UTM

UNIVERSITI TEKNOLOGI MALAYSIA

87

SECI2143-01

PROBABILITY & STATISTICAL DATA ANALYSIS

ASSIGNMENT 1

NAME : HUDA NAJIHAH BINTI AHMAD ASRI (A20EC0045)
MARNISHA BINTI MUSTAFA KAMAL (A20EC0075)
NUR HAZNIRAH BINTI HAZMAN (A20EC0114)
LECTURER : DR. SHARIN HAZLIN BINTI HUSPI

QUESTION 1

5

(1)
a) football clubs in Malaysiana
b) Malaysiana Super League matches
c) the total number of goals scored
d) the amount of time played before a goal is scored
e) the mean number of goals per match as 2.97
f) the mean number of goals per game was 3.08

2.5

(2)
a) Second year pupils
b) the second year pupils weighed
c) the number of pupils weighed
d) height
e) the weight of second year pupils
f) the data on first year pupils

2.5

QUESTION 2

23

① 10

Scales / Factors	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 5

Factor Price

Scale	Frequency	Relative Frequency	Cumulative Frequency	Cumulative Relative Frequency
1	14	14 ÷ 20 = 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		

Factor Condition of the Car

Scale	Frequency	Relative Frequency	Cumulative Frequency	Cumulative Relative 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		

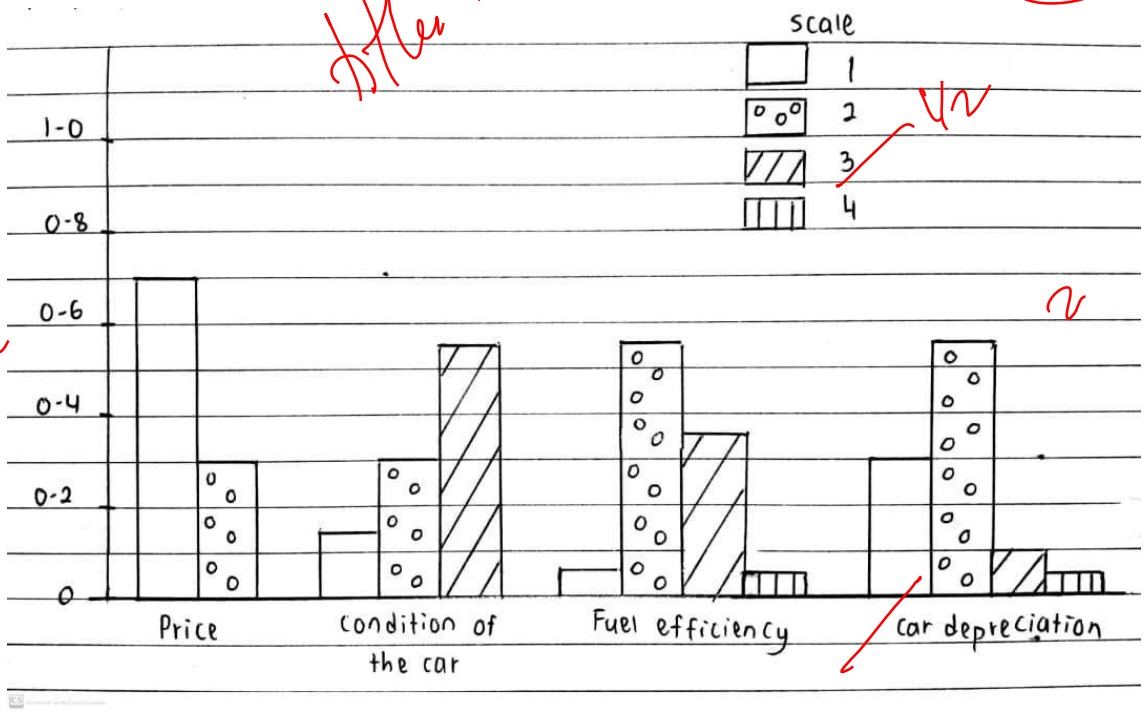
Factor Fuel Efficiency

Scale	Frequency	Relative Frequency	Cumulative Frequency	Cumulative Relative 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		

Factor Car Depreciation

Scale	Frequency	Relative Frequency	Cumulative Frequency	Cumulative Relative 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 2


scale
1
2
3
4
1·0
0·8
0·6
0·4
0·2
0
Price
condition of the car
Fuel efficiency
car depreciation

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: 5|5 = 55 days

2) a) Mean, $\bar{x} = \frac{\Sigma x}{n}$

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

$= \frac{586}{18} = 32.56$

b) Since each value occurs only once in the data, hence there is no mode for this set of data.

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

Median = $\frac{30+32}{2}$

= 31

QUESTION 3

3) a) $i = \frac{p}{100}(N)$

$Q_1 = \frac{25}{100}(18)$
$= 4.5$
$k \approx 5$
$Y[5] = 18$

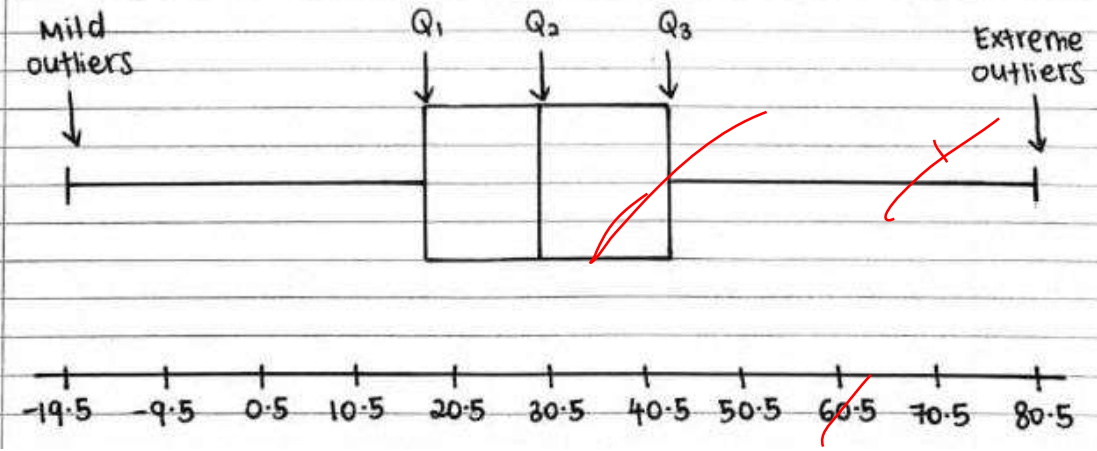
$Q_2 = \frac{50}{100}(18)$
$k = 9$
$Y[9] = 30$

$Q_3 = \frac{75}{100}(18)$
$= 13.5$
$k \approx 14$
$Y[14] = 43$

b) $IQR = Q_3 - Q_1$
$= Y[14] - Y[5]$
$= 43 - 18$
$= 25$

c) Mild outliers $= Q_1 - 1.5(IQR)$
$= 18 - 1.5(25)$
$= -19.5$

Extreme outlier $= Q_3 + 1.5(IQR)$
$= 43 + 1.5(25)$
$= 80.5$



QUESTION 4

1) Mean, $\bar{x} = \frac{\sum fx}{n}$

$= \frac{(175000 \times 1) + (250000 \times 5) + (500000 \times 4) + (700000 \times 1)}{11}$

$= \frac{4125000}{11}$

$= RM375000$

4.5

2) Median $= \frac{N}{2} = \frac{11}{2}$

$= 5.5$

$k \approx 6$

$Y[6] = RM250000$

3) The mean because it is the average value of the 11 houses

QUESTION 5

1.

Age	Ratio
Current Smoker	Nominal
Body Mass Index (BMI)	Ratio
Hypertension	Nominal

2

20

2.

Class Interval	Category	Frequency
0 < BMI < 18.5	Underweight	0
18.5 <= BMI < 25.0	Normal	12
25.0 <= BMI < 30.0	Overweight	7
30.0 <= BMI	Obesity	1

2

3.

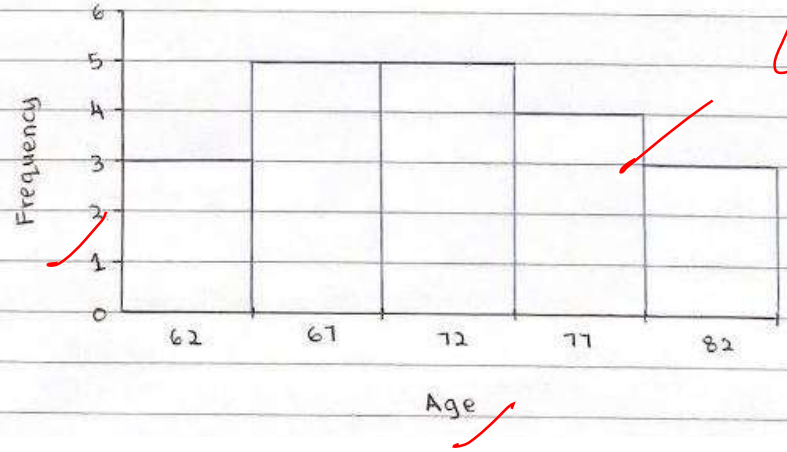
Class Interval	Class Boundaries	Class Midpoint	Frequency	Cumulative Frequency
60-64	59.5-64.5	62	3	3
65-69	64.5-69.5	67	5	8
70-74	69.5-74.5	72	5	13
75-79	74.5-79.5	77	4	17
80-84	79.5-84.5	82	3	20
Total		360	20	

10

2

4.

Histogram of Age



4

QUESTION 6

1. Time (Minutes)	Midpoint	Frequencies	fx	Relative Frequency
16-30	(16+30)/2 = 23	3	23x3 = 69	3
31-45	(31+45)/2 = 38	13	38x13 = 494	16
46-60	(46+60)/2 = 53	30	53x30 = 1590	46
61-75	(61+75)/2 = 68	25	68x25 = 1700	71
76-90	(76+90)/2 = 83	14	83x14 = 1162	85
91-105	(91+105)/2 = 98	8	98x8 = 784	93
106-120	(106+120)/2 = 113	4	113x4 = 452	97
121-135	(121+135)/2 = 128	2	128x2 = 256	99
136-150	(136+150)/2 = 143	1	143x1 = 143	100
Total		100	6650	

Median, $N/2 = 100/2 = 50$

Median class = 61-75

$$\text{Median} = L + \frac{N/2 - cf_p}{f_{med}}(w)$$

$$= 60.5 + \frac{50-46}{25}(15)$$

$$= 62.9$$

Mean, $\bar{x} = \frac{\sum fx}{n}$

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

$$= 66.5$$

Mode, Modal class = 46-60

$$\text{Mode} = L + h \times [(f_1 - f_0) \div (2f_1 - f_0 - f_2)]$$

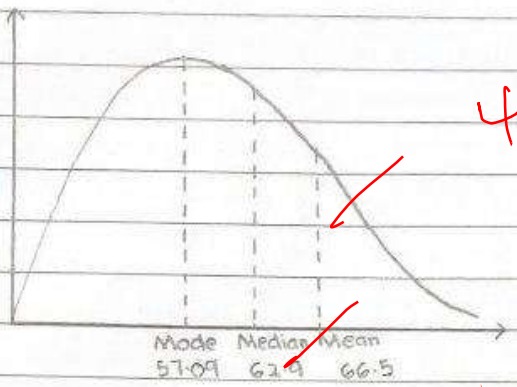
$$= 45.5 + 15 \times [(30-13) \div (2(30)-13-25)]$$

$$= 45.5 + 15 \times (17 \div 22)$$

$$= 57.09$$

QUESTION 6

2. a.



4

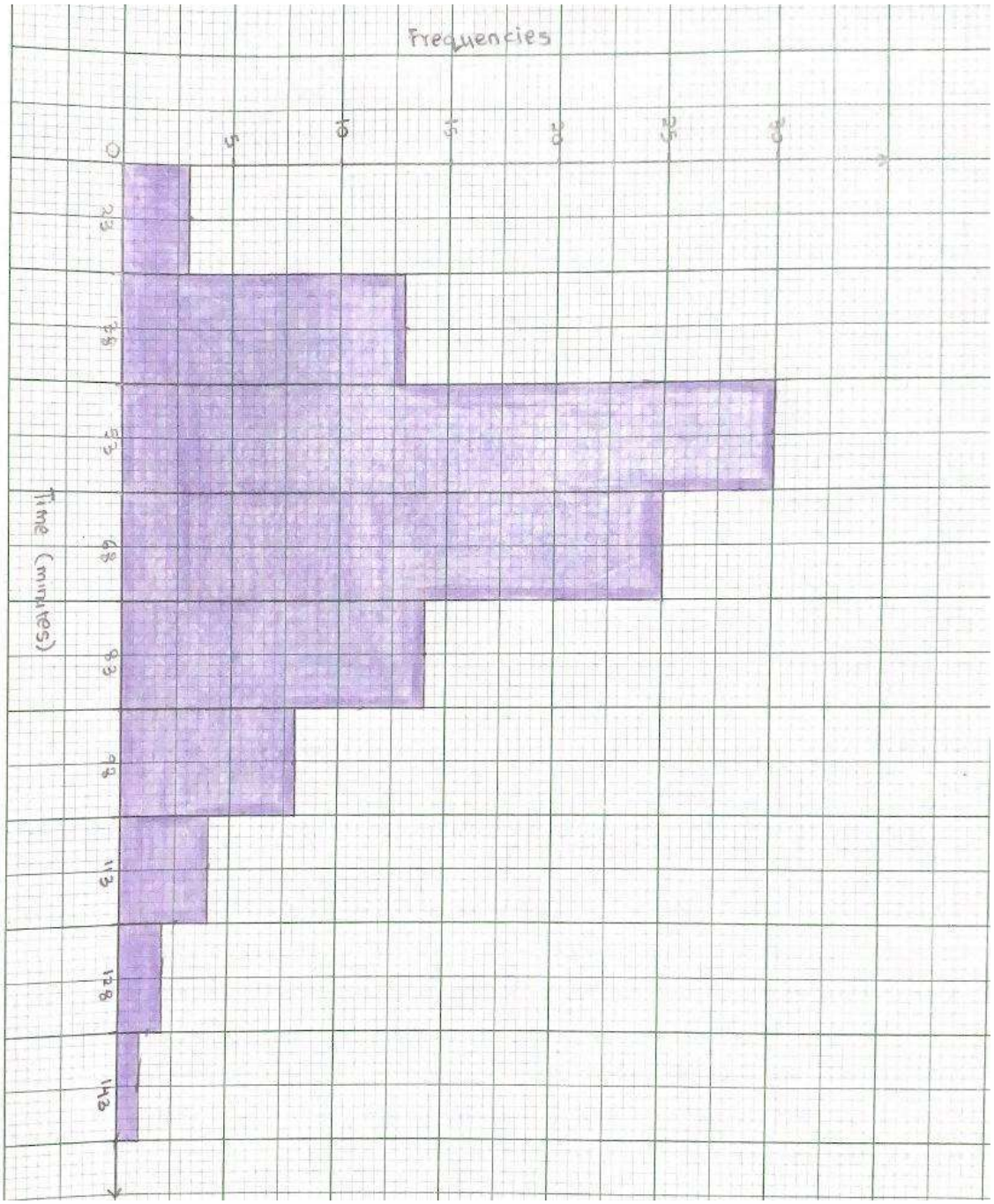
7

b. The graph is right skewed and leptokurtic

1

c. The average time for the flight delay is 66.5 minutes, but most of the flight delay is shorter in time than the average time.

2


Frequencies
0
5
10
15
20
25
30
Time (minutes)
23
38
53
68
83
98
113
128
143