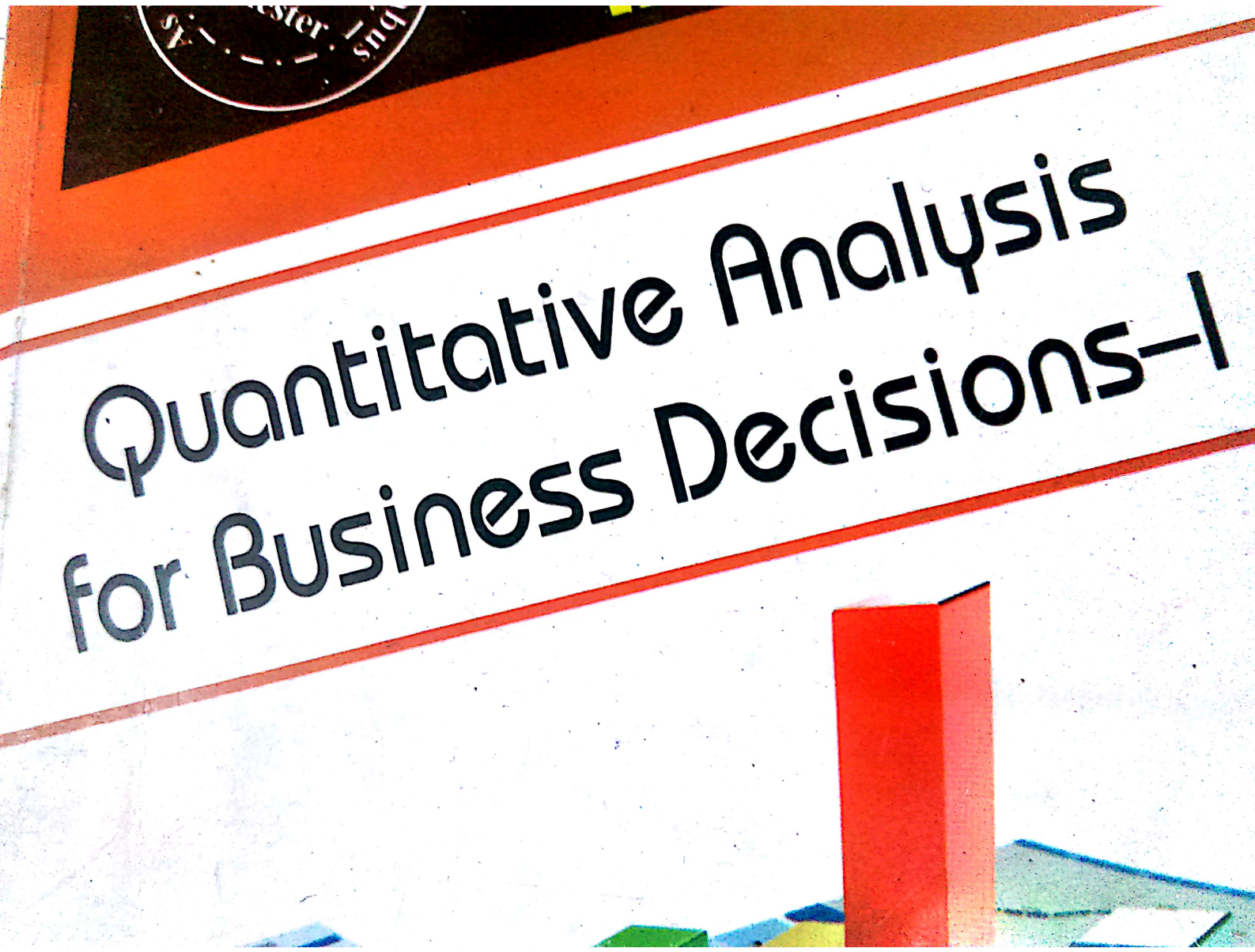
Quantitative Analysis for Business Decisions–I



1. Introduction to Statistics.

2M.

1 What do you mean by Statistics?

A. In plural sense it refers to the description of numerical facts that are presented systematically.

In a singular sense, it refers to the statistical methods and principles used for the classification and analysis of quantitative data so as to arrive at valid conclusions.

2 Give the definition of statistics?

It is the science of collection, organization, Presentation, analysis interpretation of numerical data.

3 What are the objectives of statistics?

1. To make sense from the population or mass
2. To forecast the future trend from the data
3. To prove unknown from the known data.

4. What are the functions of statistics?

1. It presents facts in a definite form.
2. It helps in formulating and testing hypotheses
3. It helps in prediction.
4. It facilitates Comparision.

5 Applications of statistics?

1. Statistics and the state
2. Statistics and Business
3. Statistics and Economics.
4. Statistics and Research.

2. Classification and Tabulation of data.

2 Mk.

1. Define Tabulation?

Tabulation is an orderly arrangement of data in columns and rows systematically in a Tabular form.

2. Mention the parts of table?

1. Table Number.
2. Title of the table
3. Caption.
4. Stub
5. Body
6. Head note
7. Foot note.

3. Format of a Table.

Title	→	Head note
Stub Heading	→	caption Heading - Column Heading
Stub entries	→	Body
Foot notes	→	Table Number.

3 Write the meaning of census?

It involves drawing inferences from the entire body of units of interest

4 Mention the types of tables?

Tables may broadly classified into two categories

1. Simple and Complex tables
2. General purpose and special purpose tables.

5. What do you mean by primary data?
The data which are originally collected for the first time are called primary data.

6. What do you mean by secondary data?
Secondary data are those data which are already collected, processed and used by some one else for their own purpose.

7. Difference between primary data & secondary data.

	primary data		secondary data.
1	It is the collection of orginal data for the first time	1	It is a collection of existing data.
2	It is relatively more costly	2	It is relatively less costly.
3	It may be used directly for the purpose of enquiry	3	It may require certain adjustments to suit the purpose of Enquiry

8. What is frquency distribution.
Frequency distribution is a summary presentation of values of variables according to the magnitude individually or in Groups.

Exclusive method.

1 Form a frequency table with the size of class interval 20 and the first belong 0-20.

5	10	15	8	20	25	28	30	36	40
42	38	50	45	47	46	59	60	70	71
69	68	81	84	75	76	41	85	86	87
90	94	96	99	100	42	45	52	55	82
65	75	95	35	25	24	26	86	72	82.

Frequency distribution table.

C.I	Tally bars	frequency.
0-20	||||	4
20-40	𝍸 𝍸	10
40-60	𝍸 𝍸 ||	12
60-80	𝍸 𝍸	10
80-100	𝍸 𝍸 ||||	14
	Total	50

In a sample study about coffee habit in two towns the following information was received

Town A :- Females were 40%.
Total coffee drinkers were 45%.
male non-coffee drinkers were 20%

Town B :- Males were 55%.
Male Non-coffee drinkers were 30%.
Female coffee drinkers were 15%.

Represent the above data in a tabular form.

	Town-A			Town-B		
Attribute	Male	Female	Total	Male	Female	Total
coffee drinkers	40	5	45	25	15	40
Non-coffee drinkers	20	35	55	30	30	60
Total	60	40	100 *	55	45	100

Measures of Central Tendency.

1 What is central value?

The objectives of statistical analysis is to get one single value that describes the characteristic of the entire mass of unwieldy data.

2 what is the word Average?

Clark → Average is an Attempt to find one single value to describe whole data.

3 Types of Averages?

1. Arithmetic mean 1) Simple 2) weighted.
2. Median
3. mode
4. Geometric mean
5. Harmonic mean

4 what is Arithmetic mean?

The value obtained by adding together all the items and by dividing this total by the number of items is called Arithmetic mean.

5. what is Median?

The median by definition is the middle value in a distribution.

6 what is mode?

The value which occurs most often in the data or The series of observations which occurs with the greatest frequency

Find the value of mean, median and mode

1. Marks below 10 20 30 40 50 60 70 80
No. of students 25 40 60 75 95 125 190 240

CI Marks	c.f No. of stud. No of students	f	x midvalue	fx
below 10	25	25	5	125
10-20	40	15	15	225
20-30	60	20	25	500
30-40	75	15	35	525
40-50	95 m	20	45	900
50-60	125 ←	(30)	55	1650
60-70	190	65	65	4225
70-80	240	50	75	3750
		$\Sigma f = 240$		$\Sigma fx = 11,900$

$$\text{Mean} = \bar{x} = \frac{\Sigma fx}{\Sigma f} = \frac{11,900}{240} = 49.58 \approx 50.$$

$$\text{median} = L + \frac{\left(\frac{N}{2} - mf\right)}{f} \times C \qquad \frac{N}{2} = \frac{240}{2} = 120$$

$$= 50 + \left(\frac{120 - 95}{30}\right) \times 10$$

$$= 50 + \left(\frac{25}{30}\right) \times 10 = 50 + 8.33 = 58.33$$

$$\text{Mode} = 3(\text{Median}) - 2(\text{mean})$$
$$= 3(58.33) - 2(50)$$
$$= 174.99 - 100$$
$$= 74.99$$
$$= 75$$

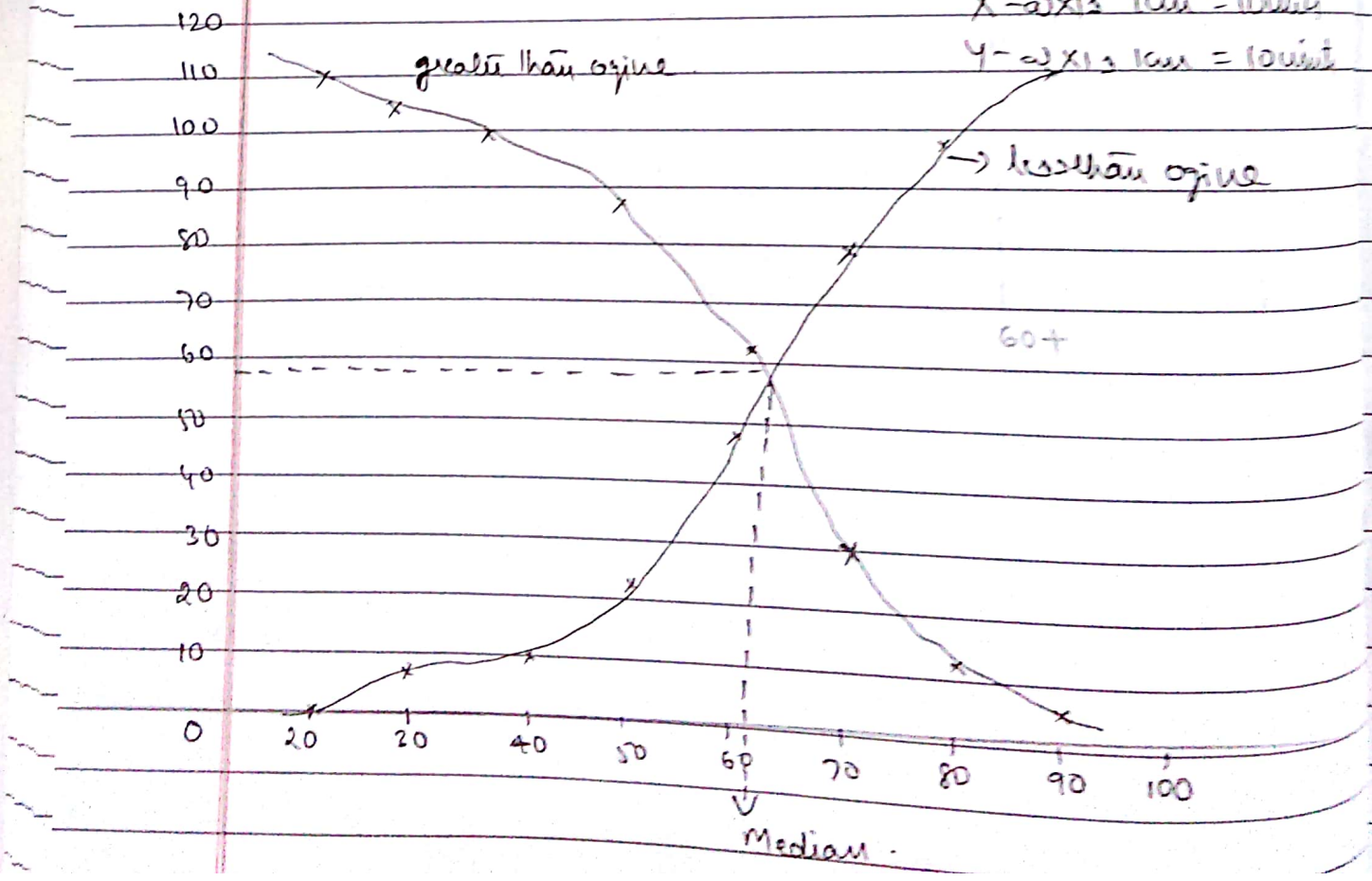
Problem on ogive curves

Draw less than and greater than ogives for the following data.

C.I	20-30	30-40	40-50	50-60	60-70	70-80	80-90	90-100
f	4	6	13	25	32	19	8	3

Sol

C.I	less than ogive	more than greater than ogive
20	0	110
30	4	106
40	10	100
50	23	87
60	48	62
70	80	30
80	99	11
90	107	3
100	110	0



UNIT-4

Measures of Dispersion and Skewness

1 What is dispersion?

Dispersion is the extent to which the magnitudes or quantities of the items differ
that is degree of diversity.

Dispersion is the measure of the variation of the items.

2 Write the properties of good measure of Dispersion.

1. It should be simple to understand
2. It should be easy to compute
3 It should be having sampling stability.

3. Difference between central tendency & Dispersion.

central tendency	Dispersion.
1. It is representative of whole data	1. It gives variable data.
2 Measures include 1) Mean - Calculate Average 2) Mode - Most frequency 3) Median - Middle value.	2. Measures include 1. Range 2. Mean deviation 3. Standard deviation 4. Quartile deviation

4 Define Range.

Range is the difference between the highest and lowest value in a series.

Symbolically $R = L - S$.

5. Coefficient of Range :-

$$C.R = \frac{\text{Largest Value} - \text{smallest value}}{\text{largest value} + \text{smallest value}}$$

6 What do you mean by Quartile deviation

Q.D. gives the average amount by which the two Quartiles differs from the median.

individual series.

Find all quartiles for the following distribution
Variables 2, 1, 6, 5, 3, 4.

Sol A-O → 1, 2, 3, 4, 5, 6.

$$Q_1 = \frac{n+1}{4} = \frac{6+1}{4} = \frac{7}{4} = 1.75$$

Sol

Q_1 = 1st term + 0.75 (2nd term − 1st term)

= 1 + 0.75 (2 − 1)
= 1 + 0.75 (1)
= 1.75.

$$Q_2 = \frac{n+1}{2} = \frac{6+1}{2} = \frac{7}{2} = 3.5^{th} \text{ term}$$

Q_2 = 3rd term + 0.5 (4th − 3rd)

= 3 + (0.5) (4 − 3)
= 3 + (0.5) (1)
= 3 + 0.5 = 3.5

$$Q_3 = \frac{3(n+1)}{4} = \frac{3(6+1)}{4} = \frac{3(7)}{4} = \frac{21}{4} = 5.25$$

Q_3 = 5th term + (0.25) (6th − 5th)

= 5 + (0.25) (6 − 5)
= 5 + (0.25) (1)
= 5 + 0.25
= 5.25.

Find all quartiles. Discrete series.

Marks :-	10	20	30	40	50	60	70
No. of students :-	2	8	4	3	2	1	0

Sol.

Marks	No. of students	Cummulative frequency
10	2	2
(20)	8	10 ←
30	4	14
40	3	17 ←
50	2	19
60	1	20
70	0	20
	N = 20	

$$Q_1 = \frac{N+1}{4} = \frac{20+1}{4} = \frac{21}{4} = 5.25^{th} \text{ item.}$$

$$Q_1 = 20.$$

$$Q_2 = \frac{N+1}{2} = \frac{20+1}{2} = \frac{21}{2} = 10.5^{th} \text{ item.}$$

$$Q_2 = 30$$

$$Q_3 = \frac{3(N+1)}{4} = \frac{3(20+1)}{4} = \frac{3(21)}{4} = \frac{63}{4} = 15.75$$

$$Q_3 = 40.$$

Continuous series

1. find all Quartiles following distribution.

C.I	0-10	10-20	20-30	30-40	40-50
f	2	4	3	1	6

Sol.

C.I	No. of persons	C.f
0-10	2	2
10-20	4 f	2+4 = 6 ←
20-30	3	6+3 = 9 *
30-40	1	9+1 = 10 ←
40-50	6	10+6 = 16 *
	N = 16	

$Q_1 = \frac{N}{4} = \frac{16}{4} = 4^{th}$ item

$Q_1 = L + \left(\frac{\frac{N}{4} - Cf}{f}\right) \times C$

$= 10 + \left(\frac{4-2}{4}\right) \times 10$

$= 10 + \left(\frac{2}{4}\right) \times 10$

$= 10 + 5 = 15.$

L = 10
f = 4
$\frac{N}{4} = 4$
Cf = 2
C = 10

$Q_2 = \frac{N}{2} = \frac{16}{2} = 8^{th}$ item.

$Q_2 = L + \left(\frac{\frac{N}{2} - Cf}{f}\right) \times C$

$= 20 + \left(\frac{8-6}{3}\right) \times 10$

$= 20 + \left(\frac{2}{3}\right) \times 10$

$= 20 + 6.67$

$= 26.67$

L = 20
f = 3
$\frac{N}{2} = 8$
Cf = 6
C = 10

$Q_3 = \frac{3N}{4} = \frac{3 \times 16}{4} = 12^{th}$ item.

$L = 40$

$f = 6$

$\frac{3N}{4} = 12$

$Cf = 10$

$C = 10.$

$$Q_3 = L + \left(\frac{\frac{3N}{4} - Cf}{f} \right) \times C$$

$$= 40 + \left(\frac{12 - 10}{6} \right) \times 10$$

$$= 40 + \left(\frac{2}{6} \right) \times 10$$

$$= 40 + 3.33$$

$$= 43.33.$$

Coefficient of Q.D $= \frac{Q_3 - Q_1}{Q_3 + Q_1}$

Calculate M.D. from the Median.

C.I	10-20	20-30	30-40	40-50	50-60	60-70	70-80	80-90	90-100
f	5	10	16	20	25	20	18	12	6.

Income	f	X	C.f	$\|d\| = X - M$	f.|d|
10-20	5	15	5	15-56 = 41	205
20-30	10	25	15	25-56 = 31	310
30-40	16	35	31	35-56 = 21	336
40-50	20	45	51	45-56 = 11	220
50-60	25	55	76 ←	55-56 = 1	25
60-70	20	65	96	65-56 = 9	180
70-80	18	75	114	75-56 = 19	342
80-90	12	85	126	85-56 = 29	348
90-100	6	95	132	95-56 = 39	234
	$\Sigma f = 132$				$\Sigma f\|d\| = 2200$

$$\text{Median} = \frac{N}{2} = \frac{132}{2} = 66$$

L = 50
f = 25
m = 51
C = 10

$$\text{Median } M = L + \frac{\left(\frac{N}{2} - m\right)}{f} \times C$$

$$= 50 + \left(\frac{66 - 51}{25}\right) \times 10$$

$$= 50 + 6 = 56.$$

$$M.D = \frac{\Sigma f|d|}{\Sigma f} = \frac{2200}{132} = 16.66$$

$$\text{Coeffi Mean Deviation} = \frac{M.D}{\text{Median}} = \frac{16.66}{56} = 0.29.$$

Unit of supplier X

Tyres of supplier 'X' are more uniform as their variation is less than

Type I:- Tyre dealer receives a sample of tyre from two suppliers X and Y. He had the sample tested for length of life which supplier tyre are more uniform regarding their length of life.

length of tyre	supplier X	supplier Y
4-8	10	2
8-12	16	42
12-16	30	12
16-20	4	4

Supplier X :-

C.I	f	x	f·x	$d = x - \bar{x}$	d^2	fd^2
4-8	10	6	60	-5.87	34.46	344.6
8-12	16	10	160	-1.87	3.50	56
12-16	30	14	420	2.13	4.54	136.2
16-20	4	18	72	6.13	37.58	150.32
	N=60		712			687.12

$$\bar{x} = \frac{\Sigma fx}{N} = \frac{712}{60} = 11.87$$

$$S.D = \sigma = \sqrt{\frac{\Sigma fd^2}{N}} = \sqrt{\frac{687.12}{60}} = \sqrt{11.45} = 3.38$$

$$C.V = \frac{\sigma}{\bar{x}} \times 100 = \frac{3.38}{11.87} \times 100 = 28.48.$$

Supplier Y

C.I	f	x	f·x	$d = x - \bar{x}$	d^2	$f \cdot d^2$
4-8	2	6	12	-5.2	27.04	54.08
8-12	42	10	420	-1.2	1.44	60.48
12-16	12	14	168	2.8	7.84	94.08
16-20	4	18	72	6.8	46.24	184.96
	N=60		672			393.6

$$\bar{x} = \frac{\Sigma fx}{N} = \frac{672}{60} = 11.2$$

$$S.D = \sigma = \sqrt{\frac{\Sigma fd^2}{N}} = \sqrt{\frac{393.6}{60}} = \sqrt{6.56} = 2.56$$

$$C.V = \frac{\sigma}{\bar{x}} \times 100 = \frac{2.56}{11.2} \times 100 = 22.86$$

Supplier X is Variable; Supplier Y is Constant

UNIT- 4

Measures and Dispersion and Skewness

1 What is skewness?

Skewness is a measure of the asymmetry of the probability distribution of a real valued random variable.

2 What is positive skewness?

The right tail is longer.

The mass of distribution is concentrated on the left

or

The value of mean is maximum and mode least. median is lies between the two

3 What is negative skewness?

The mode is maximum and mean is least. median lies between two.

4. A Distribution is said to be skewed if

1. mean ≠ median ≠ mode
2. quartiles are not equidistant from median.

UNIT- 5

Index Numbers.

1. What do you mean by Index Number?

An Index Number is a statistical measure designed to show changes in a variable or a group of related variables with respect to time, geographic location or other characteristics such as income, profession. etc.,

2. State the different type of Index Numbers.

1. Price Index
2. Quantity Index
3. Value Index
4. Special purpose index

3. What do you mean by Consumer price Index

Consumer price Index also called the cost of living Index. It represents average change over time in the prices paid by the ultimate Consumer of a specified basket of goods and services.

4. What is a Current year?

The Current year refers the year for which Comparison is done.
It is the year other than base year.

5. What is a base year?

The base year is a year which Index Numbers are to be Compared & to which they are to be referred. It should be a period which is having normal or stable economic activities. The year shall not be too far from the Current year.

Weighted Aggregative

Fisher's Ideal Index.

Construct with the help of the table below.
Fisher's Ideal Index.

Commodity	2007 Price	2007 Quantity	2011 Price	2011 Quantity
A	6	50	10	56
B	2	100	2	120
C	4	60	6	60
D	10	30	12	24
E	8	40	12	36

Sol

Commodities	Base year P_0	Base year q_0	current year P_1	current year q_1	P_0q_0	P_1q_0	P_1q_1	P_0q_1
A	6	50	10	56	300	500	560	336
B	2	100	2	120	200	200	240	240
C	4	60	6	60	240	360	360	240
D	10	30	12	24	300	360	288	240
E	8	40	12	36	320	480	432	288
					1360	1900	1880	1344.

Fisher's Ideal Index

$$P_{01} = \sqrt{\frac{\Sigma P_1 q_0}{\Sigma P_0 q_0} \times \frac{\Sigma P_1 q_1}{\Sigma P_0 q_1}} \times 100$$

$$= \sqrt{\frac{1900}{1360} \times \frac{1880}{1344}} \times 100$$

$$= \sqrt{1.954} \times 100$$

$$= 1.398 \times 100$$

$$= 139.8$$

Consumer price Index

Aggregate Expenditure method.

Commodity	Quantity Consumed	Price in 1993	Price in 1996
A	100	8	12
B	25	6	7
C	10	5	8
D	20	15	18

Sol.

Commodities	Quantity q_0	Price p_0	Price p_1	p_0q_0	p_1q_0
A	100	8	12	800	1200
B	25	6	7	150	175
C	10	5	8	50	80
D	20	15	18	300	360
				$\Sigma p_0q_0 = 1300$	$\Sigma p_1q_0 = 1815$

Consumer price index by

Aggregate Expenditure method

$$= \frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times 100$$

$$= \frac{1815}{1300} \times 100$$

$$= 139.6$$

Family Budget method or

method of weighted Relatives.

#	Commodity	Quantity	Price in 2009	Price in 2012
	A	100	8	12
	B	25	6	7
	C	90	5	8
	D	20	15	18

# Sol	Commodity	q_0 Quantity	Price P_0	Price P_1	Price Relative $P = \frac{P_1}{P_0} \times 100$	weights $W = P_0 q_0$	PxW
	A	100	8	12	150.0	800	120000
	B	25	6	7	116.6	150	17490
	C	10	5	8	160.0	50	8000
	D	20	15	18	120.0	300	36000
					546.6	$\Sigma W = 1300$	181490

$$\text{Consumer price Index} = \frac{\Sigma PW}{\Sigma W}$$

$$= \frac{181490}{1300}$$

$$= 139.6.$$

Calculate Karl-pearsons coefficient of skewness.

marks more than : 5 15 25 35 45 55 65 75 85

No. of students : 120 105 96 85 72 58 42 12 0

marks C.I	No. of student f	midpoint x	$d=\frac{x-A}{c}$	fd	d^2	fd^2	cf
5-15	15	10	-4	-60	16	240	15
15-25	9	20	-3	-27	9	81	24
25-35	11	30	-2	-22	4	44	35
35-45	13	40	-1	-13	1	13	48
45-55	(14)	50 ✓	0	0	0	0	62 ←
55-65	16	60	1	16	1	16	78
65-75	30	70	2	60	4	120	108
75-85	12	80	3	36	9	108	120
	$N=120$			$\Sigma fd=-10$		$\Sigma fd^2=622$	

Assumed mean $A = 50$.

Mean $\bar{x} = A + \frac{\Sigma fd}{\Sigma f} \times C$

$= 50 + \left(\frac{-10}{120}\right) \times 10 = 49.17$

median $= L + \frac{\frac{N}{2} - m}{f} \times C$

$= 45 + \left(\frac{60-48}{14}\right) \times 10$

$= 45 + 8.57 = 53.57.$

$\frac{N}{2} = \frac{120}{2} = 60$

$L = 45$

$m = 48$

$f = 14$

$C = 10$

S.D $\sigma = \sqrt{\frac{\Sigma fd^2}{N} - \left(\frac{\Sigma fd}{N}\right)^2} \times C$

$= \sqrt{\frac{622}{120} - \left(\frac{-10}{120}\right)^2} \times 10 = 22.75$

Karl-pearson's Skewness.

C.I	f	x	f·x	$d = x - \bar{x}$	d^2	$f \cdot d^2$
5-15	15	10	150	-40	1600	24000
15-25	9	20	180	-30	900	8100
25-35	11	30	330	-20	400	4400
35-45	13	40	520	-10	100	1300
45-55	14	50	700	0	0	0
55-65	16 f_0	60	960	+10	100	1600
← 65-75	30 f_1	70	2100	20	400	12000
75-85	12 f_2	80	960	30	900	10800
	N=120.		5900			62,200

$$\bar{x} = \frac{\Sigma fx}{\Sigma f} = \frac{5900}{120} = 49.16 = 49 = 50$$

$$\sigma = \sqrt{\frac{\Sigma fd^2}{\Sigma f}} = \sqrt{\frac{62,200}{120}} = \sqrt{518.33} = 22.7$$

$$Z = L + \left(\frac{f_1 - f_0}{2f_1 - f_0 - f_2}\right) \times C$$

$$= 65 + \left(\frac{30-16}{60-16-12}\right) \times 10$$

$$= 65 + \left(\frac{14}{32}\right) \times 10$$

$$= 65 + \left(\frac{140}{32}\right) = 65 + 4.375 = 69.375$$

$$S_k = \frac{\bar{x} - Z}{\sigma} = \frac{50 - 69.375}{22.7} = \frac{-19.375}{22.7} = -0.8535$$

Unweighted Index method.

1. Simple Aggregate method.

1. From the following data construct an Index for 2009 taking 2008 as base.

Commodity	Price 2008 (Rs)	Price in 2009 Rs
A	50	70
B	40	60
C	80	90
D	110	120
E	20	20

Sol

Commodity	P_0 Price 2008	P_1 Price 2009
A	50	70
B	40	60
C	80	90
D	110	120
E	20	20
	$\Sigma P_0 = 300$	$\Sigma P_1 = 360$

$$P_{01} = \frac{\Sigma P_1}{\Sigma P_0} \times 100$$

$$= \frac{360}{300} \times 100$$

$$= 120.$$

This means that as compared to 2008, in 2009 there is a net increase in the price of commodities included in the index to the extent of 20%.

Simple average of price Relative method

From the following data Construct an index for 2009 taking 2008 as base by the average of relative method using (a) arithmetic mean (b) Geometric mean

Commodity	price in 2008	price in 2009
A	50	70
B	40	60
C	80	90
D	110	120
E	20	20

A.M. –

Commodity	P_0 Price 2008	P_1 Price 2009	$\frac{P_1}{P_0} \times 100$ Price Relative
A	50	70	140.0
B	40	60	150.0
C	80	90	112.5
D	110	120	109.1
E	20	20	100.0
			$\Sigma \frac{P_1}{P_0} \times 100 = 611.6$

$N = 5$

$$P_{01} = \frac{\Sigma \frac{P_1}{P_0} \times 100}{N} = 122.32.$$

G.M. –

Commodity	P_0 Price 2008	P_1 Price 2009	P Price Relative	$\log P$
A	50	70	140.0	2.1461
B	40	60	150.0	2.1761
C	80	90	112.5	2.0512
D	110	120	109.1	2.0378
E	20	20	100.0	2.0000
				$\Sigma \log P = 10.4112$

$$P_{01} = \text{Antilog}\left[\frac{\Sigma \log P}{N}\right]$$

$$= \text{Antilog}\left[\frac{10.4112}{5}\right] = \text{Antilog}\,[2.0822]$$

$$= 120.9$$