18 Statistics

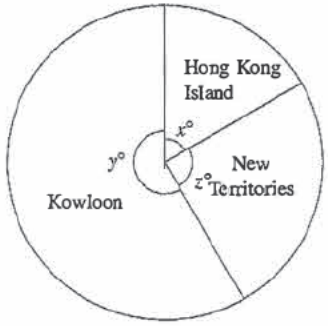
18A Presentation of data

18A.1 HKCEE MA 1982(1) – I – 7

In a certain school, the numbers of students living on Hong Kong Island, in Kowloon and the New Territories are in the ratios 2 : 7 : 3. The pie-chart in the figure shows the distribution.

(a) Find x, y and z.

(b) If the number of students living on Hong Kong Island is 240, find the total number of students in the school.



18A.2 HKCEE MA 1982(3) – I – 12

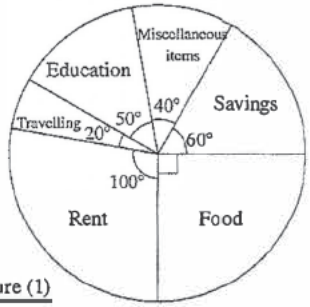


Figure (1)

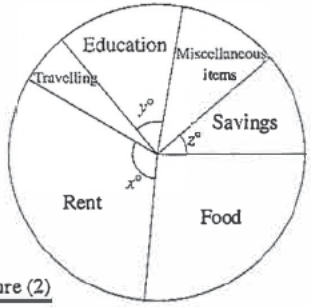


Figure (2)

(a) The pie chart in Figure (1) shows how Mr Wong's income was distributed between his expenses and savings for March. If his rent is \$2000, find Mr Wong's income for that month.

(b) The table below shows the percentage changes when each item of Mr Wong's expenses in April is compared with that in March.

Item	Food	Rent	Travelling	Education	Miscellaneous items	Savings
Percentage Change	Increased by 10%	Increased by 30%	Increased by 30%	No change	No change	?

The pie chart in Figure (2) shows how Mr Wong's income was distributed between his expenses and savings for April.

(i) Suppose that Mr Wong's income in March and April were the same.

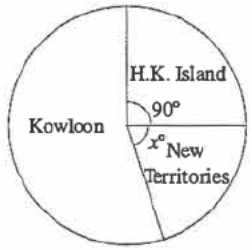
(1) Find x, y and z in Figure (2).

(2) Calculate the percentage change in Mr Wong's savings for April when compared with those for March.

(ii) If Mr Wong's income in April actually increased by 37.5%, what percentage of his income in April was spent on food?

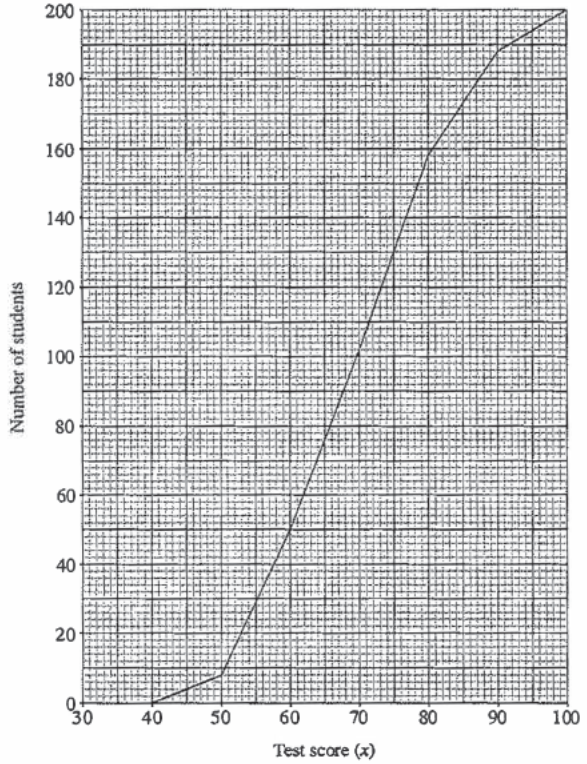
18A.3 HKCEE MA 1985(A/B) – I – 7

The pie-chart in the figure shows the distribution of traffic accidents in Hong Kong in 1983. There were 4200 traffic accidents on H.K. Island, 9240 accidents in Kowloon and n accidents in the New Territories. Find n and x.



18A.4 HKCEE MA 1998 – I – 10

The cumulative frequency polygon of the distribution of test scores of 200 students



Two hundred students took a test in Mathematics. The figure shows the cumulative frequency polygon of the distribution of the test scores.

(a) Complete the tables below.

Test score (x)	Cumulative frequency
$x \leq 50$	8
$x \leq 60$	50
$x \leq 70$	
$x \leq 80$	
$x \leq 90$	188
$x \leq 100$	200

Test score (x)	Frequency
$40 < x \leq 50$	8
$50 < x \leq 60$	42
$60 < x \leq 70$	
$70 < x \leq 80$	
$80 < x \leq 90$	30
$90 < x \leq 100$	12

(b) If the passing score is 55, estimate the passing percentage of the students in the test.

18A.5 HKCEE MA 1999 – I – 11

A school conducted a survey on the placement of her S.5 graduates last year. There were 200 graduates, of which 120 were boys and 80 were girls. The placement of the boys was shown in the figure.

(a) Find the number of boys who repeated S.5.

(b) Among all the boys promoted to S.6, what percentage of them was promoted in their own school?

(c) The result of the survey also showed that 22.5% of the girls were promoted to S.6 in their own school. Find the percentage of graduates promoted to S.6 in their own school.

Placement of the 120 boys

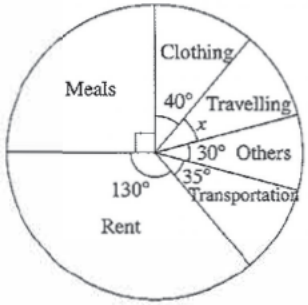
18A.6 HKCEE MA 2006 – I – 9

In the figure, the pie chart shows the expenditure of Ada in February 2006. It is given that she spent \$1750 on transportation in that month. Find

(a) x,

(b) her total expenditure in that month,

(c) her expenditure on travelling in that month.

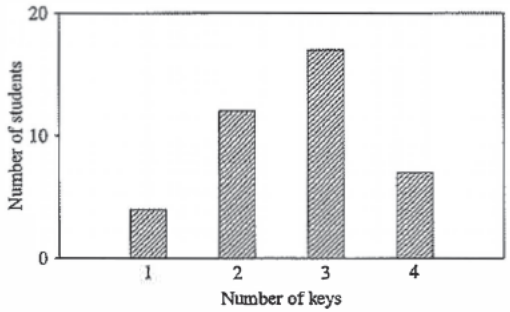


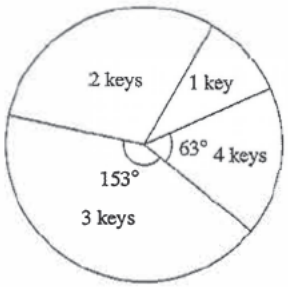
The expenditure of Ada in February 2006

18A.7 HKCEE MA 2007 – I – 12

The bar chart and pie chart in the figure show the distribution of the numbers of keys owned by the students in class A. The numbers of students having 2 keys, 3 keys and 4 keys are 12, 17 and k respectively.

Distribution of the numbers of keys owned by the students in class A





(a) Find the value of k.

(b) Find the number of students in class A.

(c) Find the probability that a randomly selected student in class A has only 1 key.

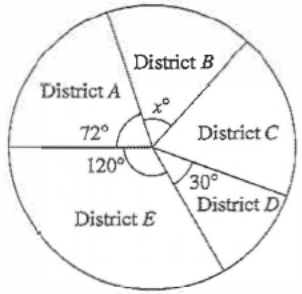
(d) It is given that the numbers of students in class A and class B are the same. The distributions of the numbers of keys owned by the students in class A and class B are also the same. The two classes are now combined to form a group. On each of the bar chart and the pie chart in the figure, is there a modification needed in order that the statistical chart can show the distribution of the numbers of keys owned by the students in this group? If your answer is 'yes', write down the modification needed.

18A.8 HKDSE MA SP – I – 9

In the figure, the pie chart shows the distribution of the numbers of traffic accidents occurred in a city in a year. In that year, the number of traffic accidents occurred in District A is 20% greater than that in District B.

(a) Find x.

(b) Is the number of traffic accidents occurred in District A greater than that in District C? Explain your answer.

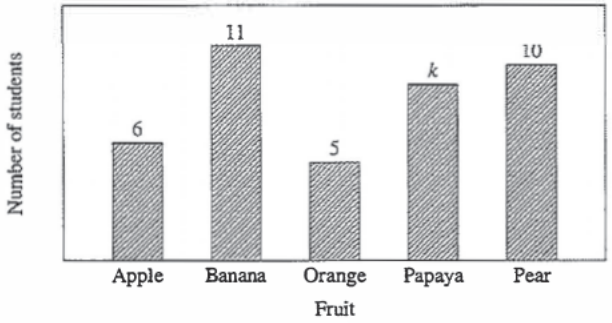


The distribution of the numbers of traffic accidents occurred in the city

18A.9 HKDSE MA PP – I – 13 (Continued from **17B.34.**)

The bar chart below shows the distribution of the most favourite fruits of the students in a group. It is given that each student has only one most favourite fruit.

Distribution of the most favourite fruits of the students in the group



If a student is randomly selected from the group, the probability that the most favourite fruit is apple is $\frac{3}{20}$.

(a) Find k.

(b) Suppose that the above distribution is represented by a pie chart.

(i) Find the angle of the sector representing that the most favourite fruit is orange.

(ii) Some new students now join the group and the most favourite fruit of each of these students is orange. Will the angle of the sector representing that the most favourite fruit is orange be doubled? Explain your answer.

18A.10 HKDSE MA 2016 – I – 9 (To continue as **17B.40.**)

The frequency distribution table and the cumulative frequency distribution table below show the distribution of the heights of the plants in a garden.

Height (m)	Frequency
0.1 – 0.3	a
0.4 – 0.6	4
0.7 – 0.9	b
1.0 – 1.2	c
1.3 – 1.5	15
1.6 – 1.8	3

Height less than (m)	Cumulative frequency
0.35	2
0.65	x
0.95	13
1.25	y
1.55	37
1.85	z

(a) Find x, y and z.

18B Measures of central tendency

18B.1 (HKCEE MA 1983(B) I 3)

The table shows the distribution of the marks of 1000 students in a mathematics test:

Class of Marks	Number of Students
40 – 49	100
50 59	300
60 – 69	400
70 79	200

(a) Find the class mark of the class 50 – 59.

(b) Estimate the mean of the distribution of marks.

18B.2 HKCEE MA 1984(A/B) – I 2

The table shows the distribution of the marks of a group of students in a short test:

If the mean of the distribution is 3, find the value of x.

Marks	1	2	3	4	5
Number of Students	10	10	5	20	x

18B.3 HKCEE MA 1986(A/B) I – 3

The table shows the number of students in three classes of a school and their average marks in a test.
If the overall average mark of the three classes is 60, find x.

Class	No. of Students	Average Mark
F.5A	40	61
F.5B	x	70
F.5C	35	50

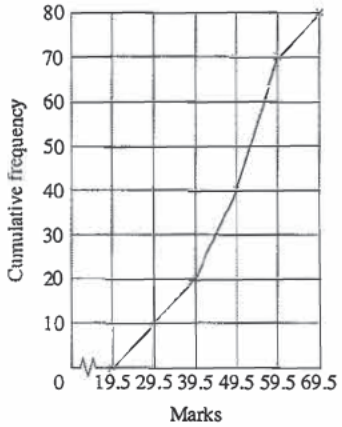
18B.4 HKCEE MA 1991 – I – 1

In the figure, the cumulative frequency polygon shows the distribution of the marks of 80 students in a Mathematics test.

(a) From the figure, write down the median of the distribution.

(b) Complete the table below.
Hence find the mean mark of the students in the test.

Marks	No. of students
20 – 29	
30 – 39	
40 49	
50 – 59	
60 – 69	



18B.5 HKCEE MA 1992 I 8

In a sports competition, the mean score of a team of m men and n women is 70.

(a) Find the total score of the team in terms of m and n.

(b) If the mean score of the men is 75 and the mean score of the women is 62, find the ratio $m : n$.

(c) If there are altogether 39 persons in the team, find the number of men.

18B.6 HKCEE MA 1994 I 1(d)

The marks scored by eleven students in a mathematics quiz are as follows:

10 20 30 45 50 60 65 65 65 70 70.

Find (i) the mean, (ii) the mode and (iii) the median of the above marks.

18B.7 HKCEE MA 1996 I 14

A youth centre has done a survey on the amount of money $\$x$ teenagers spent on buying clothes for Christmas. The results of the survey are shown in Tables (1) and (2).

Table (1) **The amount of money spent by *boys* on buying clothes for Christmas**

x	Frequency	Percentage (%)
0	70	20.0
$0 < x \leq 200$	17	4.9
$200 < x \leq 400$	48	13.7
$400 < x \leq 600$	83	[illegible]
$600 < x \leq 800$	92	26.3
$800 < x \leq 1000$	36	10.3
$x > 1000$	4	1.1
Total frequency =	350	

Table (2) **The amount of money spent by *girls* on buying clothes for Christmas**

x	Frequency	Percentage (%)
0	81	15.0
$0 < x \leq 200$	51	9.4
$200 < x \leq 400$	135	25.0
$400 < x \leq 600$	87	16.1
$600 < x \leq 800$	74	13.7
$800 < x \leq 1000$	56	10.4
$x > 1000$	57	10.5
Total frequency =	541	

(a) A number in Table (1) was accidentally covered in ink. What should this number be?

(b) Explain why the sum of the percentages in Table (2) is 100.1 instead of 100.

(c) The cumulative frequency polygon of the distribution of x ($x \leq 1000$) for girls is drawn in Figure (3).

(i) Construct the cumulative frequency table of the distribution of x ($x \leq 1000$) for boys.

(ii) On the same graph (Figure (3)), draw the cumulative frequency polygon of the distribution in (i).

(iii) Find the medians of x for boys and girls respectively in this survey.

(iv) Estimate the total number of teenagers in this survey spending not more than \$700 on buying clothes for Christmas.

(d) By considering the percentages in Tables (1) and (2), find evidence to support the statement:
"In this survey, more boys did not spend any money on buying clothes for Christmas."
Explain briefly why we have to consider the percentages instead of the frequencies.

The cumulative frequency polygon of the distribution of x ($x \leq 100$) for girls

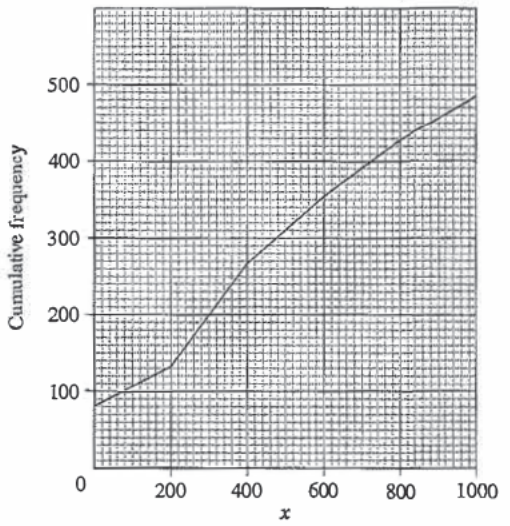


Figure (3)

18B.8 HKCEE MA 1999 – I – 8

The heights of 6 students are x cm, 161 cm, 168 cm, 159 cm, 161 cm and 152 cm. The mean height of these students is 158 cm.

(a) Find x.

(b) Find the median of the heights of these students.

18B.9 HKCEE MA 2000 – I – 11

The figure shows the cumulative frequency polygon of the distribution of the lengths of 75 songs.

(a) Complete the tables below.

Length (t seconds)	Cumulative frequency
$t \leq 220$	3
$t \leq 240$	16
$t \leq 260$	46
$t \leq 280$	
$t \leq 300$	75

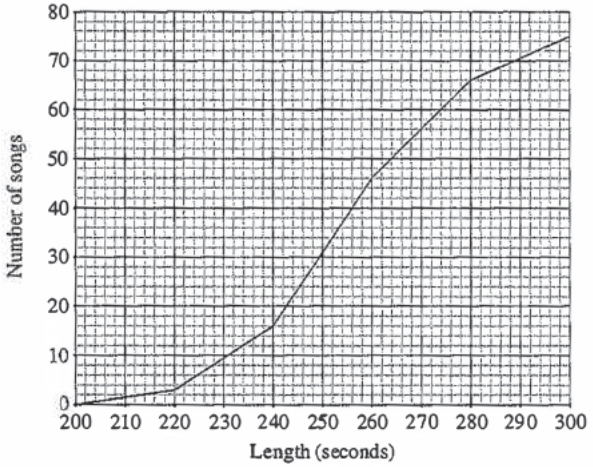
Length (t seconds)	Frequency
$200 < t \leq 220$	3
$220 < t \leq 240$	13
$240 < t \leq 260$	30
$260 < t \leq 280$	
$280 < t \leq 300$	9

(b) Find an estimate of the mean of the distribution.

(c) Estimate from the cumulative frequency polygon the median of the distribution.

(d) What percentage of these songs have lengths greater than 220 seconds but not greater than 260 seconds?

The cumulative frequency polygon of the distribution of the lengths of 75 songs



18B.10 HKCEE MA 2003 – I – 11

(a) For the set of data 10, 10, 11, 12, 13, 16, find

(i) the mode,

(ii) the median,

(iii) the mean,

(iv) the range.

(b) Four unknown data are combined with the six data in (a) to form a set of ten data.

(i) Find the least and the greatest possible values of the median of the combined set of ten data.

(ii) If the mean of the four unknown data is 11, find the mean of the combined set of ten data.

18B.11 HKCEE MA 2006 – I – 8 (To continue as **17B.8.**)

There are ten cards numbered 2, 3, 5, 8, 11, 11, 12, 15, 19 and k respectively, where k is a positive integer. It is given that the mean of the ten numbers is 11.

(a) Find the value of k.

18B.12 HKALE MS 1998 – 3 (To continue as **17C.27.**)

40 students participate in a 5-day summer camp. The stem and leaf diagram below shows the distribution of heights in cm of these students.

(a) Find the median of the distribution of heights.

Stem (tens)	Leaf (units)
13	8
14	1 5 6 9
15	0 1 3 4 4 4 5 5 6 7 8 8 9
16	1 1 2 3 3 4 5 6 7 7 8 8
17	0 2 2 3 4 5 6 7
18	1 4

18B.13 HKALE MS 2002 7 (To continue as **17C.30.**)

Twenty two students in a class attended an examination. The stem-and-leaf diagram below shows the distribution of the examination marks of these students.

(a) Find the mean of the examination marks.

(b) Two students left the class after the examination and their marks are deleted from the stem-and-leaf diagram. The mean of the remaining marks is then increased by 1.2 and there are two modes. Find the two deleted marks.

Stem (tens)	Leaf (units)
3	5 7
4	2 4 6
5	0 3 4 4 4 5
6	1 2 5 5 8
7	3 8 9
8	4 8
9	5

18B.14 HKALE MS 2010 – 5 (To continue as **17C.34.**)

The following stem-and-leaf diagram shows the distribution of the test scores of 21 students taking a statistics course. Let $\bar{x}$ be the mean of these 21 scores.

It is known that if the smallest value of these 21 scores is removed, the range is decreased by 27 and the mean is increased by 2.

(a) Find the values of a, b and $\bar{x}$.

Stem (Tens)	Leaf (Units)
2	a
3	
4	9
5	0 0 1 3 7 7
6	0 2 3 5 5 5 9
7	0 3 4 9
8	2 b

18B.15 HKDSE MA SP – I 14

The data below show the percentages of customers who bought newspaper A from a magazine stall in city H for five days randomly selected in a certain week:

62% 63% 55% 62% 58%

(a) Find the median and the mean of the above data.

(b) Let a% and b% be the percentages of customers who bought newspaper A from the stall for the other two days in that week. The two percentages are combined with the above data to form a set of seven data.

(i) Write down the least possible value of the median of the combined set of seven data.

(ii) It is known that the median and the mean of the combined set of seven data are the same as that found in (a). Write down one pair of possible values of a and b.

(c) The stall-keeper claims that since the median and the mean found in (a) exceed 50%, newspaper A has the largest market share among the newspapers in city H. Do you agree? Explain your answer.

18B.16 HKDSE MA 2012 – I – 10

Tom conducts a survey on the numbers of hours spent on doing homework in a week by secondary students. Questionnaires are sent out and twenty of them are returned. The stem-and-leaf diagram below shows the numbers of hours recorded in the twenty questionnaires:

Stem (tens)	Leaf (units)
1	0 0 1 1 2 3 4 5 5 6 6 7 7
2	0 0 0 5 8
3	4 6

(a) Find the mean and the median of the numbers of hours recorded in the twenty questionnaires.

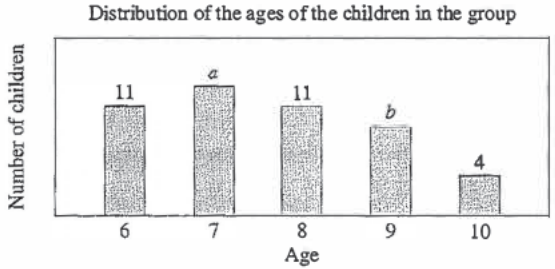
(b) Tom receives four more questionnaires. He finds that the mean of the numbers of hours recorded in these four questionnaires is 18. It is found that the numbers of hours recorded in two of these four questionnaires are 19 and 20.

(i) Write down the mean of the numbers of hours recorded in the twenty four questionnaires.

(ii) Is it possible that the median of the numbers of hours recorded in the twenty-four questionnaires is the same as the median found in (a)? Explain your answer.

18B.17 HKDSE MA 2016 – I – 12

The bar chart below shows the distribution of the ages of the children in a group, where $a > 11$ and $4 < b < 10$. The median of the ages of the children in the group is 7.5.



(a) Find a and b.

(b) Four more children now join the group. It is found that the ages of these four children are all different and the range of the ages of the children in the group remains unchanged. Find

(i) the greatest possible median of the ages of the children in the group,

(ii) the least possible mean of the ages of the children in the group.

18B.18 HKDSE MA 2018 – I – 11

The following table shows the distribution of the numbers of children of some families:

Number of children	0	1	2	3	4
Number of families	k	2	9	6	7

It is given that k is a positive integer.

(a) If the mode of the distribution is 2, write down

(i) the least possible value of k;

(ii) the greatest possible value of k.

(b) If the median of the distribution is 2, write down

(i) the least possible value of k;

(ii) the greatest possible value of k.

(c) If the mean of the distribution is 2, find the value of k.

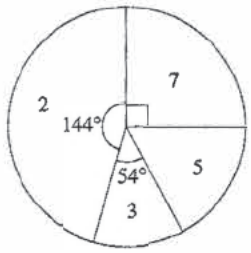
18B.19 HKDSE MA 2019 – I 8

(To continue as 17B.46.)

The pie chart below shows the distribution of the numbers of rings owned by the girls in a group.

(a) Write down the mode of the distribution.

(b) Find the mean of the distribution.



Distribution of the numbers of rings owned by the girls in the group

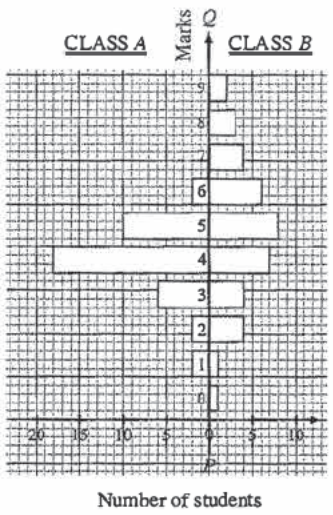
18C Measures of dispersion

18C.1 HKCEE MA 1980(3) I – 8

Two classes, A and B, each of 40 students, took a test. In the test, students may score 0, 1, 2, 3, 4, 5, 6, 7, 8 or 9 marks. In the figure, the distribution of marks of class A is shown in the bar chart on the left of PQ and that of class B is shown on the right.

(a) Find, by inspection, which class has a greater standard deviation of marks.

(b) If 70 students from the two classes pass the test, what is the minimum mark that a student should get in order to obtain a pass?



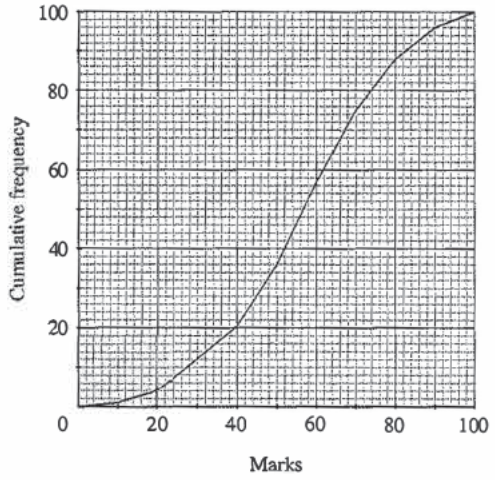
18C.2 HKCEE MA 1981(1) – I 6

The figure shows the cumulative frequency polygon of the marks obtained by 100 students taking a mathematics test.

(a) If 75% of the students pass the test, what is the pass mark, correct to the nearest integer?

(b) If the pass mark were 40, how many students would pass the test?

(c) Find the inter quartile range.



18C.3 HKCEE MA 1983(A) I – 3

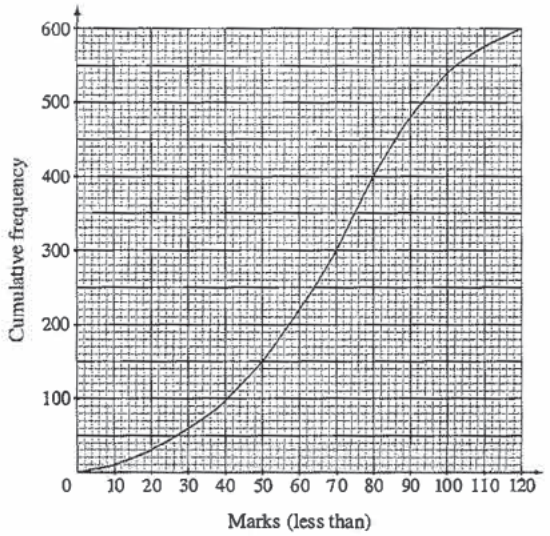
Given five real numbers $a-6$, a, $a+2$, $a+3$, $a+6$, find

(a) the mean,

(b) the standard deviation.

18C.4 HKCEE MA 1988 – I 11 (To continue as 17C.7.)

The figure below shows the cumulative frequency curve of the marks of 600 students in a mathematics contest.



(a) From the curve, find

(i) the median, and

(ii) the interquartile range of the distribution of marks.

18C.5 HKCEE MA 1990 I 12

(a) The distribution of the monthly salaries of 100 employees in a firm is shown in the histogram in the figure.

(i) Find the modal class, median, mean, interquartile range and ***mean deviation (out of syllabus)*** of the monthly salaries of the 100 employees.

(ii) Now the firm employs 10 more employees whose monthly salaries are all \$6500. Will the standard deviation of the monthly salaries of all the employees in the firm become greater, smaller or remain unchanged? Explain briefly.

(b) The mean of 7 numbers $x_1, x_2, \ldots, x_7$ is $\bar{x}$ and the *squares* of the deviations from $\bar{x}$ are 9, 4, 1, 0, 1, 4, 9 respectively. Find the standard deviation of the 7 numbers.
[not mandatory]

Distribution of monthly salaries of 100 employees

18C.6 HKCEE MA 1993 – I – 7

The following frequency table shows the distribution of the scores of 200 students in a Mathematics examination.

Frequency Table

Score	Frequency
0 – 9	20
10 – 19	40
20 – 29	60
30 39	50
40 49	20
50 59	10

Cumulative Frequency Table

Score (less than)	Cumulative Frequency
9.5	
19.5	
29.5	
39.5	
49.5	
59.5	

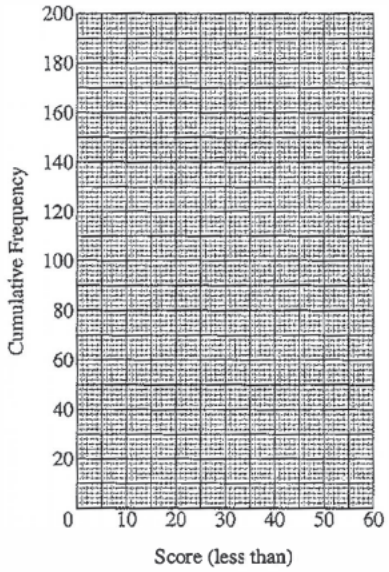
(a) Fill in the cumulative frequency table.

(b) (i) Draw the cumulative frequency polygon on the graph paper and determine the interquartile range.

(ii) If the pass percentage is set at 60%, determine the pass score from the cumulative frequency polygon.

(c) Find the mean and standard deviation of the distribution of scores. (Working steps need not be shown.)

(d) The teacher found that the scores were too low. He added 20 to each score. Write down the mean and the standard deviation of the new set of scores.



18C.7 HKCEE MA 1995 I – 9

The cumulative frequency polygon in the figure shows the distribution of the yearly average scores of all the Secondary 2 students in School A.

(a) Find

(i) the total number of Secondary 2 students in School A;

(ii) the median of the yearly average scores, correct to the nearest integer.

(b) The students will be allocated to 3 different groups in Secondary 3 according to their yearly average scores. The top 25% will be in Group I and the bottom 25% will be in Group III. The rest will be in Group II. Find, correct to the nearest integer.

(i) the minimum yearly average score for students to be allocated to Group I;

(ii) the minimum yearly average score for students to be allocated to Group II.

(c) Fill in the class marks and frequencies in the table.

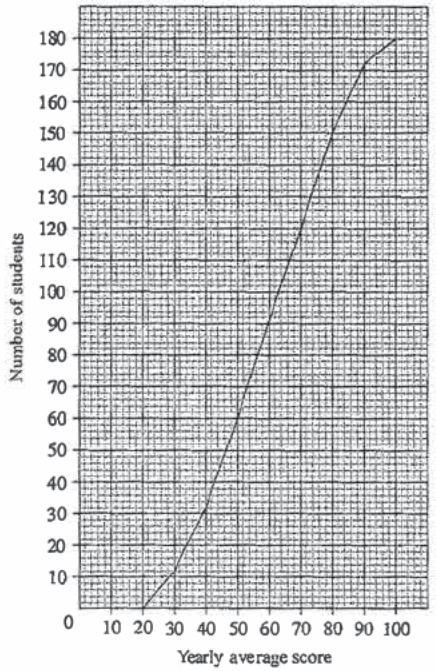
(d) From the table, find the mean and standard deviation of the yearly average scores. (Working need not be shown.)

(e) Find the percentage of students whose yearly average scores are within one standard deviation from the mean. (The distribution of the yearly average scores is not necessarily a normal distribution.)

The frequency distribution table of the yearly average scores of all the Secondary 2 students in School A

Yearly average score (x)	Class mark	Frequency
$20 < x \leq 30$	25	
$30 < x \leq 40$		20
$40 < x \leq 50$		
$50 < x \leq 60$		32
$60 < x \leq 70$		
$70 < x \leq 80$		30
$80 < x \leq 90$		22
$90 < x \leq 100$	95	

The cumulative frequency polygon of the yearly average scores of all the Secondary 2 students in School A



18C.8 HKCEE MA 1997 – I – 11

The following are the marks scored by a class of 35 students in a Mathematics test:

0	0	5	8	11	12	41	42	45	48
50	62	70	73	73	73	77	78	80	80
82	82	82	83	83	85	85	87	90	90
95	95	95	95	98					

(a) Find the mean, mode, median and standard deviation of the above marks. (Working need not be shown.)

(b) Explain briefly why the mean may not be a suitable measure of central tendency of the distribution of marks in the Mathematics test.

(c) The mean and standard deviation of the marks scored by the same class of students in an English test are 63 and 15 respectively.

(i) The standard score of a student in the English test was 0.4. Find the mark the student scored in this test.

(ii) Assume that the marks in the English test are normally distributed and the marks scored by Lai Wah in both the Mathematics and English tests are 78.

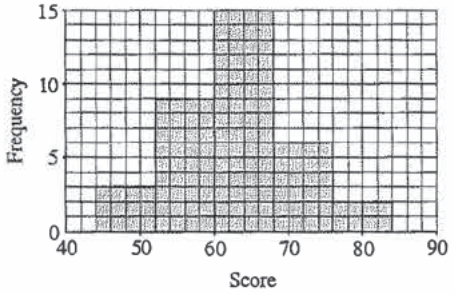
(1) What percentage of her classmates scored fewer marks than Lai Wah in the Mathematics test?

(2) Relative to her classmates, did Lai Wah perform better in the English test than in the Mathematics test?

(iii) The English teacher later found that a student was given 10 marks fewer in the English test. Find the mean of the marks in the English test after the wrong mark has been corrected.

18C.9 HKCEE MA 2001 – I – 10

Distribution of scores of 40 students



The histogram in the figure shows the distribution of scores of a class of 40 students in a test.

(a) Complete the table.

Score (x)	Class mark	Frequency
$44 \le x < 52$		3
$52 \le x < 60$		
	64	15
$68 \le x < 76$		11
	80	

(b) Estimate the mean and standard deviation of the distribution.

(c) Susan scores 76 in this test. Find her standard score.

(d) Another test is given to the same class of students. It is found that the mean and standard deviation of the scores in this second test are 58 and 10 respectively. Relative to her classmates, if Susan performs equally well in these two tests, estimate her score in the second test.

18C.10 HKCEE MA 2002 – I – 5

For the set of data 4, 4, 5, 6, 8, 12, 13, 13, 13, 18, find

(a) the mean,

(b) the mode,

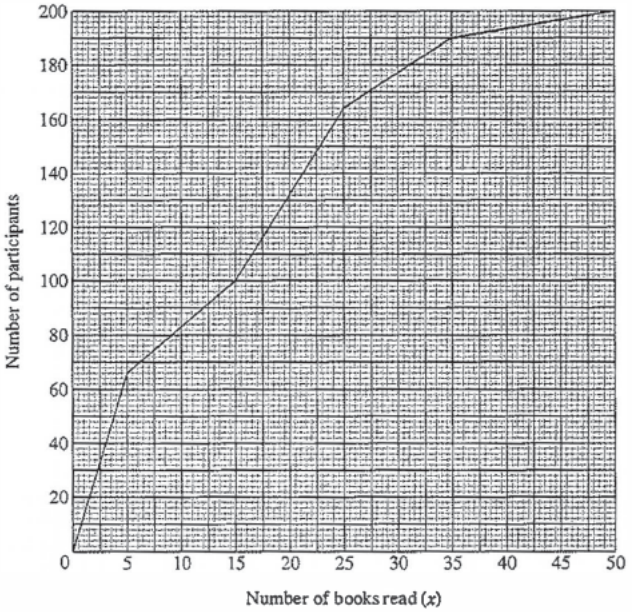
(c) the median,

(d) the standard deviation.

18C.11 HKCEE MA 2002 I – 12

(To continue as 17C.16.)

The cumulative frequency polygon of the distribution of the numbers of books read by the participants



Two hundred students participated in a summer reading programme. The figure shows the cumulative frequency polygon of the distribution of the numbers of books read by the participants.

(a) The table below shows the frequency distribution of the numbers of books read by the participants. Using the graph in the figure, complete the table.

Number of books read (x)	Number of participants	Award
$0 < x \le 5$	66	Certificate
$5 < x \le 15$		Book coupon
$15 < x \le 25$	64	Bronze medal
$25 < x \le 35$		Silver medal
$35 < x \le 50$	10	Gold medal

(b) Using the graph in the figure, find the inter quartile range of the distribution.

18C.12 HKCEE MA 2004 – I – 11

A large group of students sat in a Mathematics test consisting of two papers, Paper I and Paper II. The table below shows the mean, median, standard deviation and range of the test marks of these students in each paper:

Test paper	Mean	Median	Standard deviation	Range
Paper I	46.1 marks	46 marks	15.2 marks	91 marks
Paper II	60.3 marks	60 marks	11.6 marks	70 marks

A student, John, scored 54 marks in Paper I and 66 marks in Paper II.

(a) Assume that the marks in each paper of the Mathematics test are normally distributed. Relative to other students, did John perform better in Paper II than in Paper I? Explain your answer.

(b) In a mark adjustment, the Mathematics teacher added 4 marks to the test mark of Paper I for each of these students. Write down the mean, the median and the range of the test marks of Paper I after the mark adjustment.

18C.13 HKCEE MA 2005 – I – 15

The scores (in marks) obtained by a class of 20 students in a music test are shown below:

84 86 90 93 100
103 120 120 120 121
122 134 134 136 137
144 146 146 146 158

(a) Find the mean, the ***mean deviation (out of syllabus)*** and the standard deviation of the above scores.

(b) Mary is one of the students in the class and her standard score in the music test is 1. Is Mary one of the top 20% students of the class in the music test? Explain your answer.

(c) (i) If one student in the class withdraws, find the probability that the mean of the scores obtained by the remaining 19 students in the music test is 122 marks.

(ii) If two students in the class withdraw, find the probability that the mean of the scores obtained by the remaining 18 students in the music test is 122 marks.

18C.14 HKCEE MA 2006 – I – 14 (To continue as **17C.19.**)

The stem and leaf diagrams below show the distributions of the scores (in marks) of the students of classes A and B in a test, where a, b, c and d are non negative integers less than 10. It is given that each class consists of 25 students.

Class A

Stem (tens)	Leaf (units)
0	a 9
1	2 5 7 8 8
2	3 3 5 6 7 9
3	2 3 5 6 9 9 9
4	1 2 2 4 b

Class B

Stem (tens)	Leaf (units)
0	c 3 3 4 5
1	1 1 2 2 3 3 5 6 7 8
2	1 1 5 5 5 7 8
3	5 9
4	d

(a) (i) Find the inter quartile range of the score distribution of the students of class A and the inter quartile range of the score distribution of the students of class B.

(ii) Using the results of (a)(i), state which one of the above score distributions is less dispersed. Explain your answer.

18C.15 HKCEE MA 2007 – I – 4

The stem and leaf diagram below shows the distribution of weights (in kg) of 15 teachers in a school.

Stem (tens)	Leaf (units)
5	0 5 5 5 8
6	2 3 7 8 8 9
7	1 3 3 5

Find the median, the range and the standard deviation of the distribution.

18C.16 HKCEE MA 2008 – I – 10

The frequency distribution table and the cumulative frequency distribution table below show the distribution of the weights of the 50 babies born in a hospital during the last week, where a, b, c, k, l and m are integers.

Weight (kg)	Frequency
2.6 – 2.8	a
2.9 – 3.1	12
3.2 – 3.4	b
3.5 – 3.7	10
3.8 – 4.0	c

Weight less than (kg)	Cumulative Frequency
2.85	4
3.15	k
3.45	37
3.75	l
4.05	m

(a) Find a, b and c.

(b) Find estimates of the mean and the standard deviation of the weights of the 50 babies born in the hospital during the last week.

18C.17 HKCEE MA 2008 – I – 14 (Continued from **17C.21.**)

The stem-and-leaf diagram below shows the suggested bonuses (in dollars) of the 36 salesgirls of a boutique:

Stem (thousands)	Leaf (hundreds)
2	4 4 7
3	2 5 6 6 8
4	3 3 3 4 4 7 8 8 8
5	0 0 3 4 4 6
6	2 3 3 4 4 9 9
7	0 4 4 8
8	2 3

(a) The suggested bonus of each salesgirl of the boutique is based on her performance. The following table shows the relation between level of performance and suggested bonus:

Level of performance	Suggested bonus ($x)
Excellent	$x > 6500$
Good	$4500 < x \leq 6500$
Fair	$x \leq 4500$

(i) From the 36 salesgirl, one of them is randomly selected. Given that the level of performance of the selected salesgirl is good, find the probability that her suggested bonus is less than $5500.

(ii) From the 36 salesgirls, two of them are randomly selected.

(1) Find the probability that the level of performance of one selected salesgirl is excellent and that of the other is good.

(2) Find the probability that the levels of performance of the two selected salesgirls are different.

(b) (i) Find the median and the inter quartile range of the suggested bonuses of the 36 salesgirls.

(ii) The boutique has made a considerable profit and so the manager wants to raise the suggested bonus of each of the 36 salesgirls such that the median of the suggested bonuses will be increased by 20% and the inter-quartile range will remain unchanged. Describe how the manager should raise the suggested bonus of each of the 36 salesgirls.

18C.18 HKCEE MA 2009 – I – 10

The stem-and leaf diagram below shows the distribution of the typing speed (in words per minute) of 20 students in a school before training.

Stem (tens)	Leaf (units)
1	2 2 6 7
2	1 1 3 3 3 5 7 8
3	2 4 5 5 8 9 9 9

(a) Find the median, the range and the inter-quartile range of the above distribution.

(b) The box-and-whisker diagram below shows the distribution of the typing speed (in words per minute) of the 20 students after the training.

(i) Is the distribution of the typing speed after the training more dispersed than that before the training? Explain your answer.

(ii) The trainer claims that not less than half of these students show improvement in their typing speed after the training. Do you agree? Explain your answer.

18C.19 HKCEE MA 2010 – I – 11

The stem-and leaf diagram shows that ages of the players of a football team:

Stem (tens)	Leaf (units)
1	8 9 9
2	0 1 1 1 3 3 5 6 6 7 7 8 8 8 8
3	0 0 1 1

(a) Find the mean, the median and the range of the ages of the players of the football team.

(b) As the two oldest players leave the team, three new players join the football team. After the three players join the football team, the manager of the team finds that the mean age of the players of the football team is the same as the mean found in (a).

(i) Find the mean age of the three new players.

(ii) Furthermore, the manager finds that the median and the range of the ages of the players of the football team are the same as the median and the range found in (a) respectively. Write down two sets of possible ages of the three new players.

18C.20 HKCEE MA 2011 I – 10

The student union of a school conducts two surveys to measure the extent of the students' satisfaction on the services provided by the school library. A score from 0 to 100 is used to measure the extent of satisfaction on the services, with 0 indicating absolute dissatisfaction and 100 indicating absolute satisfaction. The stem-and-leaf diagram below shows the distribution of scores rated by 32 students in the first survey.

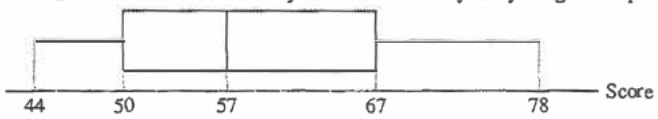
Stem (tens)	Leaf (units)
2	3 3
3	2 4 6 6
4	2 3 3 5 5 7
5	1 6 6 7 7 8 8 8 8
6	3 3 5 5 6 6 7 7 9 9
7	5

(a) Find the median, the range and the inter-quartile range of the above distribution.

(b) After six months, the student union conducts the second survey to these 32 students. The box-and-whisker diagram below shows the distribution of scores rated by these students in the second survey.

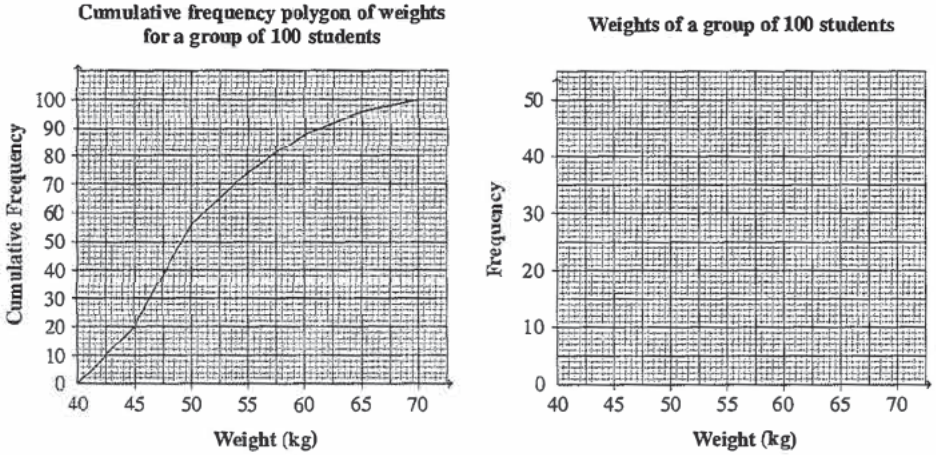
(i) Is the distribution of scores in the second survey less dispersed than the first survey? Explain your answer.

(ii) The chairman of the student union claims that at least 25% of these students have a greater extent of satisfaction shown in the second survey than the first survey. Do you agree? Explain your answer.



18C.21 HKALE MS 1994 – 4

The figure shows the cumulative frequency polygon of weights (in kg) for a group of 100 students.



(a) Use the graph paper provided to draw a histogram of the weights.

(b) Determine the inter-quartile range of the weights from the cumulative frequency polygon.

(c) Determine the mean weight from the histogram.

18C.22 HKALE MS 1995 – 1

The numbers of hours spent by 25 students in studying for an examination are as follows:

11	8	25	21	18	25	7	32	29	18
18	18	22	12	5	30	19	15	20	50
25	10	26	23	12					

Stem (in 10)	Leaf (in 1)
0	5 7 8
1	
2	
3	
4	
5	

(a) Complete the stem and leaf diagram for the above data.

(b) Find the mode, the median and the interquartile range of the numbers of hours spent by the 25 students.

18C.23 HKALE MS 1996 1

A stem-and-leaf diagram for the test scores of 30 students is shown.

Stem (tens)	Leaf (units)
1	0
2	
3	0 2
4	4 5 8 9
5	0 1 2 6 8 8 9 9
6	2 3 3 5 5 8
7	1 2 2 4 4 4
8	2 5
9	1

(a) Find the mean, mode and interquartile range of these scores.

(b) If the score 71 is an incorrect record and the correct score is 11, which of the statistics in (a) will have different values? Find the correct values of these statistics.

18C.24 HKALE MS 1997 – 2

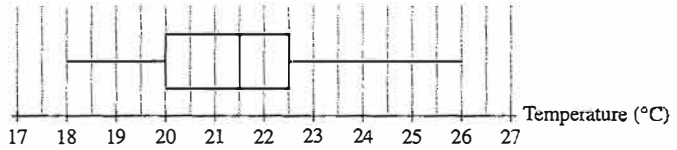
In an experiment, temperatures of a certain liquid under various experimental settings are measured. The box and whisker diagram for these temperatures (in °C) is constructed below.

(a) Find the range (in °C) of the temperatures.

(b) The temperature C (in °C) can be converted to the temperature F (in °F) according to the formula $F = \frac{9}{5}C + 32$.

(i) Find the median and interquartile range of the temperatures in °F.

(ii) If the mean and standard deviation of the temperatures are 22°C and 2°C respectively, find their values in °F.



18C.25 HKALE MS 1999 – 3

A test was carried out to see how quickly a class of students reacted to a visual instruction to press a particular key when they played a computer game. Their reaction times, measured in tenths of a second, are recorded and the statistics for the whole class are summarised below.

	Lower quartile	Upper quartile	Median	Minimum	Maximum
Boys	8	14	11	5	17
Girls	9	16	11	7	21

(a) Draw two box-and whisker diagrams comparing the reaction times of boys and girls.

(b) Suppose a boy and a girl are randomly selected from the class. Which one will have a bigger chance of having a reaction time shorter than 1.1 seconds? Explain.

18C.26 HKALE MS 2000 – 5

A fitness centre advertised a programme specifically designed for women weighing 70 kg or more, and claimed that their individual weights could be reduced by at least 20 kg on completion of the programme. 21 women joined the programme and their weights in kg when they started are shown.

Stem (tens)	Leaf (units)
7	0 0 2 3 5 5 7
8	1 1 4 5 6 6 7 8
9	0 2 5 8 9 9

(a) Find the median and the interquartile range of these weights.

(b) On completion of the programme, the median, lower quartile and upper quartile of the weights of these women are 73 kg, 68 kg and 77 kg respectively. The lightest and heaviest women weigh 60 kg and 82 kg respectively Draw two box and-whisker diagrams comparing the weights of these women before and after the programme.

(c) Referring to the box and-whisker diagram in (b), someone claimed that none of these women had reduced their individual weights by 20 kg or more on completion of the programme. Determine whether this claim is correct or not. Explain your answer briefly.

18C.27 HKALE MS 2001 – 3

The ages of 35 members of a golf club are shown below. It is known that the median and the range of the ages are 36 and 48 respectively, and the ages of the two eldest members differ by 1.

(a) Find the unknown digits a, b and c.

(b) The three members whose ages correspond to the three unknown digits a, b and c are replaced with three new members with ages 12, 38 and 68 respectively. Draw two box and whisker diagrams comparing the age distributions of the members before and after replacement.

Stem (tens)	Leaf (units)
1	a 8 8 9 9
2	0 1 2 3 3 4 7 8
3	1 2 2 5 b 9 9
4	0 2 5 5 6
5	2 2 5 5 8 8
6	0 1 c 6

18C.28 HKALE MS 2003 5

A researcher conducted a study on the time (in minutes) spent on using the Internet by university students. Thirty questionnaires were sent out and only 19 were returned. The results are as follows:

12 13 14 15 15 21 25 29
36 37 38 41 47 49 49 49
52 54 57

(a) Construct a stem and leaf diagram for these data.

(b) Suppose that the research has received eight more questionnaires. Three of them show that the time spent on using the Internet is one hour. The other show that the time spent is more than one hour.

(i) Find the revised median and the revised interquartile range of the time spent.

(ii) Describe briefly the change in the mean and the change in the range of the time spent.

18C.29 HKALE MS 2004 – 5

Some statistics from a survey on the monthly incomes (in thousands of dollars) of a group of university graduates are summarised in the table.

(a) Using the above information, construct a box and whisker diagram to describe the distribution of the monthly incomes.

(b) A student proposes to model the distribution of the monthly incomes of the group of university graduates by a normal distribution with mean and standard deviation given in the table.

(i) *[Out of syllabus]*

(ii) Is the model proposed by the student appropriate? Explain your answer.

Minimum	8
Maximum	52
Lower quartile	10
Median	17
Upper quartile	20
Mean	17.94
Standard deviation	4.7

18C.30 HKALE MS 2005 – 4

The stem and leaf diagram below shows the distribution of heights in cm of 32 students.
It is found that three records less than 150 cm are incorrect. Each of them should be 10 cm greater than the original record. Find the change in each of the following statistics after correcting the three records:

(a) the mean,

(b) the median,

(c) the mode,

(d) the range,

(e) the interquartile range.

Stem (tens)	Leaf (units)
14	5 5 6 6
15	1 2 2 4 4 5 5 7 7 7 7 7 9
16	0 2 2 5 6 7 8 8 9
17	0 1 2 3 4 4

18C.31 HKALE MS 2006 – 4

The stem-and-leaf diagram shows the distribution of the numbers of books read by 24 students of a school in the first term.

Stem (tens)	Leaf (units)
0	3 4 6 7
1	1 2 2 3 5 6 7 8 8 9
2	1 3 4 5 5 7 8 9
3	0 0

(a) Find the median and the interquartile range of the numbers of books read.

(b) The librarian of the school ran a reading award scheme in the second term. The following table shows some statistics of the distribution of the numbers of books read by these 24 students in the second term:

Minimum	Lower quartile	Median	Upper quartile	Maximum
8	26	35	41	46

(i) Draw two box-and-whisker diagrams of the same scale to compare the numbers of books read by these students in the first term and in the second term.

(ii) The librarian claims that not less than 50% of these students read at least 5 more books in the second term than that in the first term. Do you agree? Explain your answer.

18C.32 HKALE MS 2007 – 4

Albert conducted a survey on the time spent (in hours) on watching television by 16 students. The data recorded are 3.7, 1.2, 2.1, 5.1, 2.1, 4.7, 1.9, 2.4, 2.4, 2.9, 3.6, 2.3, 3.9, 2.2, 1.8 and k, where k is the missing datum.

(a) Albert assumes that the range of these data is 5.3 hours.

(i) Find the value of k.

(ii) Construct a stem-and leaf diagram for these data.

(iii) Find the mean and the median of these data.

(b) Albert finds that the assumption in (a) is incorrect and he can only assume that the range of these data is greater than 5.3 hours. Describe the change in the mean and the change in the median of these data due to the revision of Albert's assumption.

18C.33 HKALE MS 2008 6

A test is taken by a class of 18 students. The marks are as follows:

55	82	74	70	91	75	79	89	68
79	59	72	79	73	60	71	82	k

where k is Jane's mark.

It is known that the mean mark of the class is the same irrespective of including or excluding Jane's.

(a) Find the value of k.

(b) If 3 student marks are selected randomly from the set of the 18 student marks, find the probability that exactly 1 of them is the mode of the set of the 18 student marks.

(c) A student mark is classified as an *outlier* if it lies outside the interval $(\mu - 2\sigma, \mu + 2\sigma)$, where μ is the mean and σ is the standard deviation of the set of marks.

(i) Find all the *outlier*(*s*) of the set of the 18 student marks.

(ii) In order to assess the students' performance in the test, all *outliers* are removed from the set. Describe the change in the median and the standard deviation of the student marks due to such removal.

18C.34 HKALE MS 2011 6

The revision times (in minutes) of 19 students are represented by the stem and leaf diagram in the figure. It is known that the mean revision time is $(40 + b)$ minutes.

(a) Find a and b.

(b) Find the standard deviation of the revision times for the students.

(c) The revision times of 2 more students are added. If both the range and the mean do not change after the inclusion of the 2 data, find the range of possible values of the standard deviation of the revision times for the 21 students.

Tens	Units
2	6 7
3	0 0 a 2 9 9
4	b 3 3 3 6 8 8
5	6 9
6	5 9

18C.35 HKALE MS 2012 – 6 (To continue as **17C.35.**)

An educational psychologist adopts the Internet Addiction Test to measure the students' level of Internet addiction. The scores of a random sample of 30 students are presented in the following stem and-leaf diagram. Let σ be the standard deviation of the scores. It is known that the mean of the scores is 71 and the range of the scores is 56.

(a) Find the values of a, b and σ.

Stem (tens)	Leaf (units)
3	a
4	
5	2 4 6 8
6	0 1 3 5 6 7 8 8 9
7	1 2 2 4 5 5 6 8
8	0 2 3 5 8
9	0 2 b

18C.36 HKDSE MA PP – I – 9

The following table shows the distribution of the numbers of online hours spent by a group of children on a certain day.

Number of online hours	2	3	4	5
Number of children	r	8	12	s

It is given that r and s are positive numbers.

(a) Find the least possible value and the greatest possible value of the inter quartile range of the distribution.

(b) If $r = 9$ and the median of the distribution is 3, how many possible values of s are there? Explain your answer.

18C.37 HKDSE MA PP I – 15

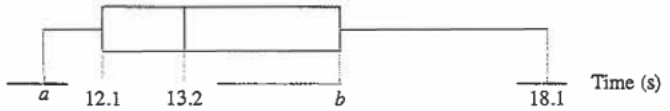
The mean score of a class of students in a test is 48 marks. The scores of Mary and John in the test are 36 marks and 66 marks respectively. The standard score of Mary in the test is -2.

(a) Find the standard score of John in the test.

(b) A student, David, withdraws from the class and his test score is then deleted. It is given that his test score is 48 marks. Will there be any change in the standard score of John due to the deletion of the test score of David? Explain your answer.

18C.38 HKDSE MA 2012 – I – 7

The box and whisker diagram below shows the distribution of the times taken by a large group of students of an athletic club to finish a 100 m race:



The inter quartile range and the range of the distribution are 3.2 s and 6.8 s respectively.

(a) Find a and b.

(b) The students join a training program. It is found that the longest time taken by the students to finish a 100 m race after the training is 2.9 s less than that before the training. The trainer claims that at least 25% of the students show improvement in the time taken to finish a 100 m race after the training. Do you agree? Explain your answer.

18C.39 HKDSE MA 2012 – I – 15

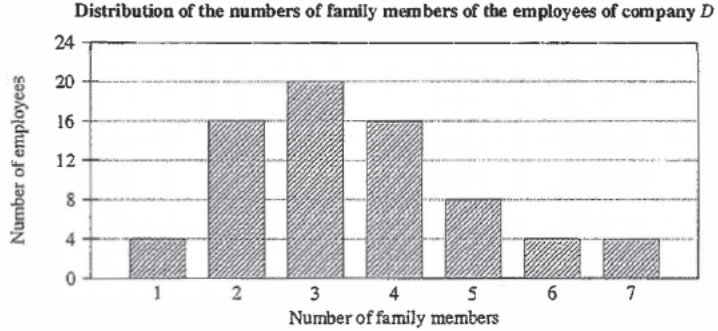
The standard deviation of the test scores obtained by a class of students in a Mathematics test is 10 marks. All the students fail in the test, so the test score of each student is adjusted such that each score is increased by 20% and then extra 5 marks are added.

(a) Find the standard deviation of the test scores after the score adjustment.

(b) Is there any change in the standard score of each student due to the score adjustment? Explain your answer.

18C.40 HKDSE MA 2013 – I – 9

The bar chart shows the distribution of the numbers of family members of the employees of company D.



(a) Find the mean, the inter quartile range and the standard deviation of the above distribution.

(b) An employee leaves company D. The number of family members of this employee is 7. Find the change in the standard deviation of the numbers of family members of the employees of company D due to the leaving of this employee.

18C.41 HKDSE MA 2013 – I – 10 (To continue as 17C.37.)

The ages of the members of Committee A are shown as follows:

17	18	21	21	22	22	23	23	23	31
31	34	35	36	47	47	58	68	69	69

(a) Write down the median and the mode of the ages of the members of Committee A.

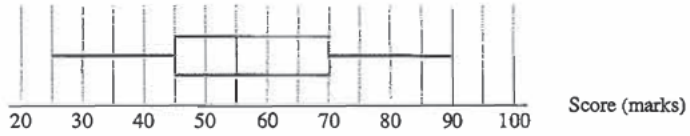
(b) The stem and leaf diagram shows the distribution of the ages of the members of Committee B. It is given that the range of this distribution is 47.

(i) Find a and b.

Stem (tens)	Leaf (units)
2	a 5 6 7
3	3 3 8
4	3
5	1 2 9
6	7 b

18C.42 HKDSE MA 2013 – I – 15

The box and whisker diagram below shows the distribution of the scores (in marks) of the students of a class in a test. Susan gets the highest score while Tom gets 65 marks in the test. The standard scores of Susan and Tom in the test are 3 and 0.5 respectively.



(a) Find the mean of the distribution.

(b) Susan claims that the standard scores of at least half of the students in the test are negative. Do you agree? Explain your answer.

18C.43 HKDSE MA 2014 – I – 4

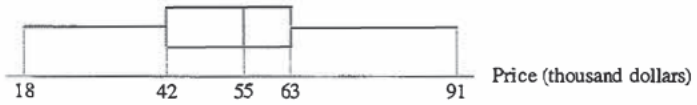
The table below shows the distribution of the numbers of calculators owned by some students.

Number of calculators	0	1	2	3
Number of students	7	14	15	4

Find the median, the mode and the standard deviation of the above distribution.

18C.44 HKDSE MA 2014 I – 11

There are 33 paintings in an art gallery. The box-and whisker diagram below shows the distribution of the prices (in thousand dollars) of the paintings in the art gallery. It is given that the mean of this distribution is 53 thousand dollars.



(a) Find the range and the inter quartile range of the above distribution.

(b) Four paintings of respective prices (in thousand dollars) 32, 34, 58 and 59 are now donated to a museum. Find the mean and the median of the prices of the remaining paintings in the art gallery.

18C.45 HKDSE MA 2015 I 12

The stem-and-leaf diagram shows the distribution of the weights (in kg) of the students in a football club.

Stem (tens)	Leaf (units)
4	0 2 3 3 3 3 9
5	1 1 2 2 3 7 9
6	3 5 8 9
7	8 9

(a) Find the mean, the median and the range of the above distribution.

(b) Two more students now join the club. It is found that both the mean and the range of the distribution of the weights are increased by 1 kg. Find the weight of each of these students.

18C.46 HKDSE MA 2015 I – 15

The table below shows the means and the standard deviations of the scores of a large group of students in a Mathematics examinations and a Science examination:

Examination	Mean	Standard deviation
Mathematics	66 marks	12 marks
Science	52 marks	10 marks

The standard score of David in the Mathematics examination is -0.5.

(a) Find the score of David in the Mathematics examination.

(b) Assume that the scores in each of the above examinations are normally distributed. David gets 49 marks in the Science examination. He claims that relative to other students, he performs better in the Science examination than in the Mathematics examination. Is the claim correct? Explain your answer.

18C.47 HKDSE MA 2016 – I – 16

In a test, the mean of the distribution of the scores of a class of students is 61 marks. The standard scores of Albert and Mary are -2.6 and 1.4 respectively. Albert gets 22 marks. A student claims that the range of the distribution is at most 59 marks. Is the claim correct? Explain your answer.

18C.48 HKDSE MA 2017 I 11 (To continue as **17B.43.**)

The stem-and-leaf diagram shows the distribution of the hourly wages (in dollars) of the workers in a group.

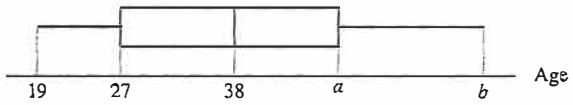
It is given that the mean and the range of the distribution are \$70 and \$22 respectively.

Stem (tens)	Leaf (units)
6	1 1 1 3 4 6 8 9 9
7	a 7 7 8
8	1 b

(a) Find the median and the standard deviation of the above distribution.

18C.49 HKDSE MA 2018 – I – 10

The box-and whisker diagram below shows the distribution of the ages of the clerks in team X of a company. It is given that the range and the inter-quartile range of this distribution are 43 and 21 respectively.



(a) Find a and b.

(b) There are five clerks in team Y of the company and three of them are of age 38. It is given that the range of the ages of the clerks in team Y is 20. Team X and team Y are now combined to form a section. The manager of the company claims that the range of the ages of the clerks in the section and the range of the ages of the clerks in team X must be the same. Do you agree? Explain your answer.

18C.50 HKDSE MA 2019 – I 12

The stem-and leaf diagram shows the distribution of the results (in seconds) of some boys in a 400 m race.

It is given that the inter-quartile range of the distribution is 8 seconds.

Stem (tens)	Leaf (units)
5	a
6	0 0 3 c c 8 9 9 9
7	0 1 1 1 2 2 5 6 9
8	b

(a) Find c.

(b) It is given that the range of the distribution exceeds 34 seconds and the mean of the distribution is 69 seconds. Find

(i) a and b,

(ii) the least possible standard deviation of the distribution.

18C.51 HKDSE MA 2020 – I – 9

The table below shows the distribution of the numbers of subjects taken by a class of students.

Number of subjects taken	4	5	6	7
Number of students	8	12	16	4

(a) Write down the mean, the median and the standard deviation of the above distribution.

(b) A new student now joins the class. The number of subjects taken by the new student is 5. Find the change in the median of the distribution due to the joining of this student.

(5 marks)

18C.52 HKDSE MA 2020 – I – 11

The stem-and-leaf diagram below shows the distribution of the weights (in grams) of the letters in a bag.

Stem (tens)	Leaf (units)
1	1 2 3 3
2	3 3 4 5 6 9 9
3	1 6 7 8 8 8
4	2
5	0 w

It is given that the range of the above distribution is the triple of its inter-quartile range.

(a) Find w. (4 marks)

(b) If a letter is randomly chosen from the bag, find the probability that the weight of the chosen letter is not less than the mode of the distribution. (2 marks)

18 Statistics

18A Presentation of data

18A.1 HKCEE MA 1982(1) – I – 7

(a) $x = 360 \times \frac{2}{2+7+3} = 60$
Similarly, $y = 210$, $z = 90$

(b) Total no. of students $= 240 \times \frac{2+7+3}{2} = 1440$

18A.2 HKCEE MA 1982(3) – I – 12

(a) Income $= \$2000 \times \frac{360°}{100°} = \$7200$

(b) (i) (1) $x = 100(1+30\%) = 130$
$y = 50$
$z = 360 - 90(1+10\%) - 130$
$20(1+30\%) - 50 - 40 = 15$

(2) % change $= \frac{15-60}{60} \times 100\% = -75\%$

(ii) Income in April $= \$7200 \times (1+37.5\%) = \$9900$
Expense on food $= \$7200 \times \frac{90°}{360°} \times (1+10\%)$
$= \$1980$
$\therefore$ Required % $= \frac{1980}{9900} \times 100\% = 20\%$

18A.3 HKCEE MA 1985(A/B) – I – 7

$\angle$ of sector representing Kowloon $= 90° \times \frac{9240}{4200} = 198°$
$\therefore x° = 360° - 90° - 198° \Rightarrow x = 72$
$n = 4200 \times \frac{72°}{90°} = 3360$

18A.4 HKCEE MA 1998 – I – 10

(a)

$x \le 70$	102	$60 < x \le 70$	52
$x \le 80$	158	$70 < x \le 80$	56

(b) The line $x = 55$ meets the c.f. polygon at around (55, 29).
$\therefore$ Passing % $= \frac{200-29}{200} \times 100\% = 85.5\%$

18A.5 HKCEE MA 1999 – I – 11

(a) $\angle$ of sector representing 'Repeated S.5' $= 72°$
$\therefore$ No of boys who repeated S.5 $= 120 \times \frac{72°}{360°} = 24$

(b) Required % $= \frac{126°}{126° + 144°} \times 100\% = 46\frac{2}{3}\%$

(c) No of boys promoted to S.6 in own school $= 120 \times \frac{126°}{360°}$
$= 42$
No of girls promoted to S.6 in own school $= 80 \times 22.5\%$
$= 18$
$\therefore$ Required % $= \frac{42+18}{200} \times 100\% = 30\%$

18A.6 HKCEE MA 2006 – I – 9

(a) $x = 360° - 40° - 90° - 130° - 35° - 30° = 35°$

(b) Total expenditure $= \$1750 \times \frac{360°}{35°} = \$18000$

(c) Expenditure on travelling = Expenditure on transportation
$= \$1750$

18A.7 HKCEE MA 2007 – I – 12

(a) $k = 17 \times \frac{63°}{153°} = 7$

(b) No of students $= 17 \times \frac{360°}{153°} = 40$

(c) No of students with 1 key $= 40 - 12 - 17 - 7 = 4$
$\therefore$ Required $p = \frac{4}{40} = \frac{1}{10}$

(d) Bar: Yes. Scales on the vertical axis should be doubled.
Pie: No

18A.8 HKDSE MA SP – 1 – 9

(a) $72 = (1+20\%)x \Rightarrow x = 60$

(b) $\angle$ of sector representing District C $= 78° > 72°$
$\therefore$ NO.

18A.9 HKDSE MA PP – I – 13

(a) Number of students $= 6 \div \frac{3}{20} = 40$
$\Rightarrow k = 40 - 6 - 11 - 5 - 10 = 8$

(b) (i) Required $\angle = 360° \times \frac{5}{40} = 45°$

(ii) Suppose n new students will double the $\angle$ for orange.
$\frac{5+n}{40+n} = \frac{45° \times 2}{360°} \Rightarrow n = \frac{20}{3}$
But since n must be an integer, there is no n satisfying the condition. NO.

18A.10 HKDSE MA 2016 – I – 9

(a) $x = 2+4 = 6$
$y = 37 - 15 = 22$
$z = 37 + 3 = 40$

18B Measures of central tendency

18B.1 HKCEE MA 1983(B) – I – 3

(a) Class mark $= 54.5$

(b) Mean $= (44.5 \times 100 + 54.5 \times 300 + 64.5 \times 400$
$+ 74.5 \times 200) \div 1000 = 61.5$

18B.2 HKCEE MA 1984(A/B) – I – 2

$1 \times 10 + 2 \times 10 + 3 \times 5 + 4 \times 20 + 5x = 3(10 + 10 + 5 + 20 + x)$
$125 + 5x = 135 + 3x \Rightarrow x = 5$

18B.3 HKCEE MA 1986(A/B) – I – 3

$61 \times 40 + 70x + 50 \times 35 = 60(40 + x + 35)$
$4190 + 70x = 4500 + 60x \Rightarrow x = 31$

18B.4 HKCEE MA 1991 – I – 1

(a) 49.5

(b)

20 – 29	10
30 – 39	10
40 – 49	20
50 – 59	30
60 – 69	10

$\therefore$ Mean mark $= (24.5 \times 10 + 34.5 \times 10 + 44.4 \times 20$
$+ 54.5 \times 30 + 64.5 \times 10) \div 80 = 47$

18B.5 HKCEE MA 1992 – I – 8

(a) $70(m+n)$

(b) $70(m+n) = 75m + 62n \Rightarrow 5m = 8n \Rightarrow m : n = 8 : 5$

(c) No of men $= 39 \times \frac{8}{8+5} = 24$

18B.6 HKCEE MA 1994 – I – 1(d)

Mean $= 50$ Mode $= 65$ Median $= 60$

18B.7 HKCEE MA 1996 – I – 14

(a) 100 $20.0 - 4.9 - 13.7 - 26.3 - 10.3 - 1.1 = 23.7$

(b) Some round-off errors have accumulated.

(c) (i)

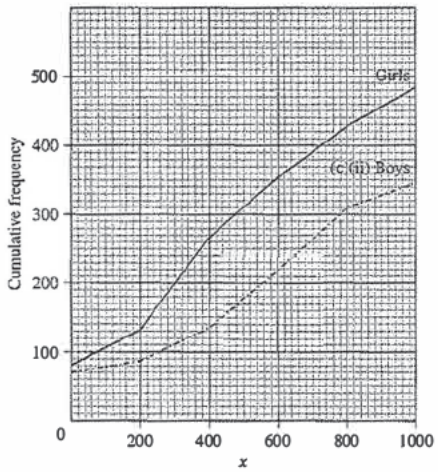
x	c.f.
$x \le 0$	70
$x \le 200$	87
$x \le 400$	135
$x \le 600$	218
$x \le 800$	310
$x \le 1000$	346

(ii) (See below)

(iii) For boys, median $= 490$; for girls, median $= 410$

(iv) Draw the vertical line $x = 700$. It meets the polygons at around (700, 390) (girls) and (700, 265) (boys).
$\therefore$ Required no $= 390 + 265 = 655$

(d) Referring to the first row of each table, the percentage of boys spending \$0 (20.0%) is indeed higher than the percentage of girls spending \$0 (15.0%).
However, the percentages have to be considered instead of the frequencies because the *total frequencies* of boys and of girls are *different*.



18B.8 HKCEE MA 1999 – I – 8

(a) $x + 161 + 168 + 159 + 161 + 152 = 158 \times 6 \Rightarrow x = 147$

(b) Median $= (159 + 161) \div 2 = 160$ (cm)

18B.9 HKCEE MA 2000 – I – 11

(a)

$t \le 280$	66	$260 < t \le 280$	20

(b) Mean $= (210 \times 3 + 230 \times 13 + 250 \times 30$
$+ 270 \times 20 + 290 \times 9) \div 75 = 255$ (s, 3 s.f.)

(c) Median $= 254$ seconds

(d) Required % $= \frac{13+30}{75} \times 100\% = 57.3\%$ (3 s.f.)

18B.10 HKCEE MA 2003 – I – 11

(a) (i) 10
(ii) 11.5
(iii) 12
(iv) $16 - 10 = 6$

(b) (i) (When all 4 new data are large,)
Least possible median $= (13 + 16) \div 2 = 14.5$
(When all 4 new data are small,)
Greatest possible median $= 10$

(ii) New mean $= (12 \times 6 + 11 \times 4) \div 10 = 11.6$

18B.11 HKCEE MA 2006 – I – 8

(a) $11 \times 10 = 86 + k \Rightarrow k = 24$

18B.12 HKALE MS 1998 – 3

(a) Median $= (161 + 162) \div 2 = 161.5$ (cm)

18B.13 HKALE MS 2002 – 7

(a) Mean $= 61$

(b) Since there are two modes, one deleted mark is 54.
The other mark $= 61 \times 22 - (61 + 1.2) \times 20 - 54 = 44$

18B.14 HKALE MS 2010 – 5

(a) $49 - (20+a) = 27 \Rightarrow a = 2$

$$\frac{49+\cdots+(80+b)}{20} - \frac{22+49+\cdots+(80+b)}{21} = 2$$

$$\frac{1274+b}{20} - \frac{1296+b}{21} \quad 2$$

$b = 6$

$\bar{x} = (1296+6) \div 21 = 62$

18B.15 HKDSE MA SP – I – 14

(a) Median = 62%
Mean = $(55+58+62+62+63) \div 5 = 60(\%)$

(b) (i) 58% (when the new data are small)
(ii) (Mean unchanged $\Rightarrow$ Mean of a and $b = 60$)
(Median unchanged $\Rightarrow$ $a \le 62$ and $b \ge 62$)
Possible pairs: (57, 63), (56, 64), (55, 65), etc.

(c) Possible reasons for NO:
- The week may not be randomly chosen.
- Only one stall is considered.

Possible reasons for YES:
- The week may be randomly chosen.
- There may only be very few stalls in H.

18B.16 HKDSE MA 2012 – I – 10

(a) Mean = $10+10+\cdots+36) \div 20 = 18$
Median = 16

(b) (i) New mean = Original mean = 18
(ii) Let the new data be 19, 20, a and b.
Mean = 18 $\Rightarrow$ $a+b = 18 \times 4$ 19 20 = 33
Since 19 and 20 exceed the original median, a and b must not exceed the original median if the median is unchanged. $\Rightarrow$ $a+b \le 16+16 = 32$
Hence it is not possible.

18B.17 HKDSE MA 2016 – I – 12

(a) Median = 7.5 $\Rightarrow$ No of 6 and 7 = No of 8, 9 and 10

$11+a = 11+b+4$

$a = b+4$

$\because a > 11$ and $4 < b < 10$

$\therefore (a, b)$ (12, 8) or (13, 9)

(b) (i) Greatest possible median = 8
(when the 4 new ages are 7, 8, 9 and 10)
(ii) Mean is least when the 4 new ages are 6, 7, 8 and 9.
If $(a, b) = (12, 8)$, mean = $(6 \times 12 + 7 \times 13 + 8 \times 12 + 9 \times 9 + 10 \times 4) \div (12 + 13 + 12 + 9 + 4) = 7.6$
If $(a, b) = (13, 9)$, mean = $(6 \times 12 + 7 \times 14 + 8 \times 12 + 9 \times 10 + 10 \times 4) \div (12 + 14 + 12 + 10 + 4) = 7.62$
$\therefore$ Least possible mean = 7.6

18B.18 HKDSE MA 2018 – I – 11

(a) (i) 1
(ii) 8

(b) (i) 3 (when the '9th 2' is the median)
(ii) 19 (when the '1st 2' is the median)

(c) $(0 \times k + 1 \times 2 + 2 \times 9 + 3 \times 6 + 4 \times 7) \div (k+2+9+6+7) = 2 \Rightarrow k = 9$

18B.19 HKDSE MA 2019 – I – 8

(a) 2

(b) Mean = $2 \times \frac{144^\circ}{360^\circ} + 3 \times \frac{54^\circ}{360^\circ} + 5 \times \frac{72^\circ}{360^\circ} + 7 \times \frac{90^\circ}{360^\circ} = 4$

18C Measures of Dispersion

18C.1 HKCEE MA 1980(3) – I – 8

(a) Class B (since its dispersion is greater)

(b) 10 students fail the test.
$\Rightarrow$ Students getting 0, 1 and 2 marks fail the test.
$\Rightarrow$ Min mark to pass test = 3

18C.2 HKCEE MA 1981(1) – I – 6

(a) The line $y = 25$ meets the polygon at around (43, 25).
$\therefore$ Pass mark = 43

(b) The line $x = 40$ meets the polygon at around (40, 20).
$\therefore$ 100 – 20 = 80 students would pass.

(c) IQR = 70 – 43 = 27

18C.3 HKCEE MA 1983(A) – I – 3

(a) Mean = $[(a \quad 6) + a + (a+2) + (a+3) + (a+6)] \div 5$
$= a+1$

(b) SD = SD of $\{ \quad 6, 0, 2, 3, 6\} = 4$

18C.4 HKCEE MA 1988 – I – 11

(a) (i) Median = 70 marks
(ii) IQR = 86 50 = 36 (marks)

(b) (i) Number of students = 600 – 540 = 60
(ii) Required $p = \frac{60}{600} = \frac{1}{10}$
(iii) (1) Required $p = \frac{C_2^{60}}{C_2^{600}} = \frac{59}{5990}$
(2) Required $p = 1 - \frac{C_2^{540}}{C_2^{600}} = \frac{1139}{5990}$

18C.5 HKCEE MA 1990 – I – 12

(a) (i) Modal class = \$6000 \$7000
Median = \$6500
Mean = \$6500 (since the distribution is symmetric)
IQR = $Q_3 - Q_1 = \frac{7500+8500}{2} - \frac{4500+5500}{2}$
= (\$)3000

(ii) $\because$ More data are close to the mean
$\therefore$ SD becomes smaller.

(b) SD = $\sqrt{\frac{9+4+1+0+1+4+9}{7}} = 2$

18C.6 HKCEE MA 1993 – I – 7

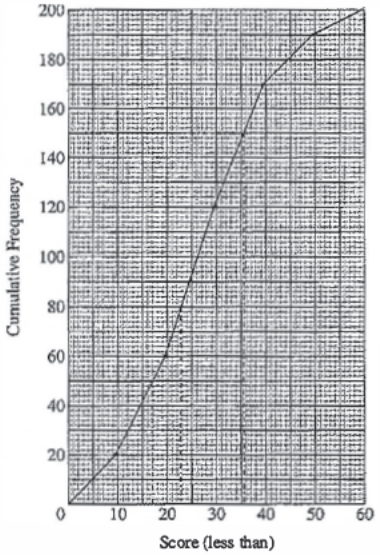
(a)

9.5	**20**
19.5	**60**
29.5	**120**
39.5	**170**
49.5	**190**
59.5	**200**

(b) (i) (See below)
Hence, $IQR = 36 - 17 = 19$ (or 35 17 = 18)
(ii) $\because 200 \times 60\% = 80$ students pass the test.
$\therefore$ The passing score should be 23.

(c) SD = 12.9

(d) SD = 12.9 (i.e. unchanged)



18C.7 HKCEE MA 1995 – I – 9

(a) (i) 180
(ii) 60

(b) (25% of students = 45)
(i) The horizontal line at 135 meets the graph at around (75, 135).
$\therefore$ 75
(ii) The horizontal line at 45 meets the graph at around (44, 45).
$\therefore$ 44

(c)

$20 < x \le 30$	25	**12**
$30 < x \le 40$	35	20
$40 < x \le 50$	45	**28**
$50 < x \le 60$	55	32
$60 < x \le 70$	65	**28**
$70 < x \le 80$	75	30
$80 < x \le 90$	85	22
$90 < x \le 100$	95	**8**

(d) Mean = 59.6, SD = 19.0 (3 s.f.)

(e) $\bar{x} \quad \sigma = 40.6$, $\bar{x} + \sigma = 78.6$
No. of students within this range = 146 34 = 112
$\therefore$ Required % = $\frac{112}{180} \times 100\% = 62.2\%$

18C.8 HKCEE MA 1997 – I – 11

(a) Mean = 64.4, Mode = 95, Median = 78, SD = 30.6

(b) There are several extremely small data.

(c) (i) Required mark = $63 + 0.4 \times 15 = 69$
(ii) (1) Required % = $\frac{17}{35} \times 100\% = 48.6\%$ (3 s.f.)
(2) *Method 1* – Standard score
S.S. in Maths = $\frac{78-64.4}{30.6} = 0.44$
S.S. in Eng = $\frac{78-63}{15} = 1 > 0.44$
$\therefore$ Performance in Eng was better.
Method 2 – Use distribution
In Maths, her score was the median. Thus, not more than half of the classmates perform worse than her.
In Eng, her score was above the mean. Thus, more than half of the classmates perform worse than her.
$\therefore$ Performance in Eng was better.

(iii) New mean = $(63 \times 35 + 10) \div 35 = 63.3$

18C.9 HKCEE MA 2001 – I – 10

(a)

Score (x)	Class mid-value (Class mark)	Frequency
$44 \le x < 52$	**48**	3
$52 \le x < 60$	**56**	**9**
$60 \le x < 68$	64	15
$68 \le x < 76$	**72**	11
$76 \le x < 84$	80	**2**

(b) Mean = 64
SD = 8

(c) S.S. = (76 64) ÷ 8 = 1.5

(d) Let x be her score in the second test.
$1.5 = \frac{x-58}{10} \Rightarrow x = 73$
The required score is 73.

18C.10 HKCEE MA 2002 – I – 5

(a) 9.6
(b) 13
(c) 10
(d) 4.59

18C.11 HKCEE MA 2002 – I – 12

(a)

$0 < x \le 5$	66	Certificate
$5 < x \le 15$	**34**	Book coupon
$15 < x \le 25$	64	Bronze medal
$25 < x \le 35$	**26**	Silver medal
$35 < x \le 50$	10	Gold medal

(b) IQR = 23 – 4 = 19

18C.12 HKCEE MA 2004 – I – 11

(a) S.S. in Paper I = $\frac{54-46.1}{15.2} = 0.520$

S.S. in Paper II = $\frac{66 \quad 60.3}{11.6} = 0.491 <$ S.S. in Paper I

$\therefore$ NO.

(b) New mean = 50.1 marks
New median = 50 marks
New range = 91 marks

18C.13 HKCEE MA 2005 – I – 15

(a) Mean = 122 marks, SD = 22 marks

(b) Top 20% = 4 students
Mary's score = 122 + 22 = 144 marks, which is not within the top 4 students.
∴ NO.

(c) (i) (Mean unchanged ⇒ Datum deleted is 122.)
Required $p = \frac{1}{20}$

(ii) (Mean unchanged ⇒ Sum of data deleted is 122×2)
Required $p = \frac{2}{C_2^{20}} = \frac{1}{95}$

18C.14 HKCEE MA 2006 – I – 14

(a) (i) Class A: IQR = 39 − 18 = 21 (marks)
Class B: IQR = 25 − 11 = 14 (marks)

(ii) ∵ IQR of B < IQR of A
∴ Class B is less dispersed.

18C.15 HKCEE MA 2007 – I – 4

Median = 67 kg
Range = 25 kg
SD = 7.65 kg

18C.16 HKCEE MA 2008 – I – 10

(a) $a = 4$
$b = 37 \quad 12 - 4 = 21$
$c = 50 - a - 12 \quad b - 10 = 3$

(b) Mean = 3.28 kg
SD = 0.299 kg

18C.17 HKCEE MA 2008 – I – 14

(a) (i) Required $p = \frac{9}{15} = \frac{3}{5}$

(ii) (1) Required $p = \frac{8 \times 15}{C_2^{36}} = \frac{4}{21}$

(2) Required $p = 1 - \frac{C_3^8}{C_3^{36}} - \frac{C_3^{15}}{C_3^{36}} - \frac{C_2^{13}}{C_2^{36}} = \frac{419}{630}$

(b) (i) Median = 5000 dollars
IQR = 6400 4300 = 2100 (dollars)

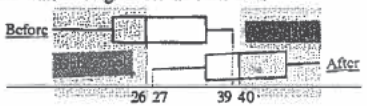
(ii) Extra \$1000 to each salesgirl

18C.18 HKCEE MA 2009 – I – 10

(a) Median = 26 wpm
Range = 27 wpm
IQR = 35 − 21 = 14 (wpm)

(b) (i) *Method 1*
Range after training = 25 wpm < 27 wpm ⇒ NO
Method 2
IQR after training = 12 wpm < 14 wpm ⇒ NO

(ii) *Method 1*
Before the training, no speed was higher than 39 wpm. After the training, at least half of the speeds are 40 wpm or above. ⇒ YES.
Method 2
Before the training, at least half of the speeds were 26 wpm or below. After the training, their speeds become at least 27 wpm. ⇒ YES.
Remarks
To look for arguments against these claims, it is often helpful to provide yourself with a sketch of the box-and-whisker diagram for the other data.



18C.19 HKCEE MA 2010 – I – 11

(a) Mean = 25
Median = 26
Range = 13

(b) (i) Let x be the mean age of the 3 new players.
$(55 \times 22 \quad 31 \quad 31 + 3x) \div 23 = 25 \Rightarrow x = 29$.
∴ The required mean is 29.

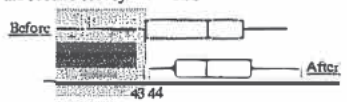
(ii) Median unchanged: If one new player is younger than the median, the other two has to be older – then the median will be the 12th datum; if two younger and one older than the mean, the median will be the 11th datum instead (which is still 26).
Range unchanged: New ages within 18 to 31
Possible ages: {25, 31, 31}, {26, 30, 31}, {27, 29, 31}, {28, 28, 31}

18C.20 HKCEE MA 2011 – I – 10

(a) Median = 57
Range = 75 23 = 52
IQR = 65 − 43 = 22

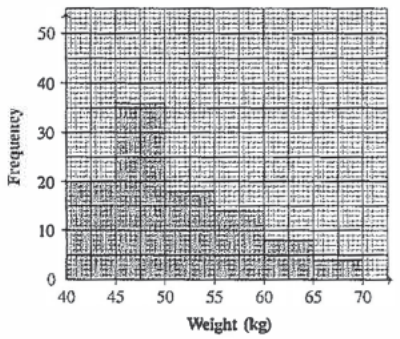
(b) (i) *Method 1*
IQR in 2nd survey = 67 − 50 = 17 < 22 ⇒ YES.
Method 2
Range in 2nd survey = 78 − 44 = 34 < 52 ⇒ YES.

(ii) At least 25% of the scores were 43 or below in the first survey. They have all improved to at least 44 in the second survey. ⇒ YES



18C.21 HKALE MS 1994 – 4

(a)



(b) IQR = 55 46 = 9 (kg)

(c) Mean $= \frac{42.5 \times 20 + \cdots + 67.5 \times 4}{100} = 50.8$ (kg)

18C.22 HKALE MS 1995 – 1

(a)

Stem (in 10)	Leaf (in 1)
0	5 7 8
1	0 1 2 2 5 8 8 8 8 9
2	0 1 2 3 5 5 5 6 9
3	0 2
4	
5	0

(b) Mode = 18
Median = 19
IQR = 25 − 12 = 13

18C.23 HKALE MS 1996 – 1

(a) Mean = 59.4
Mode = 74
IQR = 72 − 50 = 22

(b) Mean becomes 57.4.
IQR becomes 72 − 49 = 23

18C.24 HKALE MS 1997 – 2

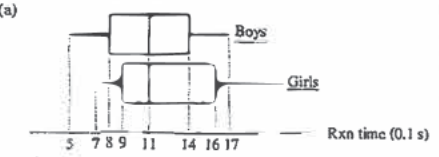
(a) 26 − 18 = 8 (°C)

(b) (i) Median = 21.5 °C = 70.7 °F
$\text{IQR} = \left[\frac{9}{5}(22.5) + 32\right] \quad \left[\frac{9}{5}(20) + 32\right] = 4.5$ (°F)

(ii) Mean $= \frac{9}{5}(22) + 32 = 71.6$ (°F)
$\text{SD} = \frac{9}{5}(2) = 3.6$ (°F)

18C.25 HKALE MS 1999 – 3

(a)

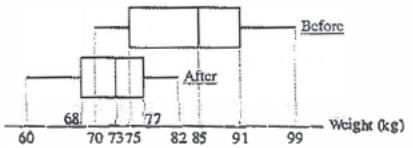


(b) Equal chance since both of the probabilities will be 0.5

18C.26 HKALE MS 2000 – 5

(a) Median = 85 kg
IQR = 91 − 75 = 16 (kg)

(b)

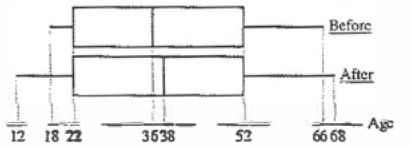


(c) No conclusion can be drawn as the diagrams show no individual difference.

18C.27 HKALE MS 2001 – 3

(a) $48 = 66 \quad (10 + a) \Rightarrow a = 8$
$30 + b = 36 \Rightarrow b = 6$
$c = 6 \quad 1 = 5$

(b)



18C.28 HKALE MS 2003 – 5

(a)

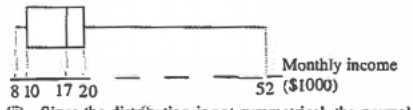
Stem (10 mins)	Leaf (1 min)
1	2 3 4 5 5
2	1 5 9
3	6 7 8
4	1 7 9 9 9
5	2 4 7

(b) (i) Revised median = 49 mins
Revised IQR = 60 − 25 = 35 (mins)

(ii) Both will become larger.

18C.29 HKALE MS 2004 – 5

(a)



(b) (ii) Since the distribution is not symmetrical, the normal distribution is not an appropriate model.

18C.30 HKALE MS 2005 – 4

(a) Change in mean = Change in sum ÷ 32
$= (3 \times 10) \div 32 = 0.9375$ (cm)

(b) Median unchanged

(c) Mode unchanged

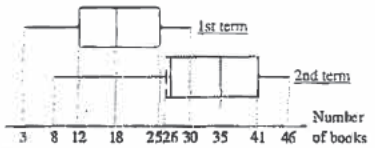
(d) Case 1: The 3 data were 145, 145 and 146.
Change in range = −1
Case 2: The 3 data were 145, 146 and 146.
Range unchanged

(e) Original IQR = 168 − 154 = 14
New IQR = 168 − 155 13
⇒ Change = −1

18C.31 HKALE MS 2006 – 4

(a) Median = 18
IQR = 25 – 12 = 13

(b) (i)



(ii) In 1st term, the maximum number was 30. In 2nd term, at least half of the numbers are 35 or above. Hence, at least 50% of students have read at least (35 – 30 =) 5 more books. Agreed.

18C.32 HKALE MS 2007 – 4

(a) (i) (k is the largest datum since 5.1 cannot be.)
$k = 1.2 + 5.3 = 6.5$

(ii)

Stem (1 hour)	Leaf (0.1 hour)
1	2 8 9
2	1 1 2 3 4 4 9
3	6 7 9
4	7
5	1
6	5

(iii) Mean = 3.05 hours
Median = 2.4 hours

(b) Mean will become larger.
Median will be unchanged.

18C.33 HKALE MS 2008 – 6

(a) k = Mean of the other 17 students = 74

(b) Required $p = \dfrac{C_1^3 C_2^{15}}{C_3^{18}} = \dfrac{105}{272}$

(c) (i) $SD = \sigma = 9.327$
Hence the interval is $(74 - 2\sigma + 2\sigma) = (55.3, 92.7)$
∴ 55 is the only outlier.

(ii) Median unchanged. SD will decrease.

18C.34 HKALE MS 2011 – 6

(a) $(40+b)19 = 743 + (30+a) + (40+b) \Rightarrow a = 18b - 53$
Since a and b are integers, $0 \le a \le 2$ and $0 \le b \le 3$,
$b = 3 \Rightarrow a = 1$

(b) 12.2 minutes

(c) Range unchanged: New data within 26 and 69
Mean unchanged: New data are $(43 - x, 43 + x)$
SD is smallest when both new data are 43.
⇒ Least possible SD = 11.6 mins
SD is greatest when the data are 26 and 60.
⇒ Greatest possible SD = 12.7 mins

18C.35 HKALE MS 2012 – 6

(a) $\dfrac{(30+a) + 52 + \cdots + 92 + (90+b)}{30} = 71$

$2120 + a + b = 2130$

$a + b = 10$

$(90+b) - (30+a) = 56 \Rightarrow a - b = 4$

Solving, $a = 7, b = 3$

$\Rightarrow \sigma = 12.7$

18C.36 HKDSE MA PP – I – 9

(a) Least possible IQR = 0
(when there are many many 2's or many many 5's)
Greatest possible IQR = 5 – 2 = 3

(b) $9 + 8 > 12 + s \Rightarrow s < 5$
∴ s = 1, 2, 3 or 4; i.e. 4 possible values of s

18C.37 HKDSE MA PP – I – 15

(a) $SD = (36 - 48) \div (-2) = 6$
∴ S.S. of John $= \dfrac{66 - 48}{6} = 3$

(b) Mean unchanged
SD increases (since 'more' data are 'far away' from mean)
∴ YES (decrease)

18C.38 HKDSE MA 2012 – I – 7

(a) $a = 18.1 - 6.8 = 11.3$
$b = 12.1 + 3.2 = 15.3$

(b) New longest time = 18.1 – 2.9 = 15.2 (s)
Before the program, at least 25% of students take 15.3 s or longer. After the program, they have shortened their time by at least 0.1 s. ⇒ YES.

18C.39 HKDSE MA 2012 – I – 15

(a) New SD = 10 × (1 + 20%) = 12

(b) Upon adjustment, the deviation of each score from the mean is increased by 20% while the SD is also increased by 20%. By the formula S.S. = $\dfrac{\text{Deviation}}{\text{SD}}$, there is no change in the standard score for each score

18C.40 HKDSE MA 2013 – I – 9

(a) Mean $= \dfrac{1 \times 4 + 2 \times 16 + \cdots + 7 \times 4}{4 + 16 + \cdots + 4} = 3.5$
IQR = 4 – 2 = 2
SD = 1.5

(b) New SD = 1.451456
∴ Change = 1.451456 – 1.5 = 0.0485

18C.41 HKDSE MA 2013 – I – 10

(a) Median = 31
Mode = 23

(b) (i) $(60+b) - (20+a) = 47 \Rightarrow b - a = 7$
∵ $0 \le a \le 5$ and $7 \le b \le 9$
∴ $(a, b) = (0, 7), (1, 8)$ or $(2, 9)$

(ii) Required $p = \dfrac{3+3+3+3+2+9+9}{20 \times 13} = \dfrac{8}{65}$

18C.42 HKDSE MA 2013 – I – 15

(a) Let $\bar{x}$ and σ be the mean and SD.

$$\begin{cases} 90 = \bar{x} + 3\sigma \\ 65 = \bar{x} + 0.5\sigma \end{cases} \Rightarrow \begin{cases} \bar{x} = 60 \\ \sigma = 10 \end{cases}$$

(b) Scores below the mean have negative standard scores. From the box-and-whisker diagram, at least half of students scored 55 or below. Hence they must have negative standards scores. ⇒ YES.

18C.43 HKDSE MA 2014 – I – 4

Median = 1
Mode = 2
SD = 0.889

18C.44 HKDSE MA 2014 – I – 11

(a) Range = 91 – 18 = 73 (000 dollars)
IQR = 63 – 42 = 21 (000 dollars)

(b) New mean = $(53 \times 33 - 32 - 34 - 58 - 59) \div 29$
= 54 (000 dollars)
New median = original median = 55 (000 dollars)

18C.45 HKDSE MA 2015 – I – 12

(a) Mean = 55 kg
Median = 52 kg
Range = 79 – 40 = 39 (kg)

(b) Let the new weights be a and b (kg).
$a + b + 55 \times 20 = 56 \times 22 \Rightarrow a + b = 132$
Since the range is increased by only 1,
If $a = 39$, then $b = 132 - 39 = 93$ (rejected)
If $b = 80$, then $a = 132 - 80 = 52$
Hence the only possibility is 52 kg and 80 kg.

18C.46 HKDSE MA 2015 – I – 15

(a) Score of David = 66 – 0.5(12) = 60

(b) S.S. in Science = $\dfrac{49 - 52}{10} = -0.3 >$ S.S. in Maths
∴ YES

18C.47 HKDSE MA 2016 – I – 16

SD = (22 – 61) ÷ (2.6) = 15
⇒ Score of Mary = 61 + 1.4(15) = 82
∵ Range ≥ 82 22 = 60
∴ The claim is wrong.

18C.48 HKDSE MA 2017 – I – 11

(a) $(80+b) - 61 = 22 \Rightarrow b = 3$
$\dfrac{61 + \cdots + (70+a) + \cdots + 83}{15} = 70 \Rightarrow a = 2$
∴ Median = \$69, SD = \$7.33

(b) Required $p = \dfrac{6}{15} = \dfrac{2}{5}$

18C.49 HKDSE MA 2018 – I – 10

(a) $a = 27 + 21 = 48$
$b = 19 + 43 = 62$

(b) Least possible age in Team Y = 38 – 20 = 18
Since 18 < 19, the range of the new section would be larger than that of Team X. Disagreed.

18C.50 HKDSE MA 2019 – I – 12

(a) IQR = 72 – (60 + c) = 8 ⇒ c = 4

(b) (i) $(80+b)$ $(50+a) > 34 \Rightarrow b - a > 4$
$(50+a) + 60 + 60 + \cdots + 79 + (80+b) = 69 \times 20$
$\Rightarrow a + b = 7$
∴ (a, b) (0, 7) or (1, 6)

(ii) SD is smaller when the data are less dispersed.
∴ Least possible SD occurs when $(a, b) = (1, 6)$
By the calculator, Least possible SD = 7.34 (3 s.f.)

18C.51 HKDSE MA 2020 – I – 9

9a The mean is 5.4.
The median is 5.5.
The standard deviation is 0.917 (corr. to 3 sig. fig.).

b The new number of students 8+12+16+4+1
= 41
Therefore, the median is the 21st smallest number of subjects taken.
Hence, the new median is 5.
The change in the median of the distribution 5 – 5.5
= –0.5

18C.52 HKDSE MA 2020 – I – 11

11a ~~The inter-quartile range = $\dfrac{38+38}{2} - \dfrac{23+23}{2}$~~
= 15
Since the range of the distribution is the triple of its inter-quartile range.
$(50 + w) - 11 = 15 \times 3$
$w = 6$

b The mode of the distribution is 38 g.
The required probability $\dfrac{6}{20}$
$\dfrac{3}{10}$