INTRODUCTION TO STATISTICS

A Data

B Mean
(Average)

C Median
(Middle Value)

D Modus
(Frequently occurring values)

Data

Statistics deals with collecting, sorting, and analyzing data. Data consists of a collection of recorded observations or values.

Data can be written in the form of single data and group data. Group data writing will help when the amount of data is very large.

The number of events in which a particular value occurs is called the frequency (f).

Example of Single Data form

18	17	21	18	19	17	18
19	19	16	17	15	19	17
17	18	19	19	18	19	18
18	15	18	17	20	18	16

Example of Group Data form

Value	f
45 – 50	4
51 – 56	4
57 – 62	8
63 – 68	30
69 – 74	31
75 – 80	20

Mean

The mean or average $(\bar{x})$ of x observations of n is as follows.

$$mean = \frac{number\ of\ observations}{numerous\ observers}$$

or

$$\bar{x} = \frac{\sum x}{n}$$

When calculating the mean of a frequency distribution, the formula used is as follows:

$$mean = \bar{x} = \frac{\sum x f}{n} = \frac{\sum xf}{\sum f}$$

For grouped data, the formula for the mean is

$$mean = \bar{x} = \frac{\sum x_i f_i}{n} = \frac{\sum x_i f_i}{\sum f_i}$$

Where x_i is midpoint of the i and f_i is the frequency of the i data interval.

EXAMPLE

Single data: 2,2,6,7,7,7,10,13

The mean is (2+2+6+7+7+7+10+13)/8

= 6.75

Value	f	x_i	x_if
45 – 50	4	47,5	190
51 – 56	4	53,5	214
57 – 62	8	59,5	476
63 – 68	30	65,5	1965
69 – 74	31	71,5	2216,5
75 – 80	20	77,5	1550
81 – 86	2	83,5	167
87 – 92	1	89,5	89,5
Jumlah	100		6868

So that

$$\bar{x} = \frac{\sum x_i f_i}{\sum f_i} = \frac{6868}{100} = 68,68$$

EXAMPLE

Value	f
45 – 50	4
51 – 56	4
57 – 62	8
63 – 68	30
69 – 74	31
75 – 80	20
81 – 86	2
87 – 92	1
total	100

The mode is determined based on the class interval with the highest frequency, which is 31, in the range of 69 to 74

$$M = 68,5 + 6\left(\frac{31 - 30}{(31 - 30) + (31 - 20)}\right)$$

$$= 68,5 + 6\left(\frac{1}{12}\right) = 69$$

Modus

The mode of a dataset is the value that appears most frequently

example :in a dataset

2,2,6,7,7,7,10,13

The mode is 7

The formula for the mode of grouped data :

$$modus = L + c\left(\frac{l}{l + u}\right)$$

L: lower class boundary of the mode class

c: class interval

l: difference between the frequency of the mode class and the one before it

u: difference between the frequency of the mode class and the one after it

Median

The median is the middle value of a dataset when the data is arranged in ascending or descending order.

Example :
The median or middle value of the data 4, 7, 8, **9,** 12, 15, 26 is 9.

When the number of data points is even, the median of the data can be found by adding the two middle terms and then dividing the result by 2.

$$median = Me = \frac{x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n}{2}+1\right)}}{2}$$

The formula for the median of grouped data is:

$$Me = b + p\left(\frac{\frac{1}{2}n - F}{f}\right)$$

n = banyak data
b = batas bawah kelas median (batas bawah -0,5)
f = frekuensi kelas median

p = panjang kelas
F = jumlah frekuensi kelas sebelum kelas median

Median   EXAMPLE (group data)

Value	f	f sum
45 – 50	4	4
51 – 56	4	8
57 – 62	8	16
63 – 68	30	46
69 – 74	31	77
75 – 80	20	97
81 – 86	2	99
87 – 92	1	100
Total	100	

The location of the median is 1/2 of all data = $n/2$ or $1/2 \times 100 = 50$ (see cumulative frequency). So the median will lie in the fifth interval class.
So that,

$$Me = 68,5 + 6\left(\frac{50 - 46}{31}\right)$$
$$= 68,5 + 0,774 = 69,27$$

$$Me = Q_2 = \begin{cases} x_{\frac{n+1}{2}}, jika\ n\ ganjil \\ \dfrac{x_{\frac{n}{2}} + x_{\frac{n}{2}+1}}{2}, jika\ n\ genap \end{cases}$$

Contoh :

Data Ganjil : 3, 4, 5, 6, 7, 8, 9
Median = 6

Data N Genap : 3, 4, 5, 6, 8, 9
Median = $\frac{5+6}{2}$ = 5.5