Mode is the value of the variable which has the largest frequency.

Mean is the measure of the position commonly known as the average. The mean π of the n numbers x… is given by $\pi = x+x/n.\pi=\sum x/Nf$

Find the mean of the data shown in table f

Table F

x	f
0	2
1	3
2	5
3	8
4	7
5	3

Note that $\sum f=28$,so there are 28 values of the variable each value of x occurs with a certain frequency f. To find the sum of all these values, we must find the sum of the products f*x and divide by $\sum f$. Hence, for a frequency distribution $\pi=\sum fx/\sum f$. Write the values of x and f in vertical columns and add a third a third column for values of fx. Complete this column(Table F).

Table F

x	f	fx
0	2	0
1	3	3
2	5	10
3	8	24
4	7	28
5	3	15
	28	80

Then $\bar{x} = \sum fx/\sum f = 80/28 \approx 2.9$. The mean of the data, that is their average value is 2.9.it is usual to give the mean to one more decimal place than the data.

Mean of a grouped frequency distribution

Find the mean mark for the distribution shown in Table G.

Table G

Marks	f
0-9	2
10-19	8
20-29	14

30-39	24
40-49	48
50-59	52
60-69	32
70-79	12
80-89	6
90-99	2

Note: We do not know the original marks as they have been grouped into the classes shown. Hence, we take the class center as the representative value of x in each class. This assumes that the values of x are generally evenly spread over each class any difference will balance out overall.it can be proved that in general, the error in the value of the mean made by grouping will be small for a reasonably large total frequency. Copy and work through the following steps referring to Table G where all values have been completed. STEP 1 Add a column for the class centers(c) ,4.5,14.5, etc., We should now find the values of fc, that is frequency (f) x class center. STEP 2 To avoid long calculations, we use an assumed mean, say 44.5.it does not matter which value we choose, but obviously it is best to use a value of the class centers near the middle of the distribution. Add a further column for values of (c-44.5). STEP 3 We now find ∑f(c-44.5). the calculations are now simple as all the values of (c-44.5) are multiples of 10.Take care of the signs. STEP 4 ∑f(c-44.5) =1020 and ∑f=200

Hence $\bar{x}$=∑f(c-44.5)/∑f+44.5=1020/200+44.5=5.1+44.5=49.5

Table G

Class Limits x	f	STEP 1 Class centers c	STEP 2 (Assumed Mean 44.5) c-44.5	STEP 3 F(c-44.5)
0-9	2	4.5	-40	-80
10-19	8	14.5	-30	-240
20-29	14	24.5	-20	-280
30-39	24	34.5	-10	-240
40-49	48	44.5	0	0
50-59	52	54.5	10	520
60-69	32	64.5	20	640
70-79	12	74.5	30	360
80=89	6	84.5	40	240
90-99	2	94.5	50	100
	200			1020

The histogram is shown in fig.? with the position of the mean marked (the model class is 50-59).

Summarizing the method of finding the mean of a grouped frequency distribution, prepare a table with columns as in fig.

Table J

Given Data

Class Limits x	f	Class centers c	Assumed mean = m) c-m	F(c-m)	
			{m is normally one of the middle values of c		
	∑f-				∑f(c-m)=

Then $\bar{x}$ = ∑f(c-m)/∑f + m.

Median

2 3 5 5 **6** 7 8 9 9. The median of the set of 9 numbers shown is the middle number of the set, when they are arranged in order. median=6.

Note for n numbers, the median is the n+1/2th in order. however, if there is an even number in the set, none of the can be the median number, in this case we take the mean of the two middle numbers as being the median.2 3 5 6 7 8 9. Median=5+6/2=5.5.

A sample of 200 spanners is taken from the daily production in a factory and the distance between the faces measured to the nearest 0.1 mm. The result are shown in Table H.(a) State the model and median classes.(b)Estimate by calculation the mean and median of this distribution.(c) Draw a histogram of the distribution and mark the results of (b) on it.

Table H

Distance (mm)	f
28.1-28.5	1
28.6-29.0	6
29.1-29.5	14
29.6-30.0	22
30.1-30.5	34
30.6-31.0	54
31.1-31.5	56
31.6-32.0	13

(a) The model class is 31.5 mm (highest frequency 56. ∑f=200 so, the median is the 100.5^{th} value. Counting the frequencies,1+6+14+22+34=77,so the median will fall into the 30.6-31.0mm class, which is the median class.

(b) Copy and complete the Table for the calculation of the mean.

Table

Class	f	Class center c	(Assumed mean 30.3)	F(c-30.3)

			C-30.3	
28.1-28.5	1	28.3	-2	-2
28.6-29.0	6	28.8	-1.5	-9
30.1-30.5	34	30.3	0	0
31.1-31.5	56	31.3	1	56
31.6-32.0	13	31.8	1.5	19.5

∑f = 200 ∑f(c-30.3) =

Mean = 30.63mm

Median = 30.55 + 23.5/54 * .5 = 30.77mm

(c) the histogram is shown in fig. with the positions of the mean and median marked

Cumulative Frequency

Percentile

Quartiles

Table T

Marks in classes	f	Cumulative frequencies	
1-10	8	8 i.e. 8	<10.5
11-20	17	25 .i.e. 8+17	,<20.5
21-30	39	64 .i.e. 25+39	<30.5
31-40	65	129 .i.e. 64+6540	<40.5
41-50	100	229 i.e. 129+100	< 50.5
51-60	122	351 i.e. 229+122	< 60.5
61-70	81	432 i.e. 351+81	<70.5
71-80	42	474 i.e. 432+42	<80.5
81-90	17	491 i.e. 474+17	<90.5
91-100	9	500 i.e. 491+9	<100.5

Candidate with less than 10.5 marks(<10.5) and so on.

∑f=500

Draw a cumulative frequency polygon, plotting cumulative frequencies against the upper class boundaries.

Plotting the point: note that the cumulative frequencies are plotted against the upper-class boundaries (10.5, etc.) not the upper-class limits. The horizontal axis here is marked 10s so it is easy to read. At the left-hand end we add the end point of the polygon, 0 frequency at 0.5 marks, as 0 candidates had <0.5 marks. Join up the point with straight lines to form a polygon. Show Ogive and Mark percentage frequencies on the vertical axis

From fig. find (a) the pass mark if 70% of the candidates are to pass, (b) the number of candidates who obtained 75 marks or more. (c) the number of candidates whose marks lie between 25 and 75 inclusive.

(a) If 70% of the candidates are to pass, then 30% are to fail. The pass mark will be therefore, at the 30^{th} percentile. From the graph this is seen to be 42.5. Hence, the pass mark will be 43. candidates obtaining 43 or more will pass. Note: if 70% pass, these are upper 70%, i.e. above the 30^{th} percentile.
(b) Read upwards from 75 on the mark scale. This gives 450 on the cumulative frequency scale. Then 50 candidates (500-450) are above this level and obtained 75 or more marks.
(c) Read upwards from 25 and 75 on the mark scale. From 25 we reach approximately 43 on the cumulative frequency scale and from 75 we reach 450.hence 450-43=407 candidates had marks

Given the frequency distribution shown in Table Q. draw a cumulative frequency curve. From the curve, find (a)the median and interquartile range (b) how many children had heights between 171 cm and 188 cm.

Note that the class limits are given in such a away the boundaries of the first class are 159.5 and 165 and so on.

The cumulative frequencies are as follows:

10<165 cm

35<170 cm

75<175 cm

131<180cm

175<185cm

195<190cm

200<195cm

Table Q

Heights of 200 children to the nearest cm

Heights h cm	f
160≤h<165	10
165≤h<170	25
170≤h<175	40
175≤h<180	56
180≤h<185	44
185≤h<190	20
190≤h<195	5

The given by these are plotted and a curve drawn throw them smoothly as possible

(a) The median is read from the 100th cumulative frequency value giving 177.5 cm.the lower quartile (50th frequency value) =172.2 cm. the upper quartile (150th frequency value) = 182.2.the interquartile range = 182.2-172.2=10 cm
(b) The frequency corresponding to 171 cm=40, and the frequency corresponding to 188 cm=188 hence 188-40=148 children had heights between 171 cm and 188 cm.

Standard Deviation (S.D)

Math's marks given in Table. Their mean $\bar{x}$ =58 marks. Find the deviation of each mark from the mean:

Marks x	Deviation d from the mean $d=x-\bar{x}$
35	-23
36	-22
46	-12
48	-10
58	0
62	4
65	7
68	10
78	20
84	26

What should the sum of these deviation be? Why?

To remove the -sign, let us square each deviation. We now have squared deviation, d^2: 529 484 144 100 0 16 49 100 400 676 and their total, $\sum d^2 = 2498$.

Now find the mean squared deviation, i.e. $\sum d^2/n$where n=10.then $\sum d^2/n=249.8$ this number was derived from the sum of the squares of the deviations of the marks from the mean, so it is in the unlikely units squared marks. To overcome this, we now take the square root which is 15.8 marks. This is the S.D of the standard deviation from the original marks from the mean. Note that it is now the correct units,i.e. marks and that each mark has been included in the calculation of the S.D. .it gives a measure of the dispersion of the marks about the mean

Note: follow this procedure in symbols. Suppose π is the mean of a set of n numbers and x is any number of the set.

I. The deviation of x from the is $d = (x-\pi)$. (this may be + or -)
II. The deviations are squared: $d^2 = (x-\pi)^2$. (This removes the -sign.)
III. Find the sum of the squared deviations: $\sum d^2 = \sum(x-\pi)^2$.
IV. Take the mean squared deviation, $\sum d^2/n$.
V. Then the S.D. s is found: $s=\sqrt{\sum d^2/n} - \sqrt{\sum(x-\pi)^2/n}$

To describe the S.D. briefly, we can say it is the root mean squared deviation from the mean. Now find the S.D. of the Economics marks similarly. You should find s=6.9, thus showing that these marks are less dispersed about the mean.

S.D. of a frequency distribution.

Find the mean and the S.D. of the distribution given in Table Z.

Table Z

x	f
1	4
2	9
3	11
4	13
5	10
6	3
	$\sum f=50$

Mean $(\pi)=4+18+33+52+50+18/50$

$=\sum fx/\sum f=175/50=3.5$

<table>
<tr>
<td>Given data
<table><tr><td>x</td><td>f</td></tr><tr><td>1</td><td>4</td></tr><tr><td>2</td><td>9</td></tr><tr><td>3</td><td>11</td></tr><tr><td>4</td><td>13</td></tr><tr><td>5</td><td>10</td></tr><tr><td>6</td><td>3</td></tr><tr><td></td><td>n = ∑f
=50</td></tr></table></td>
<td>STEP 1
(assumed mean4)
d=x-4
-3
-2
-1
0
1
2</td>
<td>STEP 2
Fd
-12
-18
-11
0
10
6
∑fd=-25</td>
<td>STEP 3
fd²
36
36
11
0
10
12
∑fd²=105</td>
</tr>
</table>

Given data		Step 1 (assumed mean 4) d = x-4	Step 2 fd	Step 3 fd^2
x	f			
1	4	-3	-12	36
2	9	-2	-18	36
3	11	-1	-11	11
4	13	0	0	0
5	10	1	10	10
6	3	2	6	12
	$n=\sum f=50$		$\sum fd=-25$	$\sum fd^2$

Mean $\pi = 4+\sum fd/n = 4-25/60=3.5$

S.D. $=\sqrt{}\sum fd^2/n - (\sum fd/n)^2 = \sqrt{}105/50 - (-25/50)^2 = \sqrt{}2.1 - 0.25 = \sqrt{}1.85 = 136$

The first two columns, x and f are the given data.

STEP 1 Take an assumed mean 4,d=x-4S

STEP 2 Calculate each value of fd (+ or -) and find their sum,∑fd.

This gives the mean as π = assumed mean + ∑fd/n,and also gives the second term in the formular for the S.D., that is $(\sum fd/n)^2$

STEP 3 Calculate fd^2 = fd*d (product of the last two columns) and find $\sum fd^2$.this gives the first term in the fomular for the S.D.,$\sum fd^2/n$

STEP 4 Now work out the S.D.,Take care to get the terms in the correct order $\sum fd^2/n - (\sum fd/n)^2$

STEP 5 Take the square root $\sqrt{}\sum fd^2 - (\sum fd/n)^2$

Calculate the mean and S.D of the distribution show in the Table below.

Table M

x	f
15	1
25	4
i35	12
45	24
55	18
65	8
75	3

Set up the columns required,x,f,d,fd,fd^2

x	f	(Assumed mean = 45) d = x-45	fd	fd^2
		-30	-30	900
		-20	-80	1600
		-10	-1200	1200
		0	0	0
		10	180	1800
		20	160	3200
		30	90	2700
	n=70		∑fd=430-230=200	$\sum fd^2$=11400

The π = 45+200/70 = 45+2.86=47.9

S=$\sqrt{}\sum fd^2/n-(\sum fd/n)^2$ = $\sqrt{}11400/70-(200/70)^2$ = $\sqrt{}162.9-8.16$ = $\sqrt{}154.7$=12.44.the unit of the mean and the S.D. will be the same as those of x.

S.D. of a grouped frequency distribution

Find the mean and S.D. of the distribution given in Table P

Table

Height (cm)	f
150-154	4
155=159	14
160-164	26
165-169	32
170-174	21
175-179	10
180-184	3

Note:

Class limits	Class centers	f	(Assumed mean=167) d=c-167	fd	fd²
150-154	152	4	-15	-60	900
155-159	157	14	-10	-140	1400
160-164	162	26	-5	-130	650
165-169	167	32	0	0	0
170-174	172	21	5	105	525
175-179	177	10	10	100	1000
180-184	182	3	15	45	675
		n=110		∑fd=80	∑fd²=5150

Then mean $\pi=167-80/110=167-0.73=166.3$ cm. and $s=\surd(5150/110-(-80/-110)^2=\surd 46.82-0.53=\surd 46.29=6.89$ cm.

In fig. :

The cumulative frequency polygon has been drawn with the mean marked in position.the limits for +/- S.D. from the mean are 166.3-6.80=159.5 cm and 166.3+6.80=173.5 cm.and these are also marked on the graph.within these limits,there are approximately 90-18=72 values, which is 65% of the total frequency.from the graph find the median and the quartiles and compare the interquartile range with the S.D.

Standard scores

The scores obtained by Agnes,Bayo,Chike and Dede in English,Maths and Geography are shown in Table R

Table R

	Agnes	Bayo	Chike	Dede
English	65	75	70	60
Maths	35	65	90	50
Geography	45	58	50	40

Find the total mark for each one.who came top,2^{nd} ,3^{rd} and 4^{th} according to these totals?

What we must do is to look at marks in each subject and see how they compare on the same standard.We convert them to standard marks or score,taking into account the mean and S.D. of each subject as follows:

Table R

	Mean π	S.D. s
English	70	10
Maths	65	20
Geography	50	8

Convert each mark into the number of standard deviations above or below the mean.for example,65 in English is 5 below the mean,that is -5 marks,and is thus -5/10=-0.5 standard deviations from the mean.

If x is any mark,then it is (x-π)marks away from the mean and therefore x-π/x standard deviations from the mean.this will be +ve if x>π,-ve if x<π.

Now convert all the marks in this way:

Table R

	Agnes	Bayo	Chike	Dede
English	-0.5	0.5	0	-1
maths	-1.5	0	1.25	-0.75
Geography	-.6	1	0	-1.25
Total	-2.6	1.5	1.25	-3

So Bayo is 1^{st} ,Cheke 2^{nd} ,Agnes 3^{rd} and Dede 4^{th} .the means and standard deviations cwere of course deliberately chosen to show up the differences,but the example shows we must be careful in adding marks from distributions with different means and standard deviations.we overcome this by using standard scores.

If x is the original score (raw score),then z = x-π/s is the standard score (or z-score),where π is the mean and s the standard deviation of the raw scores.the mean of the standard scores will now be 0 and the standard deviation 1.standard scores are completely comparable with each other.

If Ojo got 65 marks in Art and 72 in Maths in which subject did he do better relatively?suppose the mean and S.D.were 68 and 10 in Art and 78 and 15 in Maths respectively.

Then z(Art) = 65-68/10 = -0.3 and z(maths) = 72-78/15 = -0.4.Ojo did better in Art than in Maths relatively as -0.3 > -0.4.

The marks in a certain subject have a mean of 64 and a standard deviation of 12.to make them comparable with another subject,the teacher wishes to convert them so that they will have a mean of 58 and a standard deviation of 10.(a) find a formular he can use to do this.(b) if the old marks ranged from 12 to 96 what is the range of the new marks?(c) which mark will be unchanged?

(a) Let x be any mark in the subject and y its new value when converted.Each of this will have the same standard score.then x-64/12 = y-58/10 =5x-320 = 6y-348 or y = 5x+28/6.this is the formula the teacher should use.he could then convert the marks by drawing a straight line graph (fig.):

whose gradient is 5/6 and y-intercept 28/6.for each old mark x he can then read off the converted mark y.

(b) Using the formula 12 becomes (5x12)+28/6=15 and 96 becomes (5x96+28/6=85.so the new range is 85-15=70.note that the gradient of the line y=5x+28/6, i.e. y=5/6x + 28/6 is 5/6 which eguals the ratio of the standard deviations (10/12).note also that the range of y=(5x96)+28/6 – (5x12)+28/6 = (5x84)/6 = 70 and so range of y/range of x = 70/84 = 10/12 =standard deviation of y/standard deviation of x.converting the marks in this way does not of course alter their order.the top mark is still the top mark and so on.

(c) If x = y,then x=5x+28/6 which gives x=28.so the unchanged mark is 28.if we add the line y=x,show dotted on the graph the two lines intersect at x=y=28'

S