Measures of Central Tendency and Location

Measures of Central Tendency

- ✓ a single value that represents a data set.
- ✓ Its purpose is to locate the center of a data set.
- ✓ commonly referred to as an **average.**

Properties of Mean

- ✓ A set of data has only one mean
- ✓ Applied for interval and ratio data
- ✓ All values in the data set are included
- ✓ Very useful in comparing two or more data sets.
- ✓ Affected by the extreme small or large values on a data set
- ✓ Cannot be computed for the data in a frequency distribution with an open-ended class

Arithmetic Mean (Mean)

- ✓ The only common measure in which **all values plays an equal role** meaning to determine its values you would need to consider all the values of any given data set.

$\bar{X}$ X bar (for sample)

μ mu (for population)

Mean for Ungrouped Data

$$\text{Mean} = \frac{\text{Sum of all values}}{\text{Number of values}}$$

Sample Mean:

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

Population Mean:

$$\mu = \frac{\sum X}{N}$$

Example 1: Mean for Ungrouped Data

The daily rates of a sample of eight employees at GMS Inc. are 550, 420, 560, 500, 700, 670, 860, 480. Find the mean daily rate of employee.

Solution:

$$\overline{X} = \frac{\sum X}{n} = \frac{X_1 + X_2 + X_3 + \cdots + X_n}{n}$$

$$= \frac{550 + 420 + 560 + 500 + 700 + 670 + 860 + 480}{8} = \frac{4,740}{8} = 592.50$$

The sample mean daily salary of employees is 592.50

Example 2: Mean for Ungrouped Data

Find the population mean of the ages of 9 middle-management employees of a certain company. The ages are 53, 45, 59, 48, 54, 46, 51, 58, and 55.

Solution:

$$\mu = \frac{\sum X}{N} = \frac{X_1 + X_2 + X_3 + \cdots + X_n}{N}$$

$$= \frac{53 + 45 + 59 + 48 + 54 + 46 + 51 + 58 + 55}{9} = \frac{469}{9} = 52.11$$

The mean population age of middle-management employee is 52.11.

Mean for Grouped Data

$$\text{Mean} = \frac{\text{Sum of all values}}{\text{Number of values}}$$

Sample Mean:

$$\bar{X} = \frac{\sum fX}{n}$$

Population Mean:

$$\mu = \frac{\sum fX}{N}$$

Example 3: Mean for Grouped Data

Using the example provided in on SJS Travel Agency. Determine the mean of the frequency distribution on the ages of 50 people taking travel tours.

Class Limits	Frequency
18 – 26	3
27 – 35	5
36 – 44	9
45 – 53	14
54 – 62	11
63 – 71	6
72 – 80	2

Solution for Example 3

Get the midpoint

Get the product of f and X

Get the sum

Class Limits	f	X	fX
18 – 26	3	22	66
27 – 35	5	31	155
36 – 44	9	40	360
45 – 53	14	49	686
54 – 62	11	58	638
63 – 71	6	67	402
72 – 80	2	76	152
Total	50		2,459

$$\overline{X} = \frac{\sum fX}{n} = \frac{2{,}459}{50} = 49.18$$

Thus, the mean age of people taking travel is 49.18.

When arithmetic mean should not be used

In the following cases arithmetic mean should not be used:

- In highly - skewed distributions.
- When the distribution is unevenly spread. Concentration being small or large at irregular points.
- When an average rate of growth or change over a period of time is required.
- When the observation are from geometric progression.
- When averaging rates (that is speed, fluctuations in the prices of articles, etc.)
- When there are very large and very small values of observations.

Other Types of Mean

- ✓ **Weighted Mean**
- ✓ **Geometric Mean**
- ✓ **Combined Mean**

Weighted Mean

- ✓ It is useful when various classes or groups contribute differently to the total.
- ✓ It is found by multiplying each value by its corresponding weight and dividing by the sum of the weights.

$$\bar{X}_w = \frac{\sum_{i=1}^{n} w_i X_i}{\sum_{i=1}^{n} w_i} = \frac{w_1 X_1 + w_2 X_2 + \cdots + w_n X_n}{w_1 + w_2 + \cdots + w_n}$$

w = weight of each entry
--weights may not sum to 100%
--if weights sum to 100%, then $\sum w = 1$

Example for Weighted Mean

Ex 1) A class is graded based on weighted mean as follows:

Homework: 20%
Quizzes: 35%
Tests: 45%

Let's say scores are 95 on homework, 82 on quizzes and 79 on tests. What is the weighted mean?

$$\bar{x} = \frac{\sum(x \bullet w)}{\sum w} = \frac{95(.2) + 82(.35) + 79(.45)}{1} = 83.5$$

What if there were no test scores yet?

$$\bar{x} = \frac{\sum(x \bullet w)}{\sum w} = \frac{95(.2) + 82(.35)}{.55} = 86.7$$

Example for Weighted Mean

Ex 2) Rachel is taking a class in which her grade is determined as follows:

50% from her test mean
15% from her midterm
20% from her final
15% from her homework

Her scores are 86 (tests), 96 (midterm), and 100 (homework). What does Rachel need to get on her final exam to receive a 90% in class?

$$\bar{x} = \frac{\sum(x \bullet \mathrm{w})}{\sum w} = \frac{86(.5) + 96(.15) + 100(.15) + \mathrm{x}(.2)}{1} = 90$$

70.9 + .2x = 90
.2x = 19.1
X = 95.5

Rachel needs to get a 95.5% on the final to earn a 90% in the class.

Example for Weighted Mean

Ex 3) For the month of April, a checking account has a balance of $523 for 24 days, $2415 for 2 days and $250 for 4 days. What is the account's mean daily balance for April?

$$523\left(\frac{24}{30}\right)+2415\left(\frac{2}{30}\right)+250\left(\frac{4}{30}\right)=612.73$$

Geometric Mean

- ✓ To determine the average percents, indexes, and relatives;

$$GM = \sqrt[n]{(X_1)(X_2)(X_3)\cdots(X_n)}$$

- ✓ to establish the average percent increase in production, sales, or other business transaction or economic series from one period of time to another.

$$GM = \sqrt[n-1]{\frac{\text{value at the end of the period}}{\text{value at the start of the period}}} - 1$$

Example 1 for Geometric Mean

Suppose the profits earned by the MSS Construction Company on five projects were 5, 6, 4, 8, and 10 percent, respectively. What is the geometric mean profit?

Solution:

$$GM = \sqrt[n]{(X_1)(X_2)(X_3)(X_4)(X_5)}$$

$$= \sqrt[5]{(5)(6)(4)(8)(10)} = \sqrt[5]{9,600} = 6.26$$

The geometric mean profit is 6.26%.

Example 2 for Geometric Mean

Badminton as a sport grew rapidly in 2008. From January to December 2008 the number of badminton clubs in Metro Manila increased from 20 to 155. Compute the mean monthly percent increase in the number of badminton clubs.

Solution:

$$GM = \sqrt[n-1]{\frac{\text{value at the end of the period}}{\text{value at the start of the period}}} - 1$$

$$= \sqrt[12-1]{\frac{155}{20}} - 1 = \sqrt[11]{7.75} - 1 = 0.2046$$

Hence, badminton clubs are increasing at a rate of almost 0.2046 or 20.46% per month.

When geometric mean should be used

The geometric mean is often used for a set of numbers whose values are meant to be multiplied together or are exponential in nature, such as a set of growth figures: values of the human population or interest rates of a financial investment over time.

Combined Mean

- ✓ The grand mean of all the values in all groups when two or more groups are combined.

Mean Group 1 — Mean Group 2 — Mean Group n

$$\overline{X}_{CM} = \frac{\sum_{i=1}^{n} N_i \overline{X}_i}{\sum_{i=1}^{n} N_i} = \frac{N_1 \overline{X}_1 + N_2 \overline{X}_2 + \cdots + N_n \overline{X}_n}{N_1 + N_2 + \cdots + N_n}$$

Example for Combined Mean

A study comparing the typical household incomes for **3 districts** in the City of Manila was initiated to see where differences in household incomes lie across districts. The mean household incomes for a sample of 45 different families in three districts of Manila are shown in the following table. Calculate a combined mean to obtain the average household income for all **45 families** in the Manila sample.

District 1	District 2	District 3
$\overline{X}_1 = 30,400$ $N_1 = 12$	$\overline{X}_2 = 27,300$ $N_2 = 18$	$\overline{X}_3 = 42,500$ $N_3 = 15$

Example for Combined Mean

Solution:

$$\overline{X}_{CM} = \frac{\sum_{i=1}^{n} N_i \overline{X}_i}{\sum_{1=i}^{n} N_i} = \frac{N_1\overline{X}_1 + N_2\overline{X}_2 + \cdots + N_n\overline{X}_n}{N_1 + N_2 + \cdots + N_n}$$

$$= \frac{30{,}400(12) + 27{,}300(18) + 42{,}500(15)}{12 + 18 + 15}$$

$$= \frac{1{,}493{,}700}{45} = 33{,}193.33$$

Thus, the combined mean in three districts of Manila is 33,193.33 Tk.

Median

- ✓ The midpoint of the data array

  Note: Data Array is a data set arranged in order whether ascending or descending

- ✓ Appropriate measure of central tendency for data that are ordinal or above, but is more valuable in an ordinal type of data.

Properties of Median

- ✓ It is unique, there is only one median for a set of data
- ✓ It is found by arranging the set of data from lowest or highest (or highest to lowest) and getting the value of the middle observation
- ✓ It is not affected by the extreme small or large values.
- ✓ It can be computed for an-open ended frequency distribution
- ✓ It can be applied for ordinal, interval and ratio data

Median for Ungrouped Data

- ✓ If n is odd, the median is the middle ranked
- ✓ If n is even, then the median is the average of the two middle ranked values

$$\text{Median (Rank Value)} = \frac{n+1}{2}\text{th}$$

Example 1: Median for Ungrouped Data

Find the median of the ages of 9 middle-management employees of a certain company. The ages are 53, 45, 59, 48, 54, 46, 51, 58, and 55.

Solution:

Step 1: Arranged the data set in order.

45, 46, 48, 51, 53, 54, 55, 58, 59

Example 1: Median for Ungrouped Data

Step 2: Select the middle rank.

$$\text{Median (Rank Value)} = \frac{n+1}{2} = \frac{9+1}{2} = \frac{10}{2} = 5$$

Step 3: Identify the median in the data set.

45, 46, 48, 51, 53, 54, 55, 58, 59

5^{th}

Hence, the median age is **53 years.**

Example 2: Median for Ungrouped Data

The daily rates of a sample of eight employees at GMS Inc. are 550 Tk, 420 Tk, 560 Tk, 500 Tk, 700 Tk, 670 Tk, 860 Tk, 480 Tk. Find the median daily rate of employee.

Solution:

Step 1: Arranged the data set in order.

420, 480, 500, 550, 560, 670, 700, 860

Step 2: Select the middle rank.

$$\text{Median (Rank Value)} = \frac{n+1}{2} = \frac{8+1}{2} = \frac{9}{2} = 4.5$$

Example 2: Median for Ungrouped Data

Step 3: Identify the median in the data set.

420, 480, 500, 550, 560, 670, 700, 860

4.5^{th}

Get the average of the two values.

$$\text{Median} = \frac{550 + 560}{2} = \frac{1,110}{2} = 555$$

Therefore, the median daily rate is 555 Tk.

Example 3: Median for Grouped Data

Using the example provided in on SJS Travel Agency. Determine the median of the frequency distribution on the ages of 50 people taking travel tours.

Class Limits	Frequency
18 – 26	3
27 – 35	5
36 – 44	9
45 – 53	14
54 – 62	11
63 – 71	6
72 – 80	2

Solution for Example 3

$$\text{Median (Ranked Value)} = \frac{N}{2} = \frac{50}{2} = 25$$

Class Limits	f	cf
18 – 26	3	3
27 – 35	5	8
36 – 44	9	17
45 – 53	14	31
54 – 62	11	42
63 – 71	6	48
72 – 80	2	50
Total	50	

LB = 45 – 0.5
= 44.5

cf

f

Median Class

$$\text{Median} = LB + \left(\frac{\frac{N}{2} - cf}{f}\right)(i) = 44.5 + \left(\frac{\frac{50}{2} - 17}{14}\right)(9) = 49.64$$

Median

Mode

- ✓ The value in a data set that appears most frequently
- ✓ A data may not contain any mode if none of the values is **most typical**.

Unimodal	$\Rightarrow$ With 1 mode
Bimodal	$\Rightarrow$ With 2 modes
Multimodal	$\Rightarrow$ With more than 2 modes
No mode	$\Rightarrow$ Without mode

Properties of Mode

- ✓ It is found by locating the most frequently occurring value
- ✓ the easiest average to compute
- ✓ There can be more than one mode or even no mode in any given data set
- ✓ It is not affected by the extreme small or large values
- ✓ It can be applied for nominal, ordinal, interval and ratio data

Example 1: Mode

The following data represent the total sales for PSP 2000 from a sample of 10 Gaming Centers for the month of August: 15, 17, 10, 12, 13, 10, 14, 10, 8, and 9. Find the mode.

Solution:

The ordered array for these data is

8, 9, 10, 10, 10, 12, 13, 14, 15, 17

Lowest to Highest

Therefore the mode is 10.

Example 2: Mode

An operations manager in charge of a company's manufacturing keeps track of the number of manufactured LCD television in a day. Compute for the following data that represents the number of LCD television manufactured for the past three weeks: 20, 18, 19, 25, 20, 21, 20, 25, 30, 29, 28, 29, 25, 25, 27, 26, 22, and 20. Find the mode of the given data set.

Solution:

The ordered array for these data is
18, 19, 20, 20, 20, 20, 21, 22, 25, 25, 25, 25, 26, 27, 28, 29, 29, 30

There are two modes **20** and **25**.

Example 3: Mode

Find the mode of the ages of 9 middle-management employees of a certain company. The ages are 53, 45, 59, 48, 54, 46, 51, 58, and 55.

Solution:

The ordered array for these data is

45, 46, 48, 51, 53, 54, 55, 58, 59

Therefore is **no mode**.

Example 4: Mode for Grouped Data

Using the example provided in on SJS Travel Agency. Determine the **mode** of the frequency distribution on the ages of 50 people taking travel tours.

Class Limits	Frequency
18 – 26	3
27 – 35	5
36 – 44	9
45 – 53	14
54 – 62	11
63 – 71	6
72 – 80	2

Solution for Example 3

$Highest\ class\ frequency = 14$

Class Limits	f
18 – 26	3
27 – 35	5
36 – 44	9
45 – 53	14
54 – 62	11
63 – 71	6
72 – 80	2
Total	50

LB = 45 – 0.5 = 44.5

f↑ (9), f (14), f↓ (11)

Modal Class: 45 – 53

$\Delta 1 = f - f\uparrow = 14 - 9 = 5$

$\Delta 2 = f - f\downarrow = 14 - 11 = 3$

$$Mode = LB + \left(\frac{\Delta 1}{\Delta 1 + \Delta 2}\right)(i) = 44.5 + \left(\frac{5}{5+3}\right)(9) = 50.125\ or\ 50.13$$

Mode = 50.13

When to use Mode:

Generally speaking mode can be used to describe qualitative data. Mode is particularly useful average for discrete data.

Midrange

- ✓ The average of the lowest and highest value in a data set

$$\text{Midrange} = \frac{X_{\text{lowest}} + X_{\text{highest}}}{2}$$

Properties of Midrange

- ✓ It is easy to compute.
- ✓ It gives the midpoint.
- ✓ It is unique.
- ✓ It is affected by the extreme small or large values.
- ✓ It can be applied for interval and ratio data.

Example for Midrange

Find the midrange of the ages of 9 middle-management employees of a certain company. The ages are 53, 45, 59, 48, 54, 46, 51, 58, and 55.

Solution:

45, 46, 48, 51, 53, 54, 55, 58, 59

$$\text{Midrange} = \frac{X_{lowest} + X_{highest}}{2} = \frac{45+59}{2} = \frac{104}{2} = 52$$

The midrange age is 52.

Effects of Changing Units on Mean & Median

Let X_1, X_2, X_3, ... X_n be n observations. If we added a constant k to each observation or multiplied each observation by a number (h, $h \neq 0$), then

Added Constant

Constant Multiplier

Summary Measure	$Y_i = X_i + k$	$Y_i = h\,X_i$
Mean	Mean(Y) = Mean(X) + k	Mean(Y) = h Mean(X)
Median	Median(Y) = Median(X) + k	Median(Y) = h Median(X)

Example 1 for Effects on Mean & Median

A business professor gave a test to a set of students. The test had 40 questions, each worth 2 points. The summary for the students' scores on the test are as follows: Mean = 70 and Median = 68. After grading the test, the professor realized that, because he had a typographical error in question number 4, no students was able to answer the question. So he decided to adjust the students' scores by adding 2 points to each one. What will be the mean and median for the new adjusted scores?

Example 1 for Effects on Mean and Median

Solution:

Mean (X) = 70, Median (X) = 68, k = 2

Mean (Y) = Mean (X) + k = 70 + 2 = 72

Median (Y) = Median (X) + k = 68 + 2 = 70

The new mean and median are 72 and 70.

Example 2 for Effects on Mean & Median

The summary statistics for the monthly contribution of employees on cooperative are as follows: mean = 1,250 Tk and median = 1,100 Tk. This month, employees voted to increase the contribution by 5 percent (5%), in order to support their new project. What will be the mean and median for the new increased of contributions?

Example 2 for Effects on Mean and Median

Solution:

Mean (X) = 1,250 Tk

h = 105% = 1.05

Median (X) = 1,100 Tk

Mean (Y) = h Mean (X) = 1.05(1,250 Tk) = 1,312.50 Tk

Median (Y) = h Median (X) = 1.05(1,100 Tk) = 1,155 Tk

The new mean is 1,312.50 Tk and the new median is 1,155 Tk.

Quartiles, Deciles & Percentiles

	Ungrouped	Grouped
Quartiles	$Q_k = \frac{k(N+1)}{4}$	$Q_k = LB + \left(\frac{\frac{kN}{4} - cf}{f} \right)(i)$
Deciles	$D_k = \frac{k(N+1)}{10}$	$D_k = LB + \left(\frac{\frac{kN}{10} - cf}{f} \right)(i)$
Percentiles	$P_k = \frac{k(N+1)}{100}$	$P_k = LB + \left(\frac{\frac{kN}{100} - cf}{f} \right)(i)$

Example for Quartiles

Find the 1st, 2nd, and 3rd quartiles of the ages of 9 middle-management employees of a certain company. The ages are 53, 45, 59, 48, 54, 46, 51, 58, and 55.

Solution:

$$Q_1 = \frac{1(N+1)}{4} = \frac{1(9+1)}{4} = \frac{10}{4} = 2.5$$

$$Q_2 = \frac{2(N+1)}{4} = \frac{2(9+1)}{4} = \frac{2(10)}{4} = 5$$

$$Q_3 = \frac{3(N+1)}{4} = \frac{3(9+1)}{4} = \frac{3(10)}{4} = 7.5$$

Example for Quartiles

45, 46, 48, 51, 53, 54, 55, 58, 59

2.5th — 5th — 7.5th

$$Q_1 = \frac{46 + 48}{2} = \frac{94}{2} = 47$$

$$Q_3 = \frac{55 + 58}{2} = \frac{113}{2} = 56.5$$

Therefore, Q_1 = 47, Q_2 = 53, and Q_3 = 56.5.

Example for Q_k, D_k, P_k

Using the example provided in on SJS Travel Agency. Determine the Q_1, Q_2, Q_3, D_7, P_{22} of the frequency distribution on the ages of 50 people taking travel tours.

Class Limits	Frequency
18 – 26	3
27 – 35	5
36 – 44	9
45 – 53	14
54 – 62	11
63 – 71	6
72 – 80	2

Solution for Q_1

$$Q_1(\text{Ranked Value}) = \frac{N}{4} = \frac{50}{4} = 12.5$$

Class Limits	f	cf
18 – 26	3	3
27 – 35	5	8
36 – 44	9	17
45 – 53	14	31
54 – 62	11	42
63 – 71	6	48
72 – 80	2	50
Total	50	

LB = 36 – 0.5 = 35.5

cf

f

Q_1 Class

$$Q_1 = LB + \left(\frac{\frac{N}{4} - cf}{f}\right)(i) = 35.5 + \left(\frac{\frac{50}{4} - 8}{9}\right)(9) = \boxed{40}$$

Solution for Q_2

$$Q_2(\text{Ranked Value}) = \frac{2N}{4} = \frac{2(50)}{4} = 25$$

Class Limits	f	cf
18 – 26	3	3
27 – 35	5	8
36 – 44	9	17
45 – 53	14	31
54 – 62	11	42
63 – 71	6	48
72 – 80	2	50
Total	50	

LB = 45 – 0.5 = 44.5

cf

f

Q_2 Class

$$Q_2 = LB + \left(\frac{\frac{2N}{4} - cf}{f}\right)(i) = 44.5 + \left(\frac{\frac{2(50)}{4} - 17}{14}\right)(9) = \boxed{49.64}$$

Solution for Q_3

$$Q_3(\text{Ranked Value}) = \frac{3N}{4} = \frac{3(50)}{4} = 37.5$$

Class Limits	f	cf
18 – 26	3	3
27 – 35	5	8
36 – 44	9	17
45 – 53	14	31
54 – 62	11	42
63 – 71	6	48
72 – 80	2	50
Total	50	

LB = 54 – 0.5 = 53.5

cf

f

Q_3 Class

$$Q_3 = LB + \left(\frac{\frac{3N}{4} - cf}{f} \right)(i) = 53.5 + \left(\frac{\frac{3(50)}{4} - 31}{11} \right)(9) = 58.82$$

Solution for D7

$$D_7(\text{Ranked Value}) = \frac{7N}{10} = \frac{7(50)}{10} = 35$$

Class Limits	f	cf
18 – 26	3	3
27 – 35	5	8
36 – 44	9	17
45 – 53	14	31
54 – 62	11	42
63 – 71	6	48
72 – 80	2	50
Total	50	

LB = 54 – 0.5 = 53.5

cf

f

D7 Class

$$D_7 = LB + \left(\frac{\frac{7N}{10} - cf}{f} \right)(i) = 53.5 + \left(\frac{\frac{7(50)}{10} - 31}{11} \right)(9) = 56.77$$

Solution for P_{22}

$$P_{22}(\text{Ranked Value}) = \frac{22N}{100} = \frac{22(50)}{100} = 11$$

Class Limits	f	cf
18 – 26	3	3
27 – 35	5	8
36 – 44	9	17
45 – 53	14	31
54 – 62	11	42
63 – 71	6	48
72 – 80	2	50
Total	50	

LB = 36 – 0.5 = 35.5

cf

P_{22} Class

f

$$P_{22} = LB + \left(\frac{\frac{22N}{100} - cf}{f}\right)(i) = 35.5 + \left(\frac{\frac{22(50)}{100} - 8}{9}\right)(9) = 38.5$$

HARMONIC MEAN

The Harmonic Mean (H.M.) is defined as " A value that is the reciprocal of the mean of the reciprocals of a set of numbers". Typically, it is appropriate for situations when the average of rates is desired.

In other words, it is the reciprocal of the arithmetic mean and is calculated by taking reciprocal of the individual observations.

FORMULA :

- For Ungrouped data : H.M of $X = n/\sum(1/x)$
- For Grouped data : H.M of $X = \sum f/\sum(f/x)$

f = frequency of observations x1,x2,x3,.....

ADVANTAGES OF HARMONIC MEAN

- The Harmonic Mean is computed based on every observation in the data set.
- Higher weight age is given to smaller values in a data set since the reciprocal of numerical values is taken for the calculation.
- Certain algebraic changes can be made in the formula of H.M. for further analysis.

DISADVANTAGES OF HARMONIC MEAN

- The H.M. is not often used for analyzing business problems.

- The H.M. cannot be calculated if a data set values has negative or zero elements.

- Since for the calculation of H.M. the higher weight age is given to smaller values in the data set. So, it does not represent the true characteristic of the data set.

When to use Harmonic mean

The Harmonic Mean(H.M.) is useful particularly for computation of average rates and ratios.

Such rates and ratios are generally used to define relations between two different types of measuring units expressed reciprocally.

For example, distance (in Km) and time (in hours).

Whatever exist at all exist in some amount…and whatever exists in some amount can be measured.

– Edward L. Thorndike