1. A study of yellow fever disease (YFD) was conducted. The study was restricted to patients under the age of 40. The purpose was to compare the distribution of cases by age in group A to that of group B. The group data are given below:

Cl	Group A	Group B
1-5	0	2
6-10	7	4
11-15	10	7
16-20	2	20
21-25	1	16
26-30	5	10
31-35	4	4

Calculate (i) The mean and standard deviation for each group (ii) The coefficient of variation of each group (iii) which group has less variable distribution

Solution

(i) Group A

Cl	(f) Group A	x	fx	$\lvert x-\bar{x}\rvert$	$(x-\bar{x})^2$	$(\bar{x}-x)^2 f$	$f\lvert x-x_m\rvert$
1-5	0	3	0	14.8	219.04	0	0
6-10	7	8	56	9.8	96.04	693.28	68.6
11-15	10	13	130	4.8	23.04	230.4	48
16-20	2	18	36	0.2	0.04	0.08	0.4
21-25	1	23	23	5.2	27.04	27.04	5.2
26-30	5	28	140	10.2	104.04	520.2	51
31-35	4	33	132	15.2	231.04	924.16	60.8
	29		517		2574.16	2354	

mean $\bar{x}_A = \frac{\Sigma fx}{\Sigma f} = \frac{512}{29} = 17.8$

The $MD_A = \frac{\Sigma f|x - \bar{x}_A|}{\Sigma f} = \frac{234}{29} = 8.07$

The $SD_A = \sqrt{\frac{\Sigma(x-\bar{x})^2 f}{\Sigma f - 1}} = \sqrt{\frac{2374.16}{29-1}} = 9.2$ //

Group B.

CI	cf Group B	midpoint x	f(x)	$\lvert x-\bar{x}_B\rvert$	$f\lvert x-\bar{x}_B\rvert$	$(x-\bar{x}_B)^2$	$f(x-\bar{x}_B)^2$
1-5	2	3	6	17.1	34.2	292.41	584.82
6-10	4	8	32	12.1	48.4	146.41	585.64
11-15	7	13	91	7.1	49.7	50.41	352.87
16-20	26	18	360	2.1	42	4.41	88.2
21-25	16	23	368	2.9	46.4	8.41	134.56
26-30	10	28	280	7.9	79	62.41	624.1
31-35	4	33	132	12.9	51.6	166.41	665.64
	63		1269		351.3		3035.83

mean $\bar{x}_B = \frac{\Sigma fx}{\Sigma f} = \frac{1269}{63} = 20.1$

The $MD_B = \frac{\Sigma f|x-\bar{x}_B|}{\Sigma f} = \frac{351.3}{63} = 5.55$

$S.D_B = \sqrt{\frac{\Sigma f(x-\bar{x}_B)^2}{\Sigma f}} = \sqrt{\frac{3035.83}{63-1}} = 6.993 = 7$ //

(ii) The Coefficient Variation (C.V$_A$) for Group A

$= \frac{SD_A}{mean_A} \times 100$

where S.D = 9.2, mean = 17.8

$C.V_A = \frac{9.2}{17.8} \times 100 = 51.7$

for Group B

Coefficient of Variation (C.V$_B$) $= \frac{S.D_B}{mean_B} \times 100$

where S.D = 7 and mean = 20.1

$$CV_B = \frac{1}{20.1} \times 100$$

$$= \frac{100}{20.1}$$

$$= 34.1$$

(iii) The one with least coefficient variation is less variable
Therefore Group B has a better variable distribution