

Question 1

1.
 - a) Malaysian Super League matches in a season.
 - b) Home win, away win or ~~draw~~ draw.
 - c) Number of goals scored in a match.
 - d) Amount of time played before a goal is scored.
 - e) Number of goals scored in new season.
 - f) ~~mean number of goals per match.~~
Number of goals scored in past season.

2.
 - a) All students at a school in Jaykie
 - b) First year pupils
 - c) Age
 - d) Height.
 - e) Data recorded for each pupil — birth, gender, age, weight, height.
 - f) Weight of the second year students.

Question 2

1.

Factors	Scales	Frequency				Total
		1	2	3	4	
Price		14	6	0	0	20
condition of the car		3	6	11	0	20
Fuel efficiency		1	11	7	1	20
Car Depreciation		6	11	2	1	20

2 -

Factor: Price

scale	Frequency	Relative frequency	Cumulative frequency	Cumulative relative frequency
1	14	0.7	14	0.7
2	6	0.3	20	1.0
3	0	0	20	1.0
4	0	0	20	1.0
Total	20	1.0	-	-

Factor: condition of the car

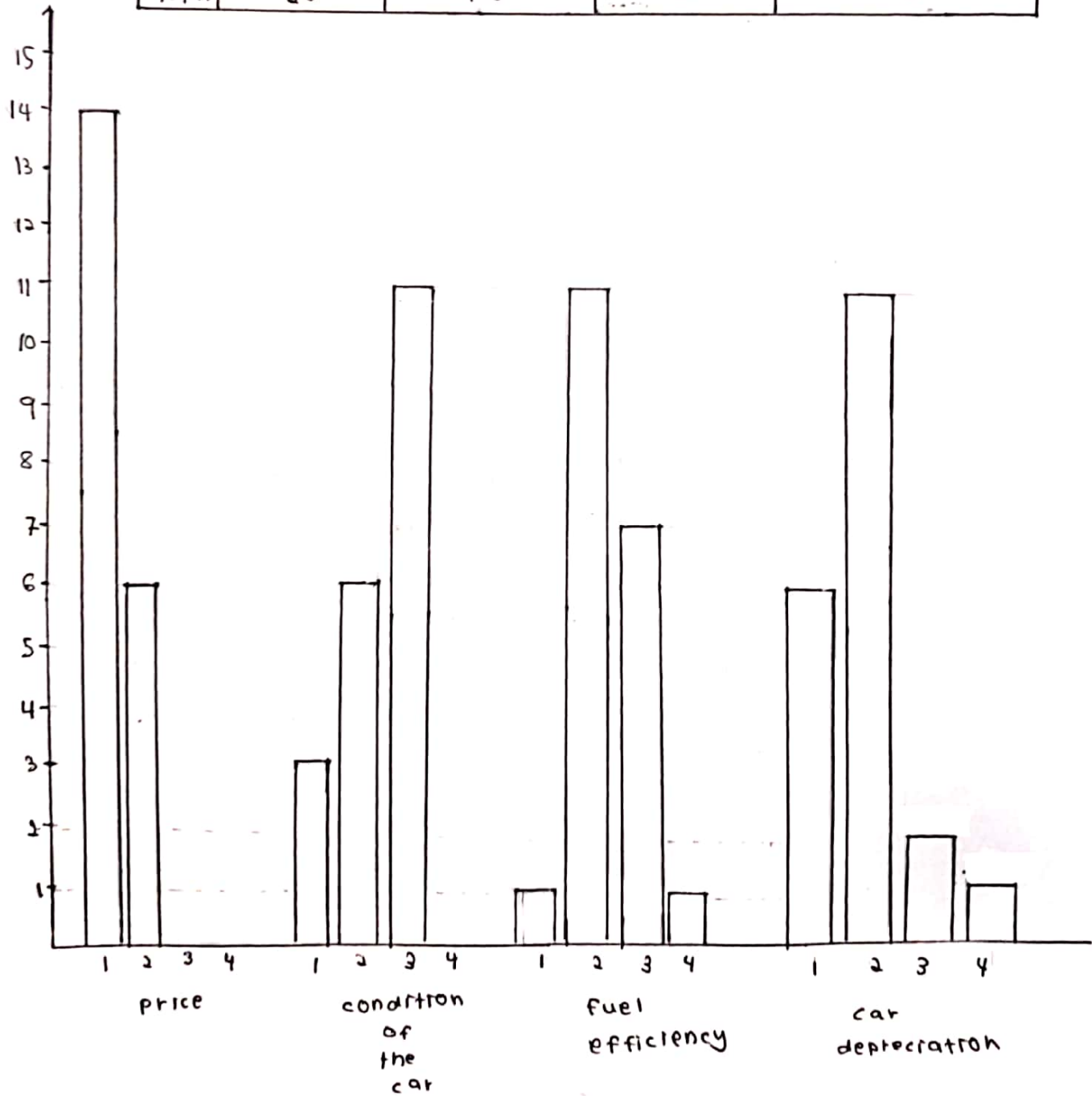
scale	Frequency	Relative frequency	Cumulative frequency	Cumulative relative freq
1	3	0.15	3	0.15
2	6	0.3	9	0.45
3	11	0.55	20	1.0
4	0	0	20	1.0
total	20	1.0	-	-

Factor: Fuel efficiency

scale	frequency	relative frequency	Cumulative freq	Cumulative relative freq
1	1	0.05	1	0.05
2	11	0.55	12	0.6
3	7	0.35	19	0.95
4	1	0.05	20	1.0
total	20	1.0	-	-

Factor : car depreciation

scale	frequency	relative frequency	cumulative frequency	cumulative relative frequency
1	6	0.3	6	0.3
2	11	0.55	17	0.85
3	2	0.1	19	0.95
4	1	0.05	20	1.0
total	20	1.0	-	-



QUESTION 3

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

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: 1 | 6 → 16

2. 4, 8, 9, 16, 18, 21, 23, 24, 30, 32, 33, 38, 42, 43, 44, 55, 65, 81

a) median

$$\text{mean} = \frac{4+8+9+16+18+21+23+24+30+32+33+38+42+43+44+55+65+81}{18} = \frac{293}{9}$$

$$1' = 32-556$$

b) mode = no mode

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

3. a.

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

$\downarrow$ $\downarrow$ $\downarrow$
 Q_1 Q_2 Q_3

a) 1st quartile, $Q_1 = Y[5] = 18$

2nd quartile $= \frac{30+32}{2} = 31$

3rd quartile, $Q_3 = Y[14] = 43$

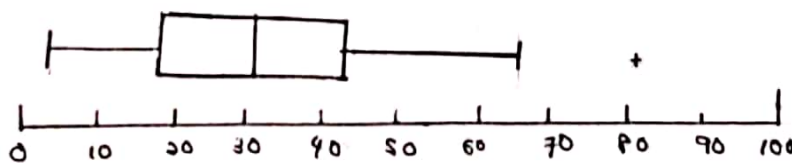
b. Interquartile range (iqr) $= Q_3 - Q_1$
 $= 43 - 18$
 $= 25$

c. maximum
whisker
length (mwl) $= iqr \times 1.5 = 37.5$

inner:
lower fence $= Q_1 - mwl = 18 - 37.5$
 $= -19.5$

inner:
upper fences $Q_3 + mwl = 43 + 37.5$
 $= 80.5$

nothing is below lower fence, but 81 is above upper fence.
 Thus 81 is a possible outlier, so ~~the~~ the maximum
 of remaining data, 65, is the new maximum.



upper outer fence $= Q_3 + 3(iqr) = 43 + 3(25)$
 $= 118$

$\therefore$ mild outlier

Question 4

1. $\text{mean} = \frac{\sum fx}{n}$

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

2. numbers of observation = $1 + 5 + 4 + 1$
 $= 11$ (odd number)

$$\therefore \text{median} = \left[\frac{(11+1) \div 2}{} \right]$$
$$= 6^{\text{th}} \text{ place}$$

$$\text{median} = \text{RM } 250\ 000$$

3. To measure the central tendency, mean is the best to represents the value of the 11 houses because mean is the average of all 11 houses. Hence, it includes all the values of the houses. ~~while~~ median

Question 5

- 1) • Age = nominal ratio
• Current Smoker = nominal
• BMI = interval
• Hypertension = nominal

2)

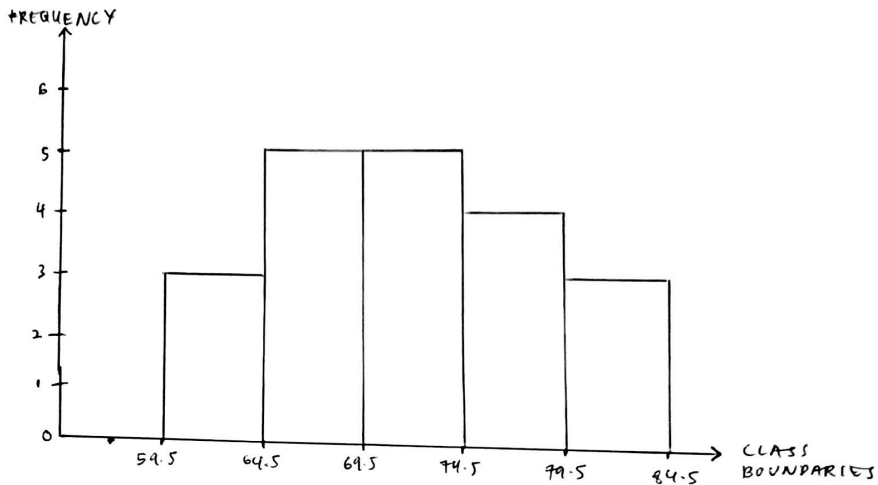
CLASS INTERVAL	CATEGORY	FREQUENCY
$0 < \text{BMI} < 18.5$	Under Weight	0
$18.5 \leq \text{BMI} < 25.0$	Normal	12
$25.0 \leq \text{BMI} < 30.0$	Over Weight	7
$30.0 \leq \text{BMI}$	obesity	1

3)

CLASS INTERVAL	CLASS BOUNDARIES	CLASS MIDPOINT	FREQUENCY	CUMULATIVE FREQUENCY
60 - 64	59.5 - 64.5	62	3	3
65 - 69	64.5 - 69.5	67	5	8
70 - 74	69.5 - 74.5	72	5	13
75 - 79	74.5 - 79.5	77	4	17
80 - 84	79.5 - 84.5	82	3	20
TOTAL		360	20	55 61

Question 5

4.)



Question 6

Class interval	Midpoint	Frequencies	Cumulative frequencies	$f_i x_i$
16 - 30	23	3	3	69
31 - 45	38	13	16	494
46 - 60	53	30	46	1590
61 - 75	68	25	71	1700
76 - 90	83	14	85	1162
91 - 105	98	8	93	784
106 - 120	113	4	97	452
121 - 135	128	2	99	256
136 - 150	143	1	100	143

$$\begin{aligned} \text{median} &= 60.5 + \frac{50 - 46}{25} (15) \\ &= 62.9 \text{ minutes} \end{aligned}$$

$$\begin{aligned} \text{mean} &= \frac{\sum f_i x_i}{\sum f} \\ &= \frac{6650}{100} \\ &= 66.5 \text{ minutes} \end{aligned}$$

$$\begin{aligned} \text{mode} &= 45.5 + (15) \left[\frac{(30 - 13)}{(2 \times 30 - 13 - 25)} \right] \\ &= 57.09 \text{ minutes} \end{aligned}$$



b) skewness of the distribution is positively skewed

$$\text{kurtosis} = \frac{\sum (x_i - \bar{x})^4 (N-1)}{(\sum (x_i - \bar{x})^2)^2}$$

$$= 22.79 \quad \text{leptokurtic}$$

c) the data does not have a symmetrical distribution about its mean and have a high degree of peakness of probability distribution relative to the benchmark of normal distribution