

13: Statistical Diagrams

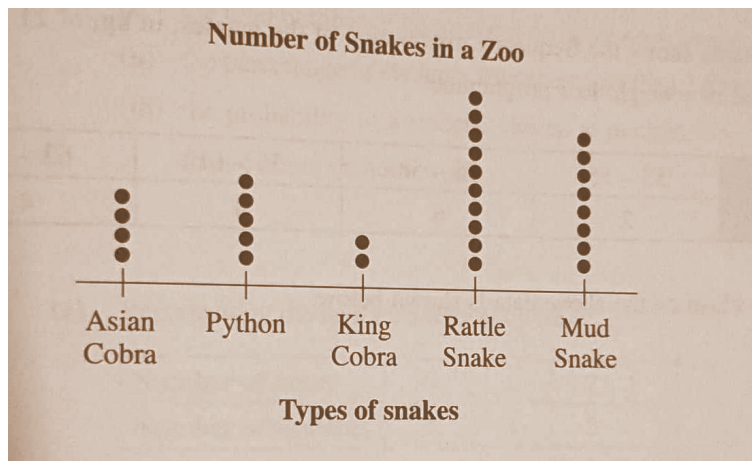
Dot Diagrams:

1, In a **dot diagram** or **dot plot**, a set of data is represented by using dots and placing them above a horizontal number line. The number of dots over the number line indicates the value of the data points. Usually, we start with making a **frequency table**.

For example:

Types of snakes	Tally	Frequency
Asian Cobra		4
Python		5
King Cobra		2
Rattle Snake		10
Mud Snake		8
Total frequency		29

After that, we can plot them on a dot diagram as such:



2. The summary of advantages and disadvantages of dot diagram:

Advantages	Disadvantages
Suitable to display a small set of numerical data	Not suitable to display a large set of numerical data (since each dot only represents one unit)
	Not suitable to display a data set with many different values or a large range

Stem-and-Leaf Diagrams:

3. In a **stem-and-leaf diagram**, the data values are split into two parts – the stem and the leaves. The ‘stem’ values are listed down while the ‘leaf values’ go right (or left) from the stem values.

4. For comparison of two data sets, we can use a **back-to-back stem-and-leaf diagram** which has a common stem with two leaves, one on the left and the other on the right of the stem.

5. To construct a single ordered stem-and-leaf diagram, we can use the following steps:

- 1) Identify the smallest and greatest number in the data set.
- 2) Draw a vertical line and write the digits that form the ‘**stems**’ on the left.
- 3) Write the digits that form the ‘**leaves**’ on the right of the vertical line.
- 4) Rewrite the digits in each row by arranging them from the smallest to the greatest number (i.e. in ascending order)
- 5) Include an explanation (key) to show the meaning of the numbers in the diagram.

Examples of stem-and-leaf diagrams:

Single Ordered		Back-to-Back																																																																				
<table><tr><th>Stem</th><th>Leaf</th></tr><tr><td>85</td><td>1</td></tr><tr><td>86</td><td>0 3 4</td></tr><tr><td>87</td><td>6 7</td></tr><tr><td>88</td><td>2 5</td></tr><tr><td>89</td><td>0</td></tr><tr><td>90</td><td>1 2 3 5 7</td></tr><tr><td>91</td><td>0 2 3 4</td></tr><tr><td>92</td><td>1 7</td></tr><tr><td>Key:</td><td>88 2 means 88.2 points</td></tr></table>		Stem	Leaf	85	1	86	0 3 4	87	6 7	88	2 5	89	0	90	1 2 3 5 7	91	0 2 3 4	92	1 7	Key:	88 2 means 88.2 points	<table><tr><th colspan="2">Class 2B</th><th></th><th colspan="2">Class 2A</th></tr><tr><th></th><th>Leaf</th><th>Stem</th><th>Leaf</th><th></th></tr><tr><td></td><td>8 8 8</td><td>0</td><td>5</td><td></td></tr><tr><td></td><td>5 2</td><td>1</td><td>0 3</td><td></td></tr><tr><td>9 7 7 5 4 3</td><td>2 2 0</td><td>2</td><td>0 4</td><td></td></tr><tr><td></td><td>7 6 0</td><td>3</td><td>0 1 3 4 5 6 8</td><td></td></tr><tr><td></td><td>2 0 0</td><td>4</td><td>0 5 6 7 9 9 9</td><td></td></tr><tr><td></td><td></td><td>5</td><td>0</td><td></td></tr><tr><td colspan="5">Key: 2 4 means 24 push-ups</td></tr></table>				Class 2B			Class 2A			Leaf	Stem	Leaf			8 8 8	0	5			5 2	1	0 3		9 7 7 5 4 3	2 2 0	2	0 4			7 6 0	3	0 1 3 4 5 6 8			2 0 0	4	0 5 6 7 9 9 9				5	0		Key: 2 4 means 24 push-ups				
Stem	Leaf																																																																					
85	1																																																																					
86	0 3 4																																																																					
87	6 7																																																																					
88	2 5																																																																					
89	0																																																																					
90	1 2 3 5 7																																																																					
91	0 2 3 4																																																																					
92	1 7																																																																					
Key:	88 2 means 88.2 points																																																																					
Class 2B			Class 2A																																																																			
	Leaf	Stem	Leaf																																																																			
	8 8 8	0	5																																																																			
	5 2	1	0 3																																																																			
9 7 7 5 4 3	2 2 0	2	0 4																																																																			
	7 6 0	3	0 1 3 4 5 6 8																																																																			
	2 0 0	4	0 5 6 7 9 9 9																																																																			
		5	0																																																																			
Key: 2 4 means 24 push-ups																																																																						

Continue to the next page.

6. The summary of advantages and disadvantages of stem-and-leaf diagram:

Advantages	Disadvantages
Suitable to display a data set with many different values by grouping the values into equal class intervals	Tedious to construct if there are too many data values
Individual data values are retained (as leaves)	Cannot group into class intervals such as 141-150, 151-160, 161-170, etc. (because of the stem)
Can display two sets of data back-to-back for easy comparison	

Histograms:

7. A **histogram** is similar to a vertical bar chart except that there are no spaces between the adjacent bars. In a histogram, the frequency corresponding to a class is represented by the area of a bar (or rectangle) whose base is the class width.

6. A histogram is different from a bar chart as it is the area of the bar, not the height, that denotes the value (or frequency). However, if the widths of the bars are equal then only the heights need to be considered since the heights of the bar are proportional to the values (or frequencies.)

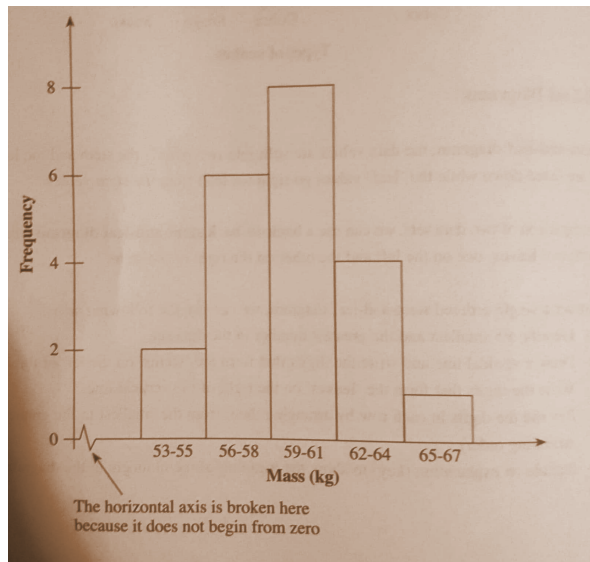
7. In a histogram,
- (a) there are no gaps between the rectangles,
 - (b) the y -axis is the frequency and it always starts at 0.

Example of constructing a histogram:

The following table shows the frequency distribution of the masses, in kg , of 21 members of a health club who are enrolled in a weight loss programme.

Mass (kg)	53 – 55	56 – 58	59 – 61	62 – 64	65 – 67
Frequency	2	6	8	4	1

A histogram to illustrate the above data is shown on the next page.



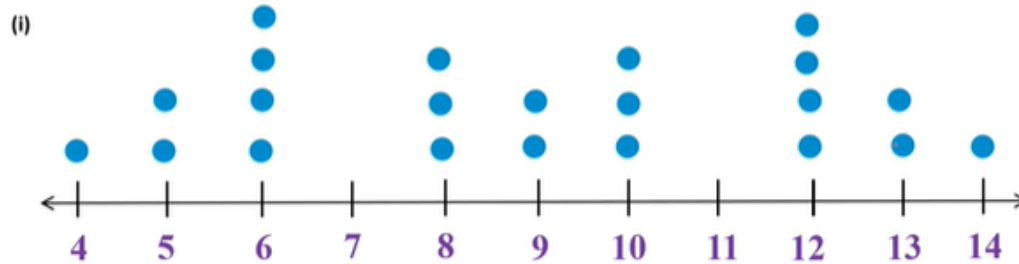
8. The summary of advantages and disadvantages of histograms:

Histogram for ungrouped data	
Advantages	Disadvantages
Suitable for displaying a large set of numerical data	More abstract because frequency is represented by area of each column
	If the frequency axis does not start from zero, the data may be misinterpreted

Histogram for grouped data	
Advantages	Disadvantages
Suitable to display a data set with many different values by grouping the values into equal class intervals	More abstract because frequency is represented by area of each column
Suitable to display a data set when individual data values are not important	Individual data values are lost
Easier to construct than a stem-and-leaf diagram, especially when there are many values	
Can group into class intervals such as 141-150, 151-160, 161-170 etc. (unlike stem-and-leaf diagrams)	

Example Questions:

This dot diagram represents the number of art pieces a group of students submit for an exhibition.



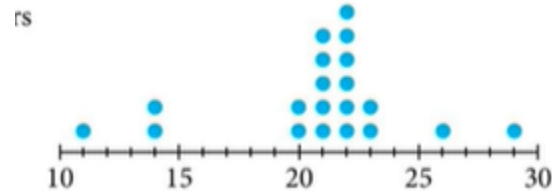
(ii) If each graduating student submits at least 10 pieces, find the number of graduating students in this group.

- $3 + 4 + 2 + 1 = 10 //$

(iii) Briefly describe the distribution of the data.

- **The number of art pieces submitted is from a range of 4 to 14. The most common number of art pieces submitted is 6 and 12. The distribution is also symmetrical about 9.**

The dot diagram represents the ages, in years, of contestants who join a karaoke contest.



(i) A contestant is selected at random. The probability of the age of the contestant is less than or equal to x years is $\frac{1}{4}$. Find the value of x .

- **Total number of contestants:** $1 + 2 + 2 + 5 + 6 + 2 + 1 + 1 = 20 //$

- **Number of contestants that can be selected:** $\frac{1}{4} \times 20 = 5 //$

- **Since there is**
- **1 10 year old**
- **2 14 year olds**
- **2 20 year olds**
- **= 5 that can be selected.**

- $\therefore x = 20$

(ii) Briefly describe the distribution of the data.

- **The age, in years, of 20 contestants who join the karaoke contest ranges from 11 years to 29 years. The age clustered around 20 to 23 years. The extreme ages of 11 years, 14 years and 29 years deviate considerably from the other ages recorded. The distribution is also not symmetrical.**

Copyright Issues:

Content shared for educational purposes only. These notes are based on materials originally published by © SHING LEE PUBLISHERS PTE LTD (credit given). The content has been modified and expanded by Om Mahto, who retains ownership of these modifications.

Version 3