q_3 = 18 \times \frac{3}{4}$
 $= 13.5$
 ≈ 14

$Q_3: 43$

b) IQR = $Q_3 - Q_1$
 $= 43 - 18$
 $= 25$



mild = $1.5 \times \text{IQR}$
 $= 37.5$

extreme = $3 \times \text{IQR}$
 $= 75$

Lower mild = $18 - 37.5 = -19.5$
 Lower extreme = $18 - 75 = -57$
 Upper mild = $43 + 37.5 = 80.5$
 Upper extreme = $43 + 75 = 118$

Question 4

1.
$$\bar{x} = \frac{175\,000 + 250\,000 \times 5 + 500\,000 (4) + 700\,000}{11}$$
$$= \frac{4\,125\,000}{11}$$
$$\bar{x} = \text{RM } 375\,000$$

No. of houses = $1 + 5 + 4 + 1$
 $= 11$

2.
$$\bar{j} = \frac{50}{100} (11)$$
$$j = 5.5$$
$$k = 6$$

$Q_2 / \text{median} = \text{RM } 250\,000$

3. Mean. This is because mean is the average value of the houses while median is only show the middle house's value.

Question 5

18.5

1.

Age	Ratio
Current smoker	Nominal
Body Mass Index (BMI)	Interval
Hypertension	Nominal

1.5

2.

Class Interval	Category	Frequency
$0 < \text{BMI} < 18.5$	Underweight	0
$18.5 \leq \text{BMI} < 25.0$	Normal	12
$25.0 \leq \text{BMI} < 30.0$	Overweight	7
$30.0 \leq \text{BMI}$	Obesity	1

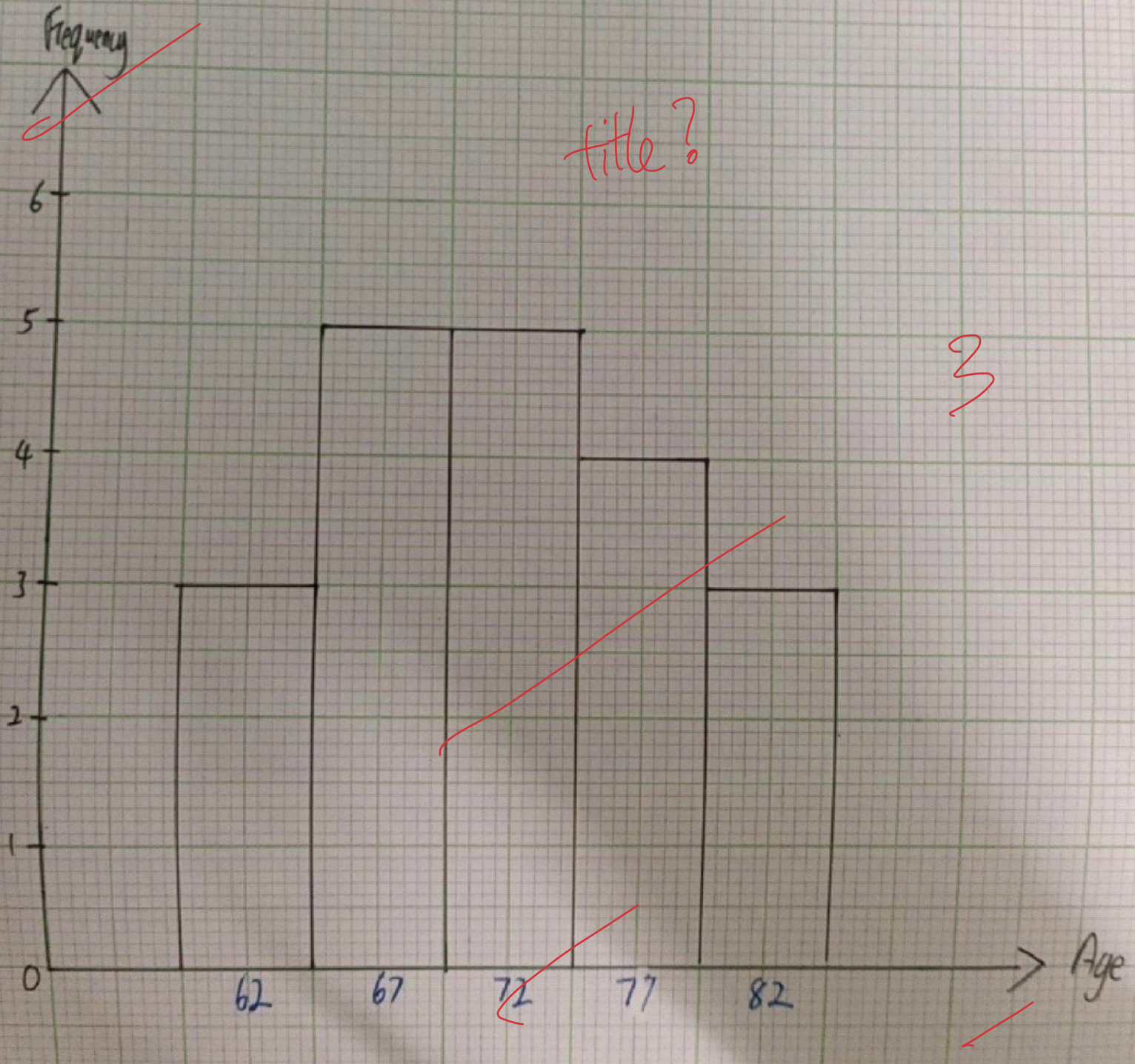
2

3.

Class Interval	Class Boundaries	Class Midpoint	Frequency	Cumulative Frequency, F
60 - 64	57.5 - 64.5	$\frac{60+64}{2} = 62$	3	3
65 - 69	64.5 - 69.5	$\frac{65+69}{2} = 67$	5	8
70 - 74	69.5 - 74.5	$\frac{70+74}{2} = 72$	5	13
75 - 79	74.5 - 79.5	$\frac{75+79}{2} = 77$	4	17
80 - 84	79.5 - 84.5	$\frac{80+84}{2} = 82$	3	20
Total			20	20

12

5.



Question 6

Time (minutes)	class boundaries	frequency	cumulative frequency	midpoint
16-30	15.5-30.5	3	3	23
31-45	30.5-45.5	13	16	38
46-60	45.5-60.5	30	46	53
61-75	60.5-75.5	25	71	68
76-90	75.5-90.5	14	85	83
91-105	90.5-105.5	8	93	98
106-120	105.5-120.5	4	97	113
121-135	120.5-135.5	2	99	128
136-150	135.5-150.5	1	100	143

20

$$1. \text{median class} = \frac{100}{2}$$

$$= 50^{\text{th}}$$

$$= 61-75$$

$$\text{median} = L + \frac{\frac{N}{2} - cf}{f_{\text{med}}} (W)$$

$$= 60.5 + \frac{\frac{100}{2} - 46}{25} (15)$$

$$= 62.9 \text{ minutes}$$

$$\bar{x} = \frac{\sum fx}{\sum f}$$

$$= \frac{3(23) + 13(38) + 30(53) + 25(68) + 14(83) + 8(98) + 4(113) + 2(128) + 1(143)}{100}$$

$$= \frac{6650}{100}$$

$$= 66.5 \text{ minutes}$$

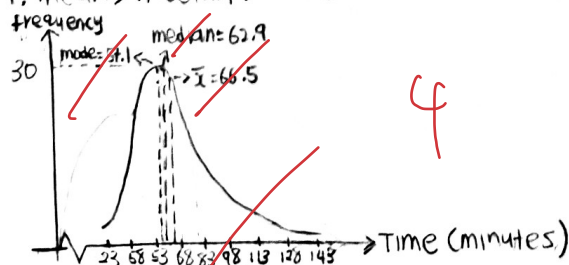
$$\text{mode class} = 46-60$$

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

$$= 45.5 + 15 \left(\frac{30 - 13}{2(30) - 13 - 25} \right)$$

$$= 57.1 \text{ minutes}$$

$$2. 9. \text{mean} > \text{median} > \text{mode}$$



b. Positively skewed / skewed to the right and the kurtosis is leptokurtic. (mean > median > mode)

c. Since average delay time of flight is higher than median delay time of flight, thus more than half of flight with delay time less than the average delay time of flight. There are more flight with less delay time than flight with high delay time.