Assignment 1

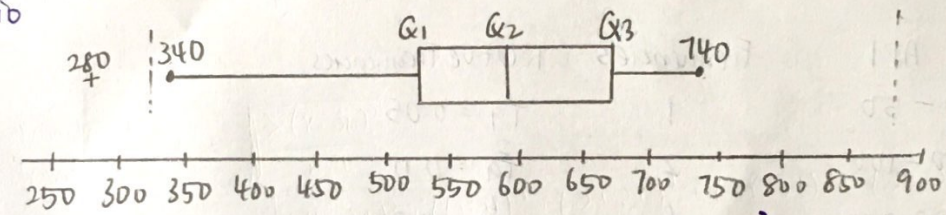
Ham Jing Yi A19EC0048
Khor Yong Xin A19EC0061

(1) a). ratio
b). ordinal
c). interval
d). nominal
e) ordinal
f). ratio
g). ratio
h). interval
i) ratio
j) ratio

(2) ai) population : the freshmen class at Lincoln High School
ii) sample : 100 students randomly selected.

bi) nominal
ii) ratio
iii) ratio

ci)



ii) $Q_1 = \frac{1}{4}(20)$ $Q_2 = \frac{1}{2}(20)$ $Q_3 = \frac{3}{4}(20)$

$= 5$ $= 10$ $= 15$

$Q_1 = \frac{Y[5] + Y[6]}{2}$ $Q_2 = \frac{Y[10] + Y[11]}{2}$ $Q_3 = \frac{Y[15] + Y[16]}{2}$

$= \frac{520 + 540}{2}$ $= \frac{580 + 600}{2}$ $= \frac{660 + 680}{2}$

$= 530$ $= 590$ $= 670$

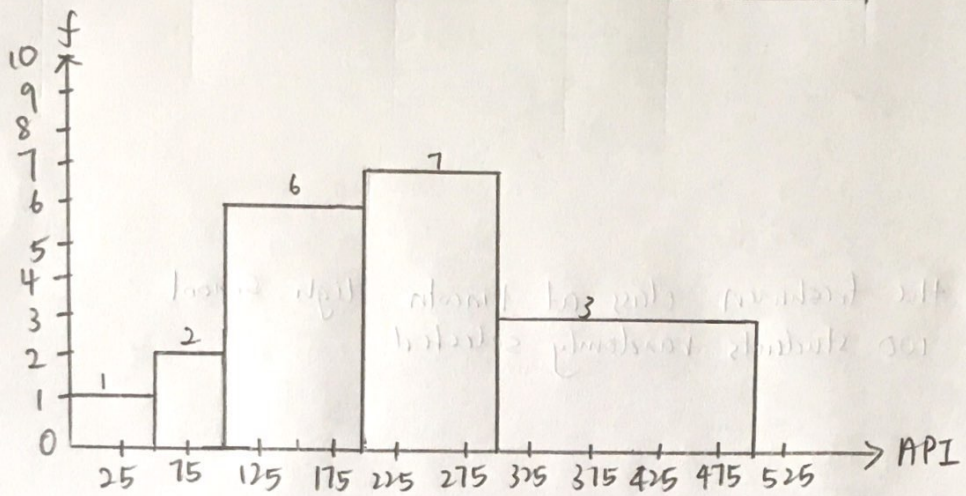
Interquartile range (IQR) $= Q_3 - Q_1$
$= 670 - 530$
$= 140$

Min: $Q_1 - 1.5(IQR)$
$= 530 - 1.5(140)$
$= 320$

Max: $Q_3 + 1.5(IQR)$
$= 670 + 1.5(140)$
$= 880$

3a).

API	lower boundaries	upper boundaies	f	Midpoint
0-50	0	50	1	25
50-100	50	100	2	75
100-200	100	200	6	150
200-300	200	300	7	250
300-500	300	500	3	400



b).

API	Frequencies	relative Frequencies
0 - 50	1	$\frac{1}{19} = 0.05$
50 - 100	2	$\frac{2}{19} = 0.11$
100 - 200	6	$\frac{6}{19} = 0.32$
200 - 300	7	$\frac{7}{19} = 0.37$
300 - 500	3	$\frac{3}{19} = 0.16$
Total	19	1.00

Percentage with API ≥ 100

$= (0.32 + 0.37 + 0.16) \times 100\%$

$= 85\%$

c). Yes. Because mean can be affected by outliers.
Besides, all the values are taken into account.

(4)(a) $\text{mean} = \dfrac{92+98+112+88+134+36+35+92+215+336+237}{11}$

$= \dfrac{1475}{11}$

$= 134.09$

mode = 92

35, 36, 88, 92, 92, 98, 112, 134, 215, 237, 336

$\frac{11}{2} = 5.5$

median = X[6]

= 98

standard deviation, $s = \sqrt{\dfrac{\Sigma(x-\bar{x})^2}{n-1}}$

$= \sqrt{\dfrac{(92-134.09)^2+(98-134.09)^2+(112-134.09)^2+(88-134.09)^2+(134-134.09)^2+(36-134.09)^2+(35-134.09)^2+(92-134.09)^2+(215-134.09)^2+(336-134.09)^2+(237-134.09)^2}{11-1}}$

$= \sqrt{\dfrac{84802.9091}{10}}$

$= 92.09$

(b) Percentile of the deaths 215 $= \frac{8}{11} \times 100\%$

$= 72.73\%$

(c) Mean. Because it can be affected by outliers and all the values are taken to be account

(5)(a) $\text{mean} = \dfrac{14+14+10+12+11+13+11+11+14+10+13+8}{12}$

$= \dfrac{141}{12}$

$= 11.75$

mode = 11 and 14

8, 10, 10, 11, 11, 11, 12, 13, 13, 14, 14, 14

$\frac{12}{2} = 6$

$\text{median} = \dfrac{X[6]+X[7]}{2}$

$= \dfrac{11+12}{2}$

$= 11.5$

range = 14 - 8
= 6

(b) median < mean
∴ It is positively skew (positive distribution)

(c) standard deviation, $s = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}}$

$$= \sqrt{\frac{(14-11.75)^2+(14-11.75)^2+(10-11.75)^2+(13-11.75)^2+(11-11.75)^2+(13-11.75)^2+(11-11.75)^2+(14-11.75)^2+(10-11.75)^2+(13-11.75)^2+(8-11.75)^2}{12-1}}$$

= 1.9129

(d) $12 \times \frac{1}{4} = 3$

$Q_1 = \frac{X[3]+X[4]}{2}$

$= \frac{10+11}{2}$

$= 10.5$

$12 \times \frac{3}{4} = 9$

$Q_3 = \frac{X[9]+X[10]}{2}$

$= \frac{13+14}{2}$

$= 13.5$

Lower limit $= Q_1 - 1.5IQR$

$= 10.5 - 1.5(13.5 - 10.5)$

$= 6 < 7$

∴ The data is not considered an outlier

(6) (a) mean $= \frac{14+12+21+28+30}{5}$

$= 21$

standard deviation, $s = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}}$

$$= \sqrt{\frac{(14-21)^2+(12-21)^2+(21-21)^2+(28-21)^2+(30-21)^2}{5-1}}$$

= 8.0623

skewness $= \frac{\Sigma(x-\bar{x})^3}{(n-1)s^3}$

$$= \frac{(14-21)^3+(12-21)^3+(21-21)^3+(28-21)^3+(30-21)^3}{(5-1)(8.0623)^3}$$

= 0

(b) It is neither positively skewed nor negatively skewed. The skewness value is zero. Hence it is symmetrical distribution