STATISTICS

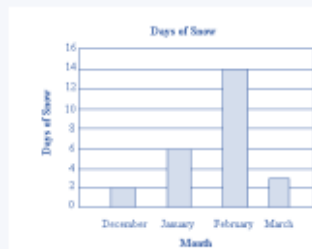| S.no	Term	Description
1	Statistics	Statistics is a broad mathematical discipline which studies ways to collect, summarize, and draw conclusions from data
2	Data	A systematic record of facts or different values of a quantity is called **data**. Data is of two types - Primary data and Secondary data. **Primary Data**: The data collected by a researcher with a specific purpose in mind is called primary data. **Secondary Data**: The data gathered from a source where it already exists is called secondary data
3	Features of data	• Statistics deals with collection, presentation, analysis and interpretation of numerical data. • Arranging data in an order to study their salient features is called presentation of data. • Data arranged in ascending or descending order is called arrayed data or an array • **Range** of the data is the difference between the maximum and the minimum values of the observations • Table that shows the frequency of different values in the given data is called a **frequency distribution table** • A frequency distribution table that shows the frequency of each individual value in the given data is called an ungrouped frequency distribution table. • A table that shows the frequency of groups of values in the given data is called a grouped frequency distribution table • The groupings used to group the values in given data are called classes or class-intervals. The number of values that each class contains is called the class size or class width. The lower value in a class is called the lower class limit. The higher value in a class is called the upper class limit.

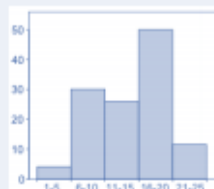|		
		• **Class mark** of a class is the mid value of the two limits of that class. • A frequency distribution in which the upper limit of one class differs from the lower limit of the succeeding class is called an **Inclusive or discontinuous Frequency Distribution**. • A frequency distribution in which the upper limit of one class coincides from the lower limit of the succeeding class is called an **exclusive or continuous Frequency Distribution** • .
4	Bar graph	A bar graph is a pictorial representation of data in which rectangular bars of uniform width are drawn with equal spacing between them on one axis, usually the x axis. The value of the variable is shown on the other axis that is the y axis.
5	Histogram	A histogram is a set of adjacent rectangles whose areas are proportional to the frequencies of a given continuous frequency distribution
6	Mean	The mean value of a variable is defined as the sum of all the values of the variable divided by the number of values. $$a_m = \frac{a_1 + a_2 + a_3 + a_4}{4} = \frac{\sum_0^n a}{n}$$

7	Median	The **median** of a set of data values is the middle value of the data set when it has been arranged in ascending order. That is, from the smallest value to the highest value
		Median is calculated as
		$$\frac{1}{2}(n+1)$$ Where n is the number of values in the data If the number of values in the data set is even, then the **median** is the average of the two middle values.
8	Mode	Mode of a statistical data is the value of that variable which has the maximum frequency

S.no	Term	Description
1	**Mean for Ungroup Frequency table**	Mean is given by $$M = \frac{\sum f_i x_i}{\sum f_i}$$
2	**Mean for group Frequency table**	In these distribution, it is assumed that frequency of each class interval is centered around its mid-point i.e. class marks $$Class\ mark = \frac{Upper\ class\ limit + Lower\ class\ limit}{2}$$ **Mean can be calculated using three method** **a) Direct method** $$M = \frac{\sum f_i x_i}{\sum f_i}$$ **b) Assumed mean method** $$M = a + \frac{\sum f_i d_i}{\sum f_i}$$

Where

a=> Assumed mean

d_i => $x_i - a$

c) Step deviation Method

$$M = a + \frac{\sum f_i u_i}{\sum f_i} h$$

Where

a=> Assumed mean

u_i => $(x_i - a)/h$

3	Mode for grouped frequency table	*Modal class: The class interval having highest frequency is called the modal class and Mode is obtained using the modal class*
		$$M_o = l + \left(\frac{f_1 - f_0}{2f_1 - f_0 - f_2}\right) h$$
		Where l = lower limit of the modal class, h = size of the class interval (assuming all class sizes to be equal), f_1 = frequency of the modal class, f_0 = frequency of the class preceding the modal class, f_2 = frequency of the class succeeding the modal class.
4	Median of a grouped data frequency table	For the given data, we need to have class interval, frequency distribution and cumulative frequency distribution Median is calculated as
		$$M_m = l + \left(\frac{\frac{n}{2} - cf}{f}\right) h$$
		Where l = lower limit of median class, n = number of observations, cf = cumulative frequency of class preceding the median class, f = frequency of median class, h = class size (assuming class size to be equal)

| 5 | Empirical Formula between Mode, Mean and Median | *3 Median=Mode +2 Mean* |