



SECI2143-01 KEBARANGKALIAN STATISTIK DAN ANALISIS DATA

(PROBABILITY & STATISTICAL DATA
ANALYSIS)

ASSIGNMENT 1

DR SHARIN

DONE BY:

PHANG CHENG YI (A20EC0131)

SHAHRLIL BIN SAIFUL BAHRI (A20EC0144)

DATE OF SUBMISSION: 11/4/2021

PSDA Assignment 1

Question 1

1. (a) The population : Football clubs in Malaysia
(b) The sample : Malaysia Super League matches
(c) A discrete variable : The total number of goals scored
(d) A continuous variable : The mean number of goals per game
(e) Primary data :
- Records for each match the total number of goals scored
- Records for each match the amount of time played before a goal is scored.
(f) Secondary data : Old newspapers articles showed that in the previous season, the mean number of goals per game was 3.08.

2. (a) The population : Second year pupils
(b) The sample : The second year pupils weighed during the lunchtime
(c) A discrete variable : The number of pupils weighed
(d) A continuous variable : Height
(e) Primary data : The weight of second year pupils
(f) Secondary data : The data on first year pupils in medical examination

Relative Frequency	Cumulative Frequency	Relative Frequency	Cumulative Frequency	Relative Frequency	Cumulative Frequency
$\frac{1}{20} = 0.05$	1	$\frac{1}{20} = 0.05$	1	1	1
$\frac{11}{20} = 0.55$	12	$\frac{11}{20} = 0.55$	12	11	12
$\frac{7}{20} = 0.35$	19	$\frac{7}{20} = 0.35$	19	7	19
$\frac{1}{20} = 0.05$	20	$\frac{1}{20} = 0.05$	20	1	20
1.00		1.00		20	Total

Question 2

1.

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.

Scale	Frequency	Relative Frequency	Cumulative Frequency	Cumulative Relative Frequency
1	14	$\frac{14}{20} = 0.70$	14	0.70
2	6	$\frac{6}{20} = 0.30$	20	1.00
3	0	0.00	20	1.00
4	0	0.00	20	1.00
Total	20	1.00		

Price

Scale	Frequency	Relative Frequency	Cumulative Frequency	Cumulative Relative Frequency
1	3	$\frac{3}{20} = 0.15$	3	0.15
2	6	$\frac{6}{20} = 0.30$	9	0.45
3	11	$\frac{11}{20} = 0.55$	20	1.00
4	0	0.00	20	1.00
Total	20	1.00		

Condition of the Car

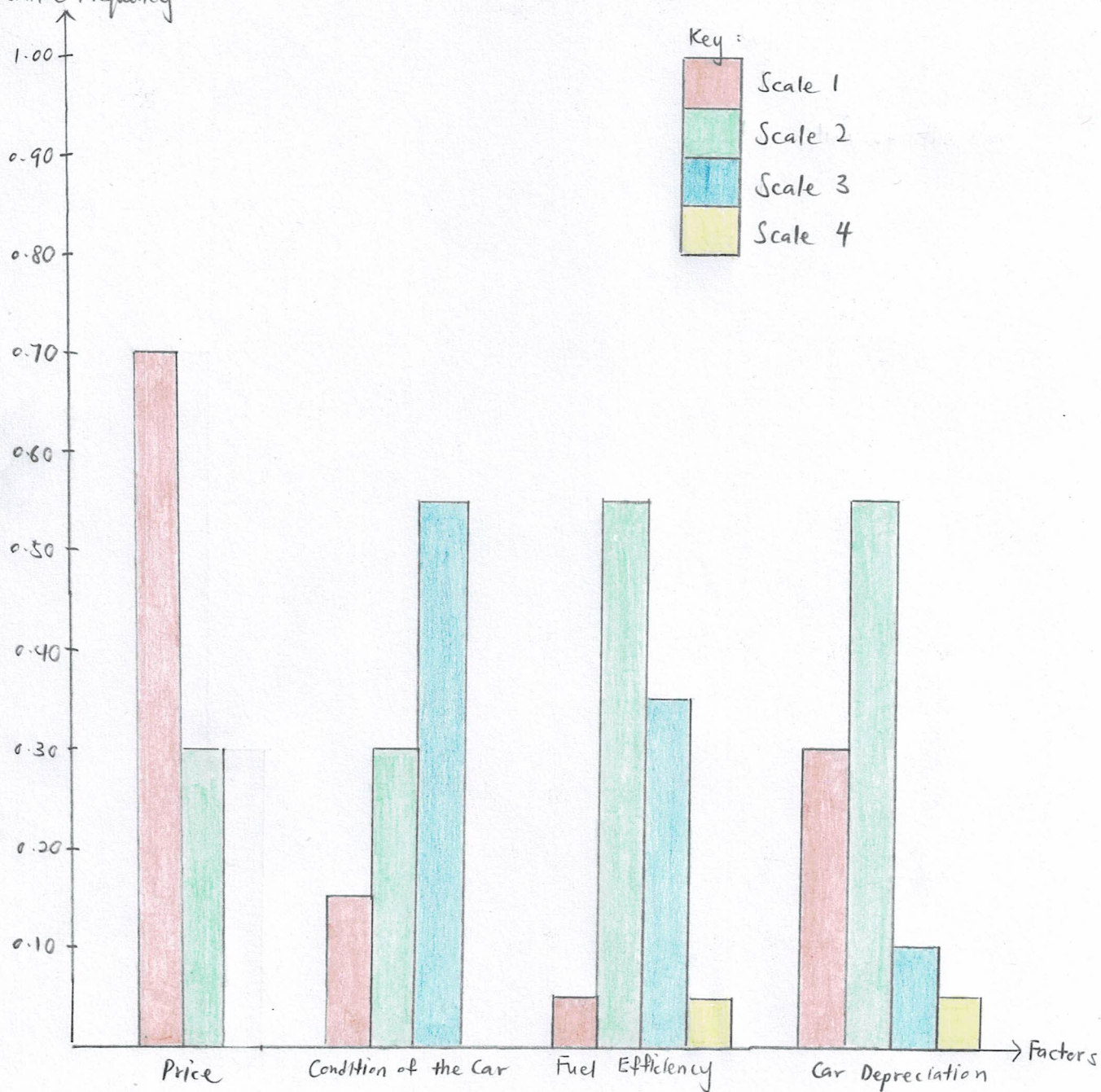
Scale	Frequency	Relative Frequency	Cumulative Frequency	Cumulative Relative Frequency
1	1	$\frac{1}{20} = 0.05$	1	0.05
2	11	$\frac{11}{20} = 0.55$	12	0.60
3	7	$\frac{7}{20} = 0.35$	19	0.95
4	1	$\frac{1}{20} = 0.05$	20	1.00
Total	20	1.00		

Fuel Efficiency

Scale	Frequency	Relative Frequency	Cumulative Frequency	Cumulative Relative Frequency
1	6	$\frac{6}{20} = 0.30$	6	0.30
2	11	$\frac{11}{20} = 0.55$	17	0.85
3	2	$\frac{2}{20} = 0.10$	19	0.95
4	1	$\frac{1}{20} = 0.05$	20	1.00
Total	20	1.00		

Car Depreciation

3. Relative Frequency



Question 3

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 : 2/4 means 24 days

2.(a) Mean, $\bar{x} = \frac{4 + 8 + 9 + 16 + 18 + 21 + 23 + 24 + 30 + 32 + 33 + 38 + 42 + 43 + 44 + 55 + 65 + 81}{18} = 32.56$

(b) no mode

(c) $i = \frac{18}{2} = 9$, $k = 9$

Median = $\frac{Y[9] + Y[10]}{2} = \frac{30 + 32}{2} = 31$

3.(a) $i = \frac{1}{4}(18) = 4.5$, $k = 5$

$Y[5] = 18$

$\therefore Q_1 = 18$

$i = \frac{3}{4}(18) = 13.5$, $k = 14$

$Y[14] = 43$

$\therefore Q_3 = 43$

$i = \frac{2}{4}(18) = 9$, $k = 9$

$Q_2 = \text{Median} = \frac{Y[9] + Y[10]}{2}$

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

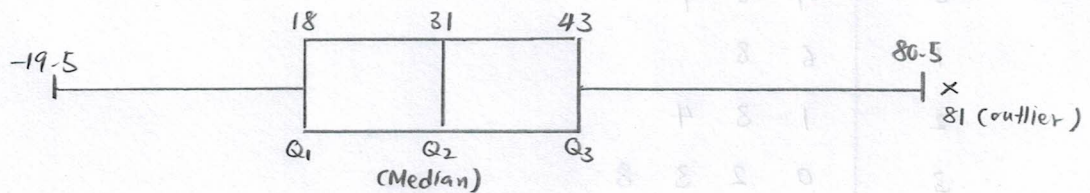
$\therefore Q_2 = 31$

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

(c) Lower Limit = $18 - 1.5(25) = -19.5$

Upper Limit = $43 + 1.5(25) = 80.5$

Outlier = 81



Question 4

$$1. \text{ Mean} = \frac{\sum fx}{n} = \frac{1(175000) + 5(250000) + 4(500000) + 1(700000)}{11} = \frac{4125000}{11} = \text{RM } 375000$$

2. $i = \frac{11}{2} = 5.5$, $k = 6$
 $Y[6] = 250000$ $\therefore \text{Median} = \text{RM } 250000$

Value per House	Number of Houses	Cumulative Frequency
RM 175 000	1	1
RM 250 000	5	6
RM 500 000	4	10
RM 700 000	1	11

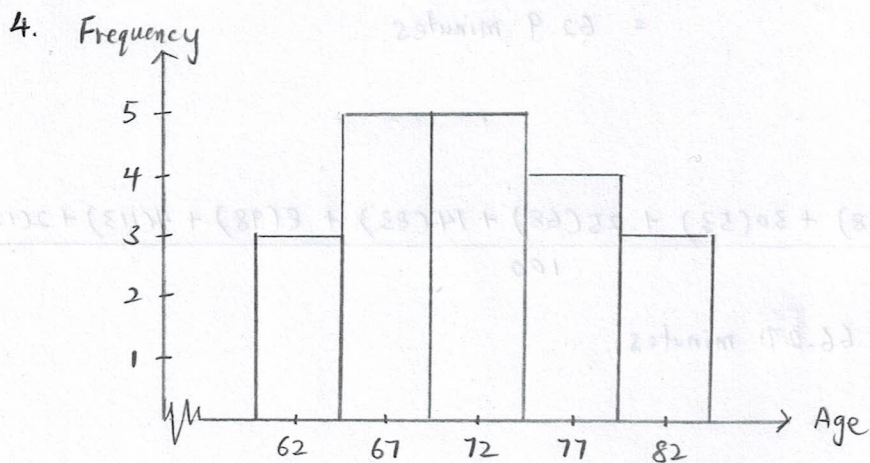
3. The mean is "best" represents the values of the 11 houses because it show the average value of houses in the new development area in Secudar.

Question 5

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

- | 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 |

- | 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 | | | 20 | |



Question 6

Time (minutes)	Frequencies	Midpoint	Cumulative Frequency
16-30	3	23	3
31-45	13	38	16
46-60	30	53	46
61-75	25	68	71
76-90	14	83	85
91-105	8	98	93
106-120	4	113	97
121-135	2	128	99
136-150	1	143	100
Total	100		

$$1. \quad \bar{i} = \frac{100}{2} = 50$$

Median class = 61-75 minutes

$$L = 60.5$$

$$N = 100$$

$$cf_p = 46$$

$$f_{med} = 25$$

$$W = 15$$

$$\text{Median} = L + \left(\frac{\frac{N}{2} - cf_p}{f_{med}} \right) W$$

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

$$= 60.5 + \left(\frac{4}{25} \right) 15$$

$$= 62.9 \text{ minutes}$$

$$\text{Mean} = \frac{\sum fx}{n}$$

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

$$= \frac{6650}{100} = 66.50 \text{ minutes}$$

Mode class = 46-60 minutes

$$l = 45.5$$

$$h = 15$$

$$f_1 = 30$$

$$f_0 = 13$$

$$f_2 = 25$$

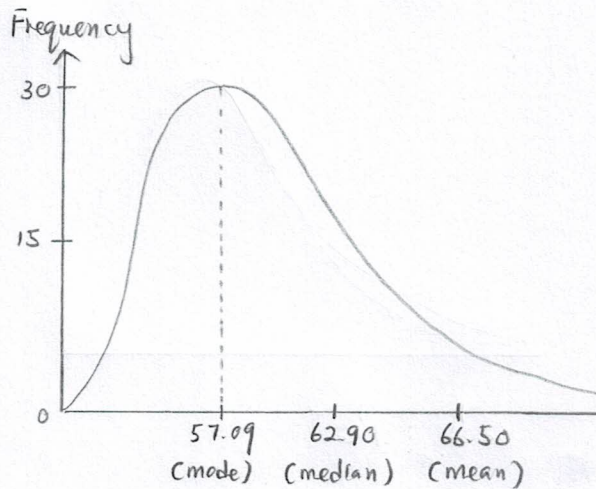
$$\text{Mode} = l + \left[\frac{(f_1 - f_0)}{(2f_1 - f_0 - f_2)} \right] h$$

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

$$= 45.5 + \left(\frac{17}{22} \right) 15$$

$$= 57.09 \text{ minutes}$$

2. (a) Mode < Median < Mean



(b) The distribution graph is skewed to the right (positively skewed).

(c) The delay time for most of the flight are less than the average time which is 66.50 minutes. Most of the flight will delay around 60 minutes to take off or land.