

NAME: JIYAD HUSSAIN

ROLL NO: 19I-1771

SECTION: BS - DS(N)

SUBJECT: PROBABILITY & STATISTICS

DOUE DATE: 9<sup>th</sup> APRIL, 2021

## ASSIGNMENT 2

Q1: SOURCE A:

LIFE [HOURS]	No of Component [F]	CUMULATIVE FREQUENCY	MIDPOINT X	X.F	$(x-\bar{M})^2$	$(x-\bar{M})^2 \cdot F$
1000 - 1020	40	40	1010	40400	2490.01	99600.40
1020 - 1040	96	136	1030	93880	894.01	85824.96
1040 - 1060	364	500	1050	36220	99.01	35675.64
1060 - 1080	372	872	1070	398040	102.01	37947.72
1080 - 1100	85	957	1090	92650	906.01	77010.85
1100 - 1120	43	1000	1110	47730	2610.01	107930.43
$\sum f = 1000$					$\sum xf = 1051160$	$\sum (x-\bar{M})^2 \cdot F = 443990$

SOURCE B:

LIFE [Hours]	No of Component [F]	CUMULATIVE FREQUENCY	MIDPOINT X	X.F	$(x-\bar{M})^2$	$(x-\bar{M})^2 \cdot F$
1030 - 1040	339	339	1035	350965	633.0256	214595.6
1040 - 1050	136	475	1045	142220	229.8256	31266.2
1050 - 1060	25	500	1055	26375	26.6256	665.64
1060 - 1070	20	520	1065	21300	23.4256	468.51
1070 - 1080	130	650	1075	139750	220.2256	28629.3
1080 - 1090	350	1000	1085	379750	617.0256	225958
$\sum f = 1000$					$\sum xf = 1060160$	$\sum (x-\bar{M})^2 \cdot F = 4926$

1. M<sub>EDIAN</sub>:

SOURCE A:

$$M_{EDIAN} \text{ CLASS } FAILS \left[ \frac{N}{2} = \frac{1000}{2} = 500 \right]$$

CUMMULATIVE FREQUENCY = 500

M<sub>EDIAN</sub> CLASS IS 1060 - 1080

$$M_{EDIAN} = L + \left[ \frac{\frac{N}{2} - C}{f} \right] \times h$$

$$L = 1060, \frac{N}{2} = 500, C = 500, f = 372, h = 20$$

$$M_{EDIAN} = 1060 + \left[ \frac{500 - 500}{372} \right] \times 20$$

$$M_{EDIAN} = 1060$$

$$\text{FIRST QUARTILE } [Q_1] = L + \left[ \frac{\frac{1}{4}N - C}{f} \right] \times h$$

$$\frac{N}{4} = \frac{1000}{4} = 250, Q_1 \text{ class is } 1040 - 1060$$

$$L = 1040, \frac{N}{4} = 250, C = 136, f = 136, f = 364, h = 20$$

$$Q_1 = 1040 + \left[ \frac{250 - 136}{364} \right] \times 20 = 1040 + 6.26 = 1046.26 \text{ HR}$$

$$Q_1 = 1046.26$$

$$\text{THIRD QUARTILE } [Q_3] = L + \left[ \frac{\frac{3}{4}N - C}{f} \right] \times h$$

$$\frac{3}{4}N = \frac{3}{4} \times 1000 = 750 ; Q_3 \text{ class is } 1060 - 1080$$

$$L = 1060, C = 500, f = 372, h = 20$$

$$Q_3 = 1060 + \frac{750 - 500}{372} \times 20 = 1060 + 13.51$$

$$Q_3 = 1073.51$$

SOURCE B:

$$M_{\text{MEDIAN}} \text{ CLASS } F_{\text{A.I.L}} \left[ \frac{N}{2} = \frac{1000}{2} = 500 \right]$$

CUMULATIVE FREQUENCY = 500

$M_{\text{MEDIAN}}$  CLASS IS 1060 - 1080

$$M_{\text{MEDIAN}} = L + \left[ \frac{\frac{N}{2} - C}{f} \right] \times h$$

$$L = 1060, \frac{N}{2} = 500, C = 500, f = 20, h = 10$$

$$M_{\text{MEDIAN}} = 1060 + \left[ \frac{500 - 500}{20} \right] \times 10$$

$$M_{\text{MEDIAN}} = 1060$$

$$\text{FIRST QUARTILE } [Q_1] = L + \left[ \frac{\frac{N}{4} - C}{f} \right] \times h$$

$$\frac{N}{4} = 250, Q_1 \text{ CLASS IS } 1030 - 1040$$

$$L = 1030, \frac{N}{4} = 250, c = 0, h = 10, f = 339$$

$$Q_1 = 1030 + \left[ \frac{250 - 0}{339} \right] \times 10 = 1030 + 7.37$$

$$Q_1 = 1037.37 \text{ Hrs}$$

$$\text{THIRD QUARTILE } [Q_3] = L + \left[ \frac{\frac{3}{4}N - c}{f} \right] \times h$$

$$\frac{3}{4}N = \frac{3}{4} \times 1000 = 750, Q_3 \text{ CLASS IS } 1080 - 1090$$

$$L = 1080, \frac{3}{4}N = 750, c = 650, f = 350, h = 10$$

$$Q_3 = L + \left[ \frac{750 - 650}{350} \right] \times 10$$

$$= 1080 + \frac{1000}{350} = 1082.86$$

## 2. SOURCE A:

SAMPLE MEAN:

$$\bar{M} = \frac{\sum xf}{\sum f} = \frac{1059900}{1000} = 1059.90$$

SAMPLE STANDARD DEVIATION:

$$M = \sqrt{\frac{\sum (x - \bar{M})^2 f}{\sum f - 1}} = \sqrt{\frac{443990}{1000 - 1}} = 21.08$$

SOURCE B:

SAMPLE MEAN:

$$M = \frac{\sum xf}{\sum f} = \frac{1060160}{1000} = 1060.16$$

SAMPLE STANDARD DEVIATION:

$$S = \sqrt{\frac{\sum (x - M)^2 f}{\sum f - 1}} = \sqrt{\frac{491574.40}{100 - 1}} = 22.183$$

Both the Mean and Standard Deviation of the component from Source B are greater

3. SOURCE B is more better than SOURCE A because

$$\text{MEAN OF SOURCE B} > \text{MEAN OF SOURCE A}$$

&

$$\text{STANDARD DEVIATION OF SOURCE B} > \text{STANDARD DEVIATION OF SOURCE A}$$

4. SOURCE A:

QUARTILE DERIVATION IN TERM OF ABSOLUTE MEASURE OF DISPERSION:

$$Q.D = \frac{Q_3 - Q_1}{2} = \frac{1073.51 - 1046.26}{2} = 13.625$$

QUARTILE DERIVATION IN TERM OF RELATIVE MEASURE OF DISPERSION:

$$Q.D = \frac{Q_3 - Q_1}{Q_3 + Q_1} = \frac{1073.51 - 1046.26}{1073.51 + 1046.26} = 0.1285516825$$

SOURCE B:

QUARTILE DERIVATION IN TERM OF ABSOLUTE MEASURE OF DISPERSION:

$$Q.D = \frac{Q_3 - Q_1}{2} = \frac{1082.86 - 1037.37}{2} = 22.745$$

QUARTILE DERIVATION IN TERM OF RELATIVE MEASURE OF DISPERSION:

$$Q.D = \frac{Q_3 - Q_1}{Q_3 + Q_1} = \frac{1082.86 - 1037.37}{1082.86 + 1037.37} = 0.02145521148$$

Q2:

### NEW ENGLAND STATES:

M<sub>MEAN</sub>:

$$\bar{x} = \frac{6768 + 3196 + 1112 + 819 + 1019 + 1795}{6}$$

$$\bar{x} = 2451.5$$

M<sub>MEDIAN</sub>:

ARRANGE THE DATA IN AN INCREASING ORDER)

819, 1019, 1112, 1795, 3196, 6768

$$MD = \frac{1112 + 1795}{2}$$
$$= 1453.5$$

M<sub>MODE</sub>:

Mode = Does Not Exist, because no value is repeated more than once

### NORTH WEST STATES:

M<sub>MEAN</sub>:

$$\bar{x} = \frac{1870 + 622 + 620 + 23 + 172 + 112}{6}$$

$$\bar{x} = 569.8333$$

M<sub>MEDIAN</sub>:

ARRANGE THE DATA IN AN INCREASING ORDER)

23, 112, 172, 620, 622, 1870

$$MD = \frac{172 + 620}{2} = 396$$

MODE:

Mode = Does Not Exist, because no value is repeated more than once

COMPARISON:

All measure of centre for New England States are greater than

All measure of centre for North West States.

Q3.

a. Mean =  $\frac{\sum x_i f_i}{\sum f_i}$

GIVEN Mean = 5 and  $\sum f_i = 20$

$$\text{Mean} = \frac{1 \times 2 + 2 \times 5 + 4 \times 7 + 6 \times 6}{20}$$

$$5 = \frac{2 + 10 + 28 + 6x}{20}$$

$$100 = 40 + 6x$$

$$6x = 100 - 40$$

$$x = \frac{60}{6}$$

$$x = 10$$

b. Upper Quartile = 110

$$\text{Median} = 61$$

$S_{KB} = 0$ ,  $\therefore$  the distribution is symmetric

$$S_{KB} = \frac{Q_3 + Q_1 - 2Q_2}{Q_3 - Q_1}$$

$$Q = Q_3 + Q_1 - 2Q_2$$

$$\begin{aligned}Q_1 &= 2Q_2 - Q_3 \\&= 2(61) - 110 \\&= 122 - 110 \\&= 12\end{aligned}$$

Lower Quartile  $Q_1 = 12$

Q4

a. Given:

Average number of days Construction Workers miss per year = 11

Standard deviation = 2.3

Average number of days Factory Workers miss per year = 8

Standard deviation = 1.8

Here from the given information:

Standard deviation of the construction workers is 2.3 which is larger than the standard deviation of factory workers which is 1.8. Hence constructions are more variable in terms of day missed. Because standard deviation is a measure variability. The larger the standard deviation value, the greater the variability.

b. GIVEN:

	MEAN	SD
ENGLISH	$\bar{x} = 85$	$s = 5$
HISTORY	$\bar{x} = 110$	$s = 8$

Coefficient of variation for English Score:

$$C\text{Var} = \frac{s}{\bar{x}} \times 100\% = \frac{5}{85} \times 100\% = 5.88\%$$

Hence, the coefficient of variation for English Class is 5.88%.

Coefficient of variation for History Class:

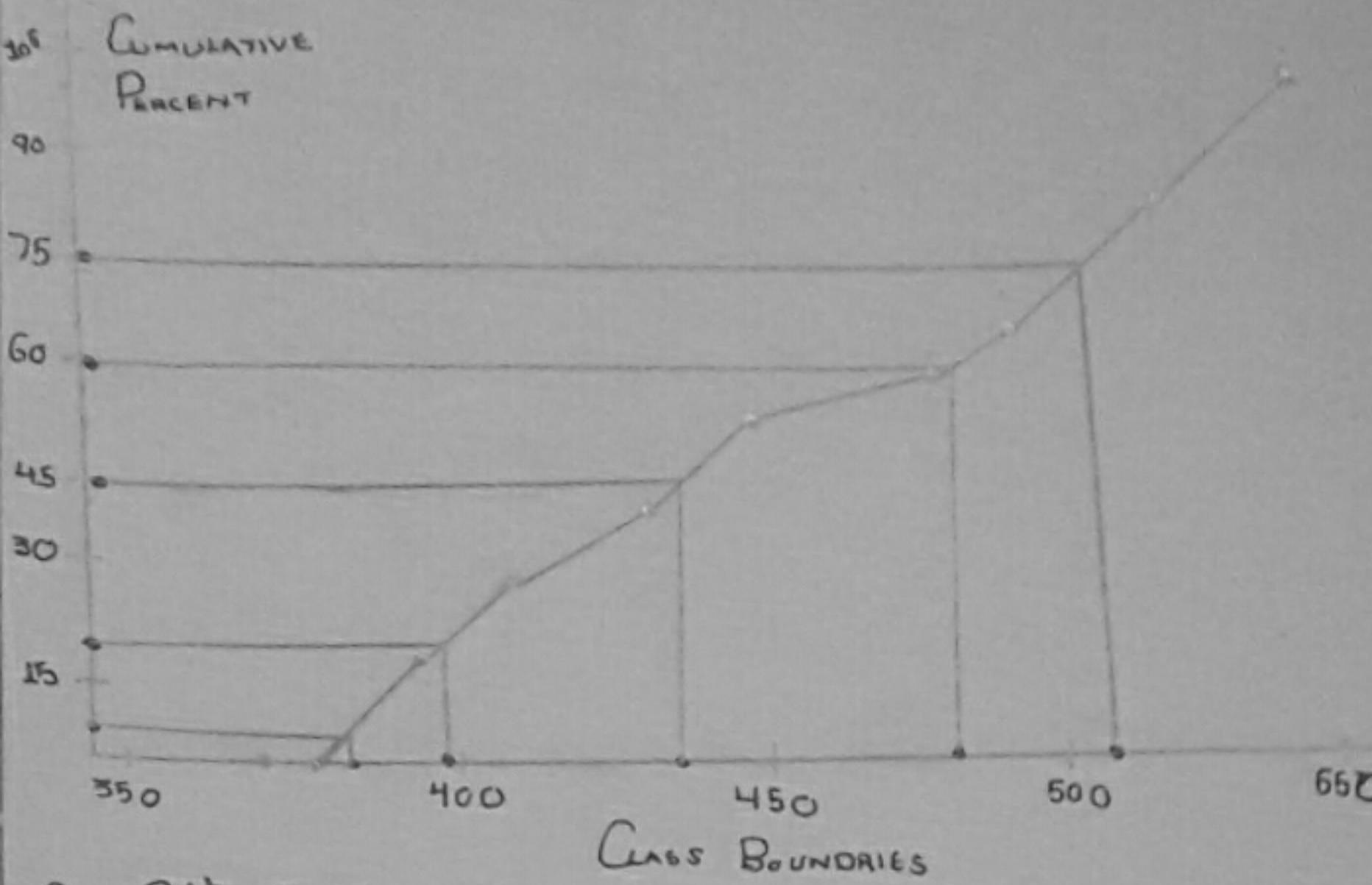
$$C\text{Var} = \frac{s}{\bar{x}} \times 100\% = \frac{8}{110} \times 100\% = 7.27\%$$

Hence, the coefficient of variation for History class is 7.27%.

It is observed from the above calculations, the History score is more variable than the score of English score

Q5.

CLASS	FREQUENCY	CUMULATIVE FREQUENCY	CUMULATIVE PERCENT
366 - 386	4	4	19.04
387 - 407	2	6	28.57
408 - 428	3	9	42.86
429 - 449	2	11	52.38
450 - 470	1	12	57.14
471 - 491	2	14	66.67
492 - 512	3	17	80.95
513 - 533	4	21	100



### a. 9<sup>th</sup> PERCENTILE:

From the above graph, it can be observed that the 9<sup>th</sup> percentile is start on Y-axis at 9 and the value corresponding to x-axis is 376. Therefore, the values at 9<sup>th</sup> percentile is 376

### b. 20<sup>th</sup> PERCENTILE:

From the above graph, it can be observed that the 20<sup>th</sup> Percentile is start on Y-axis at 20 and the value corresponding to x-axis is 389. Therefore, the values at 20<sup>th</sup> percentile is 389

### c. 45<sup>th</sup> PERCENTILE:

From the above graph, it can be observed that the 45<sup>th</sup> Percentile is start on Y-axis at 45 and the value corresponding to x-axis is 433. Therefore, the values at 45<sup>th</sup> is 433

### d. 60<sup>th</sup> PERCENTILE:

From the above graph, it can be observed that the 60<sup>th</sup> Percentile is start on Y-axis at 60 and the value corresponding to x-axis is 477. Therefore, the values at 60<sup>th</sup> is 477