Presentation of Data

Frequency: Number of ~~items~~ times an observation occurs in a set of data is called the frequency of that observation.

~~Frequency~~ distribution:

Example: Marks obtained by 20 students of a certain class.

~~54, 32, 38, 44,~~

25, 25, 25, 26, 26, 26, 27, 27, 27, 28, 28, 29, 29
30, 30, 30, 30, 36, 36, 36

Row data

Marks	tally	Frequency
25	///	3
26	///	3
27	///.	3
28	//	2
29	//	2
30	////	4
36	///	3

Grouped data

class interval	Frequency
25-30	2
30-35	5
35-40	9
40-45	10

Frequency distribution:

A frequency distribution is a tabular summary of data where observations are divided into different non-overlapping classes ~~is called the~~ on category.

class interval"
~~inter~~

class interval	Mid point	tally	frequency	cumulative frequency
(15-19)	17	111	3	3
20-24	22	ЖНТ 111	8	11
25-29	27	ЖНТ ЖНТ ЖНТ	16	27
30-34	32	ЖНТ 11	7	34

class: raw data are assigned to some choosen groups of appropriate size; the groups are called class.

class interval: Ordinarily, for numerical data, the frequencies of a ~~class~~ particular class are bounded by two values, the length of the class formed by these two boundary values is known as class interval

class midpoint:

the mid point or mid value of a class is obtained by the averaging the two class limits. For the class-interval 15-19, mid point is 17.

Class-limits:

the smallest value of a class is known as lower class limits and largest value is known as the upper class limit of the interval. For a class interval, 15-19, 15 is the lower class, 19 is the upper class.

* Important steps for constructing a frequency distribution for continous data:

1) Range: ~~Find out range~~. Range is the difference between the largest value and smallest value of the data set.

$$Range = largest\ value - smallest\ value.$$

ii) __Number of class:__ Number of classes should not be too large or too small. As a general rule, the number of classes should range from 5-25. According to M.A struge's, the number of classes can be determined using formula

$$K = 1 + 3.22 \log_{10} n$$

iii) __Width of class:__ Width of the class should be equal as far as possible, the width of the class is ~~obtained~~ by approximately the range divided by the number of classes.

Width of class: $\dfrac{Range}{number \ of \ class}$

iv) __Tally marks:__ Observations are counted and marked by tally marks.

v) __class limits:__ Each class of a frequency distribution has a lower value and a upper value. they are known as lower class limits and a upper class limits.

there are two methods of classifying the data according the class interval, namely i) Exclusive method
ii) Inclusive

a) **Exclusive method:** when class intervals are so fixed that upper limit of one class lower limit of the next class.
In this case, upper limit of each class is excluded from the count in that class.

Example: the following data relate to the audit-time of 20 clients.

10, 15, 20, 28, 13, 18, 24, 29, 12, 16, 23, 34, 14, 17, 22, 17, 21, 16, 18, 19

by Frequency distribution by exclusive method:

Audit time (hours)	number of clients
10-15	4
15-20	8
20-25	5
25-30	2
30-35	1
total	20

b) Inclusive method: Under the inclusive meth the upper limit of one class is included in that class itself.

Frequency distribution for the audit-time 20 clients by inclusive method.

Audit time hours	Number of clients
10-14	4
15-19	8
20-24	5
25-29	2
30-34.	1
~~30-34~~ Total	20

class boundaries: To ensure continuity of classes by inclusive method and get rid of confusion exclusive method, we use class boundaries instead of class limits. The adjustment consists of finding the difference between the lower limits of one class with upper limit of the previous class, dividing the difference by two, substracting the value so obtained from all the lower limits and adding to all upper limits

this can be done by the formula as follows

connection term = $\dfrac{\text{lower limit of one class} - \text{upper limit of previous class}}{2}$

$= \dfrac{d}{2}$

For example, the class 15-19, has lower limit 15 and the previous class 10-14 has upper limit 14.

the connection term is $= \dfrac{15-14}{2} = .5$

To get class boundaries, we have to deduct .5 from the lower limit of all classes and add up .5 to all the upper limits. then the frequency distribution of inclusive method will take the following form.

audit time (hours)	Number of clients
9.5 - 14.5	9
14.5 - 19.5	8
19.5 - 24.5	5
24.5 - 29.5	2
29.5 - 34.5	1
total	20

Example:

Audit time of 50 clients

36, 48, 50, 45, 49, 31, 50, 48, 43, 42,
37, 40, 39, 41, 47, 39, 45, 43, 47, 38,
39, 37, 40, 32, 52, 56, 31, 54, 36, 51,
46, 41, 55, 58, 31, 42, 53, 32, 44, 53,
36, 60, 59, 41, 53, 58, 36, 38, 40, 65

i) Construct frequency distribution
   by Exclusive and inclusive method.

Soln:

number of observation, $n = 50$

Range =
here, Highest value = 65
      lowest   "   = 31

$\therefore$ Range = 65 - 31 = 34

According to Struge's formula, number
   of classes is

$$K = 1 + 3.22 \log_{10} 50 = 6.64$$
$$\approx 7$$

Width of the class = $\dfrac{R}{K} = \dfrac{Range}{number\ of\ classes}$

$$= \dfrac{34}{7} = 4.857 = 5$$

By Exclusive method: Frequency distribution table

class interval	tally	frequency
31 - 36	ⅢⅠ	5
36 - 41	ⅢⅠ ⅢⅠ ⅠⅠⅠⅠ	14
41 - 46	ⅢⅠ ⅢⅠ	10
46 - 51	ⅢⅠ ⅠⅠⅠ	8
51 - 56	ⅢⅠ ⅠⅠ	7
56 - 61	ⅢⅠ	5
61 - 66	Ⅰ	1

By Inclusive method: Frequency distribution

class interval	tally	frequency
31 - 35	ⅢⅠ	5
36 - 40	ⅢⅠ ⅢⅠ ⅠⅠⅠⅠ	14
41 - 45	ⅢⅠ ⅢⅠ	10
46 - 50	ⅢⅠ ⅠⅠⅠ	8
51 - 55	ⅢⅠ ⅠⅠ	7
56 - 60	ⅢⅠ	5
61 - 65	Ⅰ	1