

UNIT 8 - STATISTICS

17

Frequency Distribution

POINTS TO REMEMBER

1. **Data.** *The word data means a set of given facts in numerical figures.*
2. **Statistics.** *It is the science which deals with the collection, presentation, analysis and interpretation of numerical data.*
In singular form, statistics is taken as a subject.
And, in plural form, statistics means data.
3. **Fundamental Characteristics of Data**
 - (i) Numerical facts alone form data. Qualitative characteristics, like honesty, poverty, etc., which cannot be measured numerically do not form data.
 - (ii) Data are aggregate of facts. A single observation does not form data.
 - (iii) Data collected for a definite purpose may not be suitable for another purpose.
4. **Types of Data**
 - (i) **Primary Data.** *The data collected by the investigator himself with a definite plan in mind are known as primary data.*
 - (ii) **Secondary Data.** *The data collected by someone, other than the investigator, are known as secondary data.*
5. **Raw or Ungrouped Data.** *The data obtained in original form are called raw data or ungrouped data.*
6. **Array.** *An arrangement of raw data in ascending or descending order of magnitude is called an array.*
7. **Presentation of Data.** *Putting the data in condensed form in the form of a table, is known as presentation of data.*
8. **Frequency.** *The number of times an observation occurs is called its frequency.*
9. **Frequency Distribution.** *The tabular arrangement of data showing the frequency of each observation is called its frequency distribution.*
10. **Grouped Frequency Distribution.** *To put the data in a more condensed form, we make groups of suitable size, and mention the frequency of each group. Such a table is called a grouped frequency distribution table.*
11. **Types of Grouped Frequency Distribution.** *Suppose the marks obtained by some students in an examination are given. We may put them into forms given below :*

(i) Discontinuous Interval (i.e. Inclusive Form) :

We may consider the groups such as 1 – 10, 11 – 20, etc., called **classes**.

Here, by 1 – 10 means marks obtained from 1 to 10, **both included**.

Here, 1 is called the **lower limit** and 10 is called the **upper limit** of the class 1 – 10.

The number of observations in a class is called its **frequency**.

(ii) Continuous Interval (i.e. Exclusive Form) :

We may consider the groups such as 1 – 10, 10 – 20, etc.

Here, by 1 – 10 means marks from 1 to 9, **excluding the upper limit and including the lower limit**.

12. Variable. *A quantity which can take different values is called a variable.*

Examples. *Height, Age and Weight* of pupils in a class are three variables.

If we denote them by x , y and z respectively, then values of x give the heights of the pupils ; the values of y give the ages of the pupils and the values of z give the weights of the pupils.

(i) Discrete Variable. *A variable which cannot take all possible values between two given values, is called a discrete or discontinuous variable.*

Examples. (i) Number of members in a family.

(ii) Number of workers in a factory.

(ii) Continuous Variable. *A variable which can take any numerical value within a certain range is called a continuous variable.*

Examples. (i) Wages of workers in a factory. (ii) Heights of children in a class.

(iii) Weights of persons in a group etc.

13. Range of a Variable. *The difference between the maximum and minimum values of a variable is called its range.***14. Some Definitions Related to Grouped Data**

(i) Class Interval. *Each group into which the raw data is condensed, is called a class-interval.*

(ii) Class-Limits. *Each class-interval is bounded by two figures, called class limits.*

The figure to the left of a class is called its lower-limit and that on its right is called the upper-limit of the class.

EXERCISE 17

Q. 1. Define statistics as a subject.

Sol. It is a science which deals with the collection, presentation, analysis and interpretation of numerical data.

Q. 2. What are primary data and secondary data ? Which of the two are more reliable and why ?

Sol. (i) **Primary data.** The data collected by the investigator himself with a definite plan in mind are known as primary data.

(ii) **Secondary data.** The data collected by someone, other than the investigator are known as secondary data.

Primary data is more reliable because it is collected by the investigator or user himself.

Q. 3. Fill in the blanks :

(i) The difference between the maximum and minimum observations in a data is called the.....of the data.

(ii) The number of observations in a class-interval is called the.....of the interval.

(iii) The mid-point of class-interval is called the.....of the interval.

(iv) Lower-limit of the class-interval 24–30 is.....

(v) Upper limit of the class-interval 16–20 is.....

(vi) The class-mark of the class-interval 20–30 is.....

(vii) The class-mark of the class-interval 9.5–19.5 is.....

Sol. (i) range (ii) frequency
(iii) class-mark (iv) 24
(v) 20

$$(vi) \frac{20 + 30}{2} = \frac{50}{2} = 25$$

$$(vii) \frac{9.5 + 19.5}{2} = \frac{29}{2} = 14.5.$$

Q. 4. Find the range of the data :

(i) 5, 7, 16, 21, 8, 10

(ii) 11, 13, 17, 14, 19, 14, 15, 18.

Sol. (i) Highest or maximum value = 21
and minimum value = 5

$$\therefore \text{Range } 21 - 5 = 16$$

(ii) Maximum value = 19

Minimum value = 11

$$\therefore \text{Range} = 19 - 11 = 8 \text{ Ans.}$$

Q. 5. The class marks of a frequency distribution are 28, 34, 40, 46, 52. Find the class-size and all the class-intervals.

Sol. Class-size = $34 - 28 = 6$

\therefore First class-interval-will be 25 – 31

$$(\because 28 - 3 = 25, \text{ and } 34 - 3 = 31)$$

Similarly other class intervals will be 31 – 37, 37 – 43, 43 – 49 and 49 – 55. **Ans.**

Q. 6. State which of the following variables are continuous and which are discrete :

(i) Marks obtained by the students of a class in a test.

(ii) Daily maximum temperature of a city.

(iii) I.Q. of students of a class.

(iv) Weights of players of a Volley-Ball team.

(v) Number of car-accidents in a city.

(vi) Distance travelled by a train.

(vii) Time taken by runners in a race.

(viii) Sizes of shoes sold in a shoe-store.

(ix) Number of patients in a hospital per day.

Sol. (i) discrete (ii) continuous

(iii) continuous (iv) continuous

(v) discrete (vi) continuous

(vii) continuous (viii) discrete

(ix) discrete.

Q. 7. Define the following terms :

(i) Variable (ii) Class-interval

(iii) Class-size (iv) Class-mark

(v) Class-limits (vi) True class-limits

(vii) Frequency of a class

(viii) Cumulative frequency of a class.

Sol. (i) **Variable.** A quantity which can take different values is called a variable.

(ii) **Class-interval.** Each group into which the raw data is condensed is called a class-interval.

(iii) **Class size.** The difference of true upper limit and true lower limit is called class size.

(iv) **Class mark.** $\frac{1}{2}$ (lower limit + upper limit) is called class mark.

(v) **Class limits.** Each class interval is bounded by two figures, called limits.

(vi) **True class limits.** In continuous interval, the limits of the class is called true class limits.

(vii) **Frequency of a class.** The number of times an observations in a class, occurs is called its frequency.

(viii) **Cumulative frequency of class.** The sum of the frequencies of all the previous classes and that particular class, is called cumulative frequency of the class.

Q. 8. Following data gives the number of children in 40 families :

1, 2, 6, 5, 1, 3, 2, 6, 2, 3, 4, 2, 0, 4, 4, 3, 2, 2, 0, 0, 1, 2, 2, 4, 4, 3, 2, 1, 0, 5, 1, 2, 4, 3, 4, 1, 1, 6, 2, 2

Represent it in the form of a frequency distribution.

Sol. Below is given a frequency distribution of the given data.

Number of children	Tally-Marks	Frequency
0		4
1		7
2		12
3		5
4		7
5		2
6		3

Q. 9. The marks obtained by 40 students of a class in an examination are given below. Present the data in the form of a frequency distribution using equal class-size, one such class being 10–15 (15 not included).

3, 20, 13, 1, 21, 13, 3, 23, 16, 13, 18, 12, 5, 12, 5, 24, 9, 2, 7, 18, 20, 3, 10, 12, 7, 18, 2, 5, 7, 10, 16, 8, 16, 17, 8, 23, 21, 6, 23, 15.

Sol. Below is given the frequency distribution of the given data :

Class-interval	Tally-Marks	Frequency
1 – 5		6
5 – 10		10
10 – 15		8
15 – 20		8
20 – 25		8

Q. 10. Construct a frequency table for the following ages (in years) of 30 students using equal class-intervals, one of them being 9–12, where 12 is not included.

18, 12, 7, 6, 11, 15, 21, 9, 8, 13, 15, 17, 22, 19, 14, 21, 23, 8, 12, 17, 15, 6, 18, 23, 22, 16, 9, 21, 11, 16.

Sol. Below is given the frequency distribution of the given data :

Class-interval	Tall-Marks	Frequency
6 – 9		5
9 – 12		4
12 – 15		4
15 – 18		7
18 – 21		3
21 – 24		7

Q.11. The weekly wages (in rupees) of 30 workers in a factory are given below :
630, 635, 690, 610, 635, 636, 639, 645, 698, 690, 620, 660, 632, 633, 655, 645, 604, 608, 612, 640, 685, 635, 636, 678, 640, 668, 690, 606, 640, 690

Represent the data in the form of a frequency distribution with class size 10.

Sol. From the given data, Lowest data = 604 and largest data = 698

$$\therefore \text{Range} = 698 - 604 = 94$$

Class interval	Tally Marks	Frequency
600 — 610		3
610 — 620		2
620 — 630		1
630 — 640		9
640 — 650		5
650 — 660		1
660 — 670		2
670 — 680		1
680 — 690		1
690 — 700		5
Total		30

12. The weights in grams of 50 apples picked at random from consignment are as follows :
131, 113, 82, 75, 204, 81, 84, 118, 104, 110, 80, 107, 111, 141, 136, 123, 90, 78, 90, 115,
110, 98, 106, 99, 107, 84, 76, 186, 82, 100, 109, 128, 115, 107, 115, 119, 93, 187, 139,
129, 130, 68, 195, 123, 125, 111, 92, 86, 70, 126.

From the grouped frequency table by dividing the variable range into intervals of equal width of 20g.

Sol. Smallest weight = 68 gm

greatest weight = 204 gm

\therefore Range = $204 - 68 = 136$

width of each class interval = 20 g

Now we will form frequency table as given below :

Class interval	Tally marks (f)	Number of apples
60 — 80		5
80 — 100	 	13
100 — 120	 	17
120 — 140	 	10
140 — 160		1
160 — 180	—	0
180 — 200		3
200 — 220		1
Total		50 Ans.

13. The marks obtained by 35 students in an examination are given below : 370, 290, 318, 175, 170, 410, 378, 405, 380, 375, 315, 305, 325, 275, 241, 288, 261, 355, 402, 380, 178, 253, 428, 240, 210, 175, 154, 405, 380, 370, 306, 460, 328, 440, 425.

From a cumulative frequency table with class intervals of length 50.

Sol. Smallest marks = 154

Greatest marks = 460

Range = $460 - 154 = 306$

Now the cumulative frequency table is given below :

Class interval	Tally marks	Frequency	Cumulative frequency
150 — 200		5	5
200 — 250		3	8
250 — 300		5	13
300 — 350	 	6	19
350 — 400	 	8	27
400 — 450	 	7	34
450 — 500		1	35
Total		35 Ans.	

Q.14. Construct the cumulative frequency table from the frequency table given below :

Class-interval	Frequency
0 - 6	7
6 - 12	11
12 - 18	8
18 - 24	14
24 - 30	12

Sol. Below is given cumulative frequency table from the given frequency table :

Class-interval	Frequency	Cumulative frequency
0 - 6	7	7
6 - 12	11	18
12 - 18	8	26
18 - 24	14	40
24 - 30	12	52

Q.15. Construct a frequency distribution table from the following cumulative frequency distribution :

Class-interval	Cumulative Frequency
0 - 8	8
8 - 16	21
16 - 24	26
24 - 32	33
32 - 40	42

Sol. Below is given the cumulative frequency table from the given frequency distribution table :

Class-interval	Cumulative frequency	Frequency
0 - 8	8	8
8 - 16	21	13
16 - 24	26	5
24 - 32	33	7
32 - 40	42	9

Q.16. Construct a frequency table from the following data :

Age (in years)	Number of students
Less than 10	6
Less than 20	14
Less than 30	30
Less than 40	52
Less than 50	65
Less than 60	70

Sol. Below is given the frequency table of the given data :

Class-interval	Cumulative frequency	Frequency
0 - 10	6	6
10 - 20	14	8
20 - 30	30	16
30 - 40	52	22
40 - 50	65	13
50 - 60	70	5

Q.17. Convert the following frequency distribution to exclusive form :

Class-interval	Frequency
30 - 34	7
35 - 39	9
40 - 44	13
45 - 49	6
50 - 54	3
55 - 59	10

Class-interval	Frequency
29.5 - 34.5	7
34.5 - 39.5	9
39.5 - 44.5	13
44.5 - 49.5	6
49.5 - 54.5	3
54.5 - 59.5	10

Use this table to find :

- (i) The true class-limits of the fourth class-interval.
- (ii) The class-boundaries of the fifth class-interval.
- (iii) The class-mark of the third class-interval.
- (iv) The class-size of the sixth class-interval.

Sol. Following is giving a frequency distribution to exclusive form of the given frequency distribution :

- (i) The true class limits of the fourth class interval is $44.5 - 49.5$.
- (ii) The class boundaries of the fifth class interval is $49.5 - 54.5$.
- (iii) The class mark of the third class interval

$$= \frac{39.5 + 44.5}{2} = \frac{84}{2} = 42.$$
- (iv) The class size of the sixth class-interval

$$= 59.5 - 54.5 = 5 \text{ Ans.}$$