

Data Representation and Interpretation

By studying this lesson you will be able to,

- represent data in a stem and leaf diagram,
- find the maximum value, minimum value and the range of a collection of data • using a stem and leaf diagram, and
- find the mode, median, mean and range of a collection of raw data. •

27.1 Stem and leaf diagram

In Grades 6 and 7, you learnt to represent and interpret data using picture graphs, bar graphs and multi bar graphs. Now we will consider what a stem and leaf diagram is and how data is represented in a stem and leaf diagram.

A stem and leaf diagram is a standard method of organizing numerical data to enable us to interpret the data easily.

When data is organized according to this method,

- if the values of the data are from 0 to 99, the value in the units place of a datum is indicated as the leaf and the value in the tens place is indicated as the stem.
- If the values of the data are from 100 to 999, the value in the units place is • indicated as the leaf and the values in the tens and hundreds places considered together is indicated as the stem.
 - Only the digit in the units place is indicated as the leaf.
 - For values from 0 to 9, the stem takes the value 0.
 - If a row has more than one leaf value, the values are written leaving a gap between the digits.

Example 1

- (i) Write the stem and leaf of each of the numbers 2, 43 and 225.
- (ii) Write the datum of which the stem is 3 and the leaf is 0.

(i)	Data	Stem	Leaf
	2	0	2
	43	4	3
	225	22	5
(ii) 30			



The marks obtained by 25 students in a certain class for a mathematics test paper marked out of 50 are given below.

5	7	9	11	13
16	19	20	21	22
24	25	26	26	29
31	33	35	36	38
40	43	45	48	49

Let us represent this data in a stem and leaf diagram.

In a stem and leaf diagram, the first column is called the stem and the second column is called the leaf.

Stem	Leaf
0	5 7 9
1	1 3 6 9
2	0 1 2 4 5 6 6 9
3	1 3 5 6 8
4	0 3 5 8 9 Key: 3 1 means 31.

- All the numbers are written in ascending order such that the stems of the numbers are in the first column (stem column) and the leaves of the numbers are in the second column, and with the numbers from 0 to 9 in the first column, the numbers from 10 to 19 in the second column and the numbers from 20 to 29 in the third column etc.
- The numbers in the fourth row of the above stem and leaf diagram have 3 as the stem and 1, 3, 5, 6, 8 respectively as the leaves. Their corresponding values are 31, 33, 35, 36, and 38.

The numbers represented in the other rows can also be written as shown above.

- It is easier to understand information related to the above 25 data when they are represented in a stem and leaf diagram than when they are written in a row.
- If the students who obtained less than 20 marks failed the test, then we can easily say that the number of students who failed is 3 + 4 = 7.
- If an "A" pass is given to those who have obtained 40 or more marks, then we can easily say by considering the stem and leaf diagram that there are 5 such students.

Therefore, a stem and leaf diagram can be considered as a simple method of representing and understanding data.

Now let us consider through an example how data is organized in ascending order.

Example 2

The heights of some students in a class are given below in centimetres.

141	148	142	130	152	135	157	146	140	160
151	173	139	135	144	134	151	138	137	137
169	136	143	154	146	166	131	150	145	143

(i) Represent this data in a stem and leaf diagram.

(ii) What is the least value of this collection of data?

(iii) What is the greatest value of this collection of data?

P

Stem	Leaf										
13	0	5	9	5	4	8	7	7	6	1	
14	1	8	2	6	0	4	3	6	5	3	
15	2	7	1	1	4	0					
16	0	9	6								
17	3										

Key : 14|1 means 141.

The stem and leaf diagram prepared with the data values in ascending order

is given be	low.	Stem					L	eaf	f			
		13	0	1	4	5	5	6	7	7	8	9
		14	0	1	2	3	3	4	5	6	6	8
		15	0	1	1	2	4	7				
		16	0	6	9							
		17	3									
(ii) 130	(iii) 173	3										

Now let us consider through the following examples how a collection of data consisting of decimal numbers is represented in a stem and leaf diagram.

Example 3

The birth weights of 25 animals of a certain species are given below in kilogrammes.

6.1	9.8	6.7	8.1	5.6	6.4	7.5	8.6
8.5	7.2	9.5	6.8	8.9	7.3	6.8	7.7
9.3	9.0	8.4	7.6	8.2	8.5	7.9	8.3
95							

- (i) Represent this data in a stem and leaf diagram.
- (ii) What is the minimum birth weight?
- (iii) What is the maximum birth weight? Ŕ
 - (i) In these decimal numbers, the whole number parts take values from 5 to 9. These are taken as the stems and the decimal parts are taken as the leaves.

	Stem	Leaf
	5	6
	6	1 4 7 8 8
	7	2 3 5 6 7 9
	8	1 2 3 4 5 5 6 9
	9	0 3 5 5 8
		Key : 7 3 means 7.3
(ii) 5.6 kg		
(iii) 9.8 kg		

Exercise 27.1

(1) The period of service of a group of employees of a certain company are given below in months. Represent this data in a stem and leaf diagram. 1

20	145	164	156	134	129	132	145	158	162

(2) The mass in kilogrammes of the bags of 30 pilgrims who flew to their destination in Dambadiva are given below. Represent this data in a stem and leaf diagram.

30	29	27	28	19	22	18	21	20	24
28	12	23	30	09	21	17	25	27	26
26	10	29	25	24	20	15	29	29	28

(3) The masses of the water melons for sale in a certain shop on a particular day are given below in kilogrammes.

6.5	7.8	5.7	4.3	5.8	6.2	4.3	6.9	7.8	7.2
6.9	5.5	7.7	7.8	5.2	6.7	5.7	6.1	6.0	7.3
7.1	6.7	7.7	4.3	6.5	7.3	6.7	5.8	6.8	5.4

- (i) Represent this data in a stem and leaf diagram.
- (ii) How many water melons are there for sale in this shop on this day?
- (iii) What is the mass of the heaviest water melon for sale in this shop?
- (iv) What is the mass of the water melon with the least mass?

27.2 Distribution of data represented in a stem and leaf diagram

The number of customers who bought gift items from a certain shop on each day of a period of 30 days is given below.

Stem	L	eaf	•			
0	8	9				
1	2	8	9			
2	3	2	6	6	9	
3	0	5	6	8		
4	0	1	1	4		
5	3	4	6	7		
6	2	5	8			
7	2	4	6			
8	0	1				

Key :4|0 means 40.

• The minimum value of this collection of data is 8.

This is the minimum number of customers who bought items from the shop on a day in that period of 30 days.

• The maximum value of this collection of data is 81.

This is the maximum number of customers who bought items from the shop on a day in that period of 30 days.

• Accordingly, this data is distributed from 8 to 81. To find the range of this data, we use;

```
Range = Maximum value – Minimum value
= 81 - 8
= 73
```

126

• When the groups of ten from 0 to 90 are considered, the maximum number of data, that is 5 data, is in the group 20 – 29. The minimum number of data, that is 2 data, is in the groups 0 – 9 and 80 – 89.

Exercise 27.2

(1) A cyclist had a training schedule for a month. The distance he cycled each day is given below in kilometres.

Stem	L	eaf							
1	5	5	8						
2	0	1	3	4	6	7			
3	2	4	5	6	6	8	8		
4	0	2	4	4	5	6	8	8	
5	1	2	4	6					
6	3	5							

Key : 5|1 means 51.

- (i) What is the minimum value of this data?
- (ii) What is the maximum distance he cycled in a day during this period?
- (iii) Find the range of this data.
- (2) 30 students in Grade 8 were given 40 English words to read and then write down. The number of incorrect words written by each student is given below.

16	24	12	15	10	23
23	15	13	19	14	25
26	21	31	24	19	27
35	12	17	29	18	29
32	18	27	31	21	31

- (i) Represent this data in a stem and leaf diagram.
- (ii) How many incorrect words were written by the student who wrote the least number of incorrect words?
- (iii) How many incorrect words were written by the student who wrote the most number of incorrect words?
- (iv) Find the range of the incorrect words written by this group of students.
- (v) Write the groups of ten to which the greatest and least values belong.



Fish	buns sold	Bot	tles of fruit juice sold
Stem	Leaf	Stem	Leaf
5	456889	0	8 9
6	033588	1	0 2 5
7	233599	2	0 1 3 5 8 9
8	0 0 3 4 5 7	3	5 6
9	0 1 3 4 4 5	4	3 4 5
		5	0 2 6 8
		6	1
Kev ·6	3 means 63	7	0 2 5
itey .0	jo means ob.	8	1 4
		9	0 2 4 6

Key :8|1 means 81.

- (i) What is the minimum number of fish buns sold in a day?
- (ii) What is the maximum number of fish buns sold in a day?
- (iii) Find the range of the sales of fish buns.
- (iv) What is the minimum number of bottles of fruit juice sold in a day?
- (v) What is the maximum number of bottles of fruit juice sold in a day?
- (vi) Find the range of the sales of bottles of fruit juice.
- (vii) Compare the sales of fish buns with the sales of fruit juice and write your conclusions.
- (4) The marks obtained for a mathematics test paper marked from 100, by the students of two parallel classes *A* and *B* are given below.

	Class A		Class B	
Stem	Leaf	Stem	Leaf	
5	026	0	59	
6	0135668	1	0256	
7	2 2 3 5	2	1	
8	0 2	3	2 3	
	I	4	4 5 8	
		5	1 3	
		6	08	

Key :7|2 means 72.

Key :5|1 means 51.



- (i) Write separately the number of students in class A and the number of students in class B.
- (ii) Find the minimum mark, maximum mark and the range of the marks obtained by the students in class *A*.
- (iii) Find the minimum mark, maximum mark and the range of the marks obtained by the students in class B.
- (iv) Compare the level of achievement in mathematics of the students in classes A and B for this mathematics paper based on the above data and write your conclusions.

27.3 Interpreting a collection of numerical data

Now let us consider how a given collection of data is interpreted.

- During a pluck, we can usually get 8 coconuts from each coconut tree in an estate.
- The average mark of a student for 8 subjects is 73.6.
- The average runs scored per over in a certain cricket match was 5.3.
- On a certain day, the price of 1 kg of beans displayed by most of the vendors in a market was 120 rupees.

A single value that is used to give an idea regarding a collection of data, as in the above examples is called a **representative value**.

We will now consider a few representative values that are used.

• Mode

The marks obtained by the 13 students in a certain class for a mathematics question paper are given below.

96, 81, 78, 45, 71, 57, 71, 83, 95, 68, 94, 71, 79

The 'number of data in a collection' is the total number of data in that collection.

The number of data in the above collection is 13.

Let us write these values in ascending order.

45, 57, 68, 71, 71, 71, 78, 79, 81, 83, 94, 95, 96

The mark that has been obtained the most is 71. Three students have obtained this mark.



In a collection of data, some of the values could be identical. The value which occurs most often is called the **mode** of that collection of data.

In the above collection of data, since 71 is the value that occurs most often, 71 is the mode.

Note:

It is not necessary to write the data in ascending order to find the mode.

Example 1

The ages of 10 students in Grade 8 are given below in years. Find the mode of this collection of data.

13 14 15 14 15 14 14 14 13 14

In the above collection of data, 14 years is the value that occurs most frequently. Therefore the mode of the ages of these grade 8 students is 14.

Example 2

The number of employees who took leave on each of the 15 working days of a certain month is given below. Find the mode of this collection of data.

12	14	20	16	15
16	21	19	16	18
17	15	18	19	18

Here, the values 16 and 18 have each occurred 3 times. The other values have occurred less than 3 times. Therefore, we can take 16 or 18 as the mode of this collection of data.

Such a distribution of data is known as a bimodal distribution. A collection of data may have more than two modes too.

• Median

> Let us consider a collection of data with an odd number of values.

The median of a collection of data is the value of the datum in the centre, when the data is arranged in ascending order.

3, 9, 9, 11, 15, 22, 24, 25, 31, 37, 40



There are 11 data. The 6th datum is in the centre. Its value is 22. There are 5 data values less than 22 and 5 data values greater than 22.

If the number of values in a collection of data is an odd number, then the value in the centre, when the values are arranged in ascending order, is the median of the collection of data.

Accordingly, the median of the above collection of data is 22.

When the values of the above collection of data are arranged in ascending order, the value in the centre is the $\frac{11 + 1}{2} = 6$ th value. Therefore, the median of this set of data is 22.

> Now let us consider a collection of data with an even number of values.

3, 9, 9, 11, 15 (22), (24), 25, 31, 37, 40, 41

There are 12 values in this collection of data, which is an even number. Here, we cannot find a datum in the centre. The two data in the centre are the 6th and 7th data. Their values are 22 and 24 respectively.

- If there is an even number of data, the median is half the sum of the values of the two data in the centre, when the data are arranged in ascending order.
- If there is an even number of data, and if the values of the data are arranged in ascending order, then the data in the centre are the $(\frac{\text{number of data}}{2})$ th datum and the $(\frac{\text{number of data}}{2}+1)$ th datum.
- There fore the median of the above data is $\frac{22 + 24}{2}$; that is, 23. There are 6 data less than 23 and 6 data greater than 23.

Example 3

The number of soft drink bottles sold at a certain shop on each day of a week is given below. Find the median of the number of bottles sold during a day.

44 32 60 52 48 41 40 When this data is organized in ascending order we obtain 32 40 41 44 48 52 60 The median is 44. The median of this collection of data is 44.

Example 4

The number of athletes who came to a certain sports centre for training each day of a period of 16 days is given below. Find the median of the number of athletes who came for training to the sports centre each day.

2.2 12 16 23 36 18 09 14 26 15 18 25 20 21 20 15 When this data is organized in ascending order we obtain,

09 12 14 15 15 16 18 18 20 20 21 22 23 25 26 36

There are two values in the centre.

There are two values in the centre because there are 16 scores in total. They are the value of the $\frac{16}{2}$ = 8th datum and the value of the $\frac{16}{2}$ + 1 = 9th datum.

The value of the 8th datum = 18

The value of the 9th datum = 20

: the median =
$$\frac{18+20}{2} = 19$$

The median number of athletes who came for training to the sports centre each day is 19.

Example 5

(i) A group of 15 students who are members of a drill display, are made to stand in a row as shown in the figure in ascending order of their heights, after their heights were measured in centimetres. Find the median of this collection of data.



In the figure, the student in the centre is caged. This is the 8th student. We can easily obtain the value in the centre when there is an odd number of data. The median of this distribution is the value of the $\frac{15+1}{2}$ = 8th datum. The height of the 8th student is 127 cm. Therefore, the median of this collection of data is 127 cm.



(ii) Assume that a new student of height 150 cm joined the end of this row of students. Find the median of this collection of data.



Now there are 16 values in this collection of data. If we order these students as previously, two students are in the centre. They are the 8th and 9th students. Accordingly, the median is the value that is obtained when the heights of the 8th and 9th students are added together and divided by 2. Therefore the median is

127 + 135 cm; that is, 131 cm.

Mean

The average value of a collection of data is considered as its mean.

The mean of a collection of data is the value that is obtained when the sum of all the values of the collection of data is divided by the number of values.

$$Mean = \frac{The total sum of the values of the collection of data}{number of data}$$

Example 6

The

The marks obtained by 13 students of a certain class for a mathematics test paper marked out of 100 are given below in ascending order. Find the mean of this data.

45, 57, 69, 71, 71, 71, 78, 79, 81, 81, 94, 95, 96

Mean =
$$\frac{\text{The total sum of the values of the collection of data}}{\text{number of data}}$$
The mean of the data
$$=\frac{45 + 57 + 69 + 71 + 71 + 78 + 79 + 81 + 81 + 94 + 95 + 96}{13}$$

By comparing the value 76 with the total marks of 100 allocated for this mathematics test paper, we can assess these students' knowledge and skills in mathematics.



• Range

The marks obtained by the students in three classes for a mathematics paper are given below.

A 57 58 60 60 60 62 63 Median of the marks obtained by the students in class A = 60Mean of the marks obtained by the students in class A = 60B 35 45 55 60 65 75 85 Median of the marks obtained by the students in class B = 60Mean of the marks obtained by the students in class B = 60*C* 31 42 55 60 69 73 90 Median of the marks obtained by the students in class C = 60Mean of the marks obtained by the students in class C = 60

The marks of the students vary across the classes, but the medians and the means are the same for all three classes.

In such situations, interpreting the data using only the median and the mean is insufficient We need to consider the dispersion (spread) of the data too. For this we use measurements related to the dispersion of the data.

In this lesson we learn about one measurement of dispersion, namely the range.

The marks obtained by 8 students in a certain class for a mathematics test paper marked out of 100 are given below.

```
12, 28, 56, 48, 32, 64, 80, 92
```

Let us write the above values in ascending order.

12, 28, 32, 48, 56, 64, 80, 92

The maximum value of the above data set is 92 and the minimum value is 12. The difference between the maximum value and the minimum value is 92 - 12 = 80. It indicates that the dispersion of the data is 80 units.

The difference between the greatest value and the least value of a collection of data is called **its range**.

Range = Greatest value – Least value

The range of the above set of data is 80.



- The difference between the maximum mark of 100 and the minimum mark of 0 that can be scored by as student for the above mathematics test paper is 100.
- When the value of the range is comparatively low, the data take values which are close to each other. The range 80 in the above example is relatively large, when compared with 100. Therefore we can conclude that the marks are not close to each other.

Example 7

The marks obtained by another class of 8 students for a mathematics test paper are written below in ascending order. Find the range of these marks.

The range of the above marks = 54 - 46 = 8

The range 8 in this example is relatively small when compared with 100. Therefore we can say that the marks are approximately at the same level, and conclude that these students' knowledge tested in this paper is at about the same level.

The most suitable representative value:

The runs scored by a cricketer during 8 overs of a cricket match are given below.

3, 8, 9, 12, 5, 3, 5, 3

The total number of runs he scored is 48. When the runs he scored in each of the 8 overs are written in ascending order we obtain

3, 3, 3, 5, 5, 8, 9, 12

The mode of this collection of data is 3.

The median
$$= \frac{5+5}{2} = 5$$

The mean $= \frac{48}{8}$
 $= 6$

- > The mode value of 3 indicates that the runs he scores in an over is most often 3.
- The median value of 5 indicates that the likelihood of scoring 5 or less runs per over is the same as the likelihood of scoring 5 or more runs per over.
- > The mean value of 6 indicates that the rate at which he scores runs is 6 per over.





Exercise 27.3

- (1) Find the mode, the median, the mean and the range of each collection of data.
 - (i) 8, 9, 12, 10, 12, 7, 8, 6, 10, 5, 8
 - (ii) 33, 32, 18, 33, 45, 23, 53, 32, 33
 - (iii) 78, 78, 80, 70, 78, 65, 69, 70
 - (iv) 3.5, 2.5, 4.8, 1.3, 3.9
 - (v) 12.5, 32.4, 23.6, 8.3

(2) The number of matchsticks in 10 boxes of matches is given here.

49, 50, 48, 47, 49, 50, 49, 50, 47, 51.

For these boxes of matchsticks, find

- (i) the mode,
- (ii) the median,
- (iii) the mean number of matchsticks in a box.



- (3) The temperature in the 9 provinces of Sri Lanka on a certain day are given here. 26°C, 27°C, 28°C, 32°C, 29°C, 28°C, 30°C, 29°C, 28°C. What was the mean temperature on that day?
- (4) The masses of a group of children of the same age who arrived at a clinic on a certain day are given here. 15 kg, 16 kg, 18 kg, 12 kg, 14 kg, 16 kg, 17 kg, 20 kg.
 - (i) What is the mode of the masses of this group of children?
 - (ii) If the children are kept in a row in an ascending order of their masses, what would be the mass of the child in the centre?
 - (iii) According to the given data, what is the mean mass of a child in this group?
- (5) The number of runs scored by each of the 11 batsmen in each of the two teams that played a cricket match against each other is given in the following table.

Batsman	1	2	3	4	5	6	7	8	9	10	11
Team A	34	42	58	5	32	21	16	0	9	3	12
Team B	8	0	12	33	31	60	44	36	24	12	6

(a) By considering the runs scored by the batsmen in team A, find;
(i) the minimum value
(ii) the maximum value
(iii) the range
(iv) the median
(v) the mean

(b) By considering the runs scored by the batsmen in team *B*, find;
(i) the minimum value
(ii) the maximum value
(iii) the range
(iv) the median
(v) the mean



(c) Using the above information, fill in the table given below.

Team	Range	Mean	Median
A			
В			

- (d) From which representative value is the total marks of a cricket team obtained accurately? Show how you get this answer.
- (6) The mean of the masses of 4 children is 34 kg. When another child joined, the mean mass increased to 38 kg.
 - (i) Find the total mass of the 4 children.
 - (ii) What is the mass of the child who joined later?
 - (iii) Show that the mean mass of 34 kg does not change if the mass of the child who joined later is also 34 kg.

Miscellaneous Exercise

- (1) A bowler allows the opponent batsmen to score 52 runs in his 10 overs. Find the mean number of runs he gave per over.
- (2) A group of 15 pilgrims are in an airplane. The mean mass of their luggage is 29 kg. Each person can carry up to 30 kg of luggage. If this is exceeded, an additional fee is charged.
 - (i) What is the total mass of the luggage carried by this group?
 - (ii) What is the total mass allowed for this group?





(3) The marks that Malitha and Dilitha got in the previous term test are as follows.

Subject	Sinhala	English	Mathematics	Science	Buddhism	Geography	Art	Agriculture & Food Technology	History
Malitha	39	40	65	60	56	64	70	65	54
Dilitha	64	55	42	58	70	68	49	70	45

(i) Complete the table given below.

	Malitha	Dilitha
Mode of the marks		
Median of the marks		
Number of subjects for which the mark exceeds 50		

- (ii) Find separately the median mark of each student.
- (iii) What is the most suitable representative value to compare two collections of data? Give reasons for your answer.
- (4) The total marks obtained in a term test for all the subjects offered is given below for a group of students in a certain class. Represent this data in a stem and leaf diagram.

481	706	609	689	273	538	386	525	720	356
529	513	634	713	673	224	736	281	613	496
671	381	524	591	613	729	681	673	571	351



Stem	Leaf	
25	0 2 5	
26	1 4 6 8	
27	0 0 0 5 6 7	8 9
28	0 1 5 5 5	
29	0 1 2	
30	0 0 0	Key : 28 1 means 281.

- (i) What is the minimum value of this data?
- (ii) What is the maximum value?
- (iii) Find the range.

Summary

- Data can be represented easily using a stem and leaf diagram. Understanding the data is facilitated by using a stem and leaf diagram.
- In a collection of data, some of the values could be identical. The value which occurs most often is called the mode of that collection of data.
- If the number of values in a collection of data is an odd number, then the value at the centre, when the values are arranged in ascending order, is the median of the collection of data.
- □ If there is an even number of data, the median is half the sum of the values of the two data at the centre, when arranged in ascending order.
- The mean of a collection of data is the value that is obtained when the sum of all the values of the collection of data is divided by the number of values.
- The difference between the greatest value and the least value of a collection of data is called its range.

Range = Greatest value – Least value

