Measures of central tendency

After the process of classification and tabulation the next important objective of statistical analysis to determine various numerical measures which measures inherent characteristics of the data. It can be achieved through the statistical techniques such as measures of central tendency and dispersion.

The concentration of values around central value of a distribution is known as *central tendency*.

The different measures that are used to study the characteristics of the distribution are known as *measures of central tendency*.

The following are some common measures of central tendency.

1. Arithmetic Mean (AM) – Simple and Weighted
2. Geometric Mean (GM)
3. Harmonic Mean (HM)
4. Median (M)
5. Mode (Z)

Arithmetic Mean (AM) is also known as simple average. It is denoted by x. x can be calculated by dividing the sum (total) of all observations by number of observations.

For ungrouped data

To calculate the arithmetic mean of a set of data we must first add up (sum) all of the data values (x) and then divide the result by the number of values (n). Since Σ is the symbol used to indicate that values are to be summed .

$$x^{-} = \frac{\sum x}{n}$$

Example: Find the mean of: 6, 8, 11, 5, 2, 9, 7, 8

$$x^{-} = \frac{\sum x}{n} = \frac{6+8+11+5+2+9+7+8}{8} = 7$$

For grouped data

The mean of a frequency distribution for a sample is approximated by :

$$x = \frac{\sum(x.f)}{n}, \qquad Note\ that\ n = \sum f$$

Where x and f are the midpoints and frequencies of the classes.

Example:

The following frequency distribution represents the ages of 30 students in a statistics class. Find the mean of the frequency distribution.

Class	x	f	$(x \cdot f)$
18 – 25	21.5	13	279.5
26 – 33	29.5	8	236.0
34 – 41	37.5	4	150.0
42 – 49	45.5	3	136.5
50 – 57	53.5	2	107.0
		$n = 30$	$\Sigma = 909.0$

$$x = \frac{\sum(x.f)}{n} = \frac{909}{30} = 30.3$$

The mean age of the students is 30.3 years.

Median

The median value of a set of data is the middle value of the ordered data

For ungrouped data

Step 1 Arrange the observed values of variable in a data in increasing order.

Step 2 If the total number of items n is an odd number, then the number on the $\frac{n+1}{2}$ position is the median; If n is an even number, then the average of the two numbers on the $n/2$ and $(n/2+1)$ positions is the median. (For ordinal level of data, choose any one on the two middle positions).

Example: find the median of the following:

A) 11,4,9,7,10,5,6

Ordering the data gives : 4,5,6,7,9,10,11

n=7 (Odd number)

The median position is (n+1)/2

The Median Position = (7+1)/2= 4

4,5,6, $\boxed{7}$,9,10,11

Median=7

B) 11,4,9,9,7,10,5,6

Ordering the data gives: 4,5,6,7,9,9,10,11

n=8(Even number)

The median position is n/2 , n/2+1

Median Position : 8/2=4 , (8/2) +1=5

4,5,6, $\boxed{7,9}$,9,10,11

Here there is a middle pair 7 and 9. The median is between these 2 values i.e. the mean of them

Median= (7+9)/2=8

>>

For groped data

Step 1: Construct the cumulative frequency distribution.

Step 2: Decide the class that contain the median. Class Median is the first class with the value of cumulative frequency greater than or equal to n/2.

Step 3: Find the median by using the following formula:

$$\text{Median } (M_m) = l + \left(\frac{(\frac{n}{2} - cf}{f}\right) h$$

L=lower limit of median class

n=is the sum of all frequency values

cf=the cumulative frequency before class median

f=the frequency of the class median

h=class size

Example : Based on the grouped data below, find the median:

Step 1 : Construct the cumulative frequency

Class Limit	**Frequency**	**Cumulative frequency**
135-	6	6
140-	10	16
145-	18	34
150-	22	56
155-	20	76
160-	15	91
165-	6	97
170-175	3	100
Total	100	

Median class

Step 2 : n/2=100/2=50

Class median is the 4rd class

So, cf=34, f=22, L=150 and h=5

Median (M_m)$= l + \left(\frac{(\frac{n}{2} - cf}{f}\right) h$

Median=150+ ((50-34)/22)* 5

=150 + 3.64

=153.64