

STA304
Applied Statistics
Handouts

Lecture: 01

Content:

- ▶ Index Number
- ▶ Selection of Base Period
- ▶ Fixed Base Method
- ▶ Chain Base Method
- ▶ Selection of Average

Index Number:

An index number is a statistical measure of average change in a variable or group of variables with respect to time or space.

The classical definition of an index number is provided by the English economist F.Y. Edge worth (1845-1926) who states that “An index number is a quantity which shows by its variations, the changes over time or space of a magnitude which is not subject either of accurate measurement in itself or of direct valuation in practice”.

Example

Suppose that a cup of coffee in a particular café cost Rs. 35 in 2002. In 2012, an identical cup of coffee cost Rs. 50. How has the price changed from 2002 to 2012?

Construction of index number:

The usual method of compilation of an index number of wholesale prices involves the following steps:

- ▶ Selection of commodities to be included, their number and price quotations.
- ▶ Selection of the base period and calculation of price relatives.
- ▶ Selection of average to be used.
- ▶ Selection of appropriate weights.

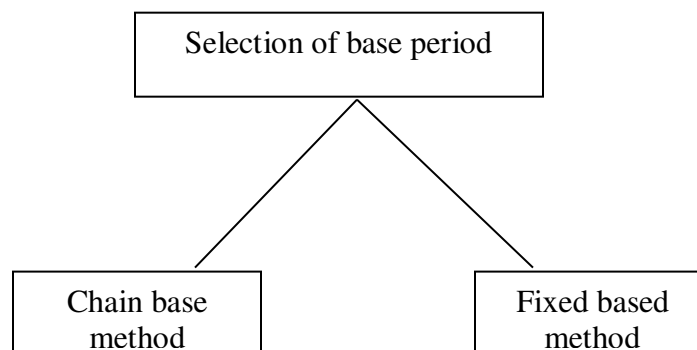
Selection of Commodities:

The following steps are included for selection of commodities:

1. A reasonable number of commodities on the basis of their evaluated importance should be used.
2. While deciding on their number, the commodities to be selected must be taken as,
 - i. Representative of the tastes, habits and requirements of the people concerned
 - ii. Unlikely to vary in quality or grade
 - iii. Comparable

Selection of Based Period:

There are two methods for selection of base period.



Notation:

- p_t : prices (of a vector of commodities) at time t
- q_t : real quantities purchase (of a vector of commodities) at time t
- Suppose 2 periods
- Base period: $t=0$
- Current period: $t=1$

Fixed Based Method:

A fixed base method is one in which a particular year is generally chosen as the base period that remains unchanged during the life term of the index.

Example:

- ▶ Suppose that a cup of coffee in a particular café cost Rs. 35 in 2002. In 2012, an identical cup of coffee cost Rs. 50. How has the price changed from 2002 to 2012?
- ▶ The particular time period of 1995 which we've chosen to compare against, is called the base period.

Price Relative:

A price relative expresses the price of a commodity in a given year as a fraction of the price in the base year. It is multiplied by 100 to make it a percentage.

$$\text{Price Relative} = \frac{p_n}{p_o} \times 100$$

Where P_o denotes the price of the base year and P_n denotes the price of the current period.

Example:

- ▶ Suppose that a cup of coffee in a particular café cost Rs. 35 in 2002. In 2012, an identical cup of coffee cost Rs. 50. How has the price changed from 2002 to 2012?
- ▶ The particular time period of 2002 which we've chosen to compare against, is called the base period.
- ▶ Here $P_o = 35$, $P_n = 50$, then the

$$\begin{aligned} \text{Price Relative} &= \frac{p_n}{p_o} \times 100 \\ \text{Price Relative} &= \frac{50}{35} \times 100 \\ &= 142\% \end{aligned}$$

Example:

Find the price relatives for each year from the following average retail prices of wheat, using

- ❖ 1957 as a base
- ❖ 1960 as a base

Year	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
Retail Price	16	18	18	20	23	25	40	70	90	150

Solution:

Year	Retail Prices (Rs.)	Price Relatives	
		1999 as base	2005 as base
1999	16	100	40
2000	18	112.5	45
2001	18	112.5	45
2002	20	125	50
2003	23	143.8	57.5
2004	25	156.3	62.5
2005	40	250	100
2006	70	437.5	175
2007	90	562.5	225
2008	150	937.5	375

Chain Based Method:

A chain base method is one in which the base period is not fixed, but moves with the given year. For example, for 2008, 2007 will be the base; for 2007, 2006 will be the base and so on.

Example:

Suppose that a cup of coffee in a particular café cost Rs. 35 in 2002. In 2012, an identical cup of coffee cost Rs. 50. In 2014, an identical cup of coffee cost Rs. 60.

Link Relative:

The relatives are computed with the immediately preceding year as the base and such relatives are called link relatives.

$$\text{Link Relative} = \frac{p_n}{p_{n-1}} \times 100$$

Where P_n denotes the price of the commodity in the given year and P_{n-1} denotes the price of the commodity in the preceding year.

Example:

From the data given below construct an index number by chain base method.

Price of a commodity from 2006-2008.

<i>Year</i>	<i>Price</i>
2006	50
2007	60
2008	65

Solution:

<i>Year</i>	<i>Price</i>	<i>Link Relative</i>
2006	50	100
2007	60	120
2008	65	108

Example:

The price relatives of three commodities are given below:

<i>Year</i>	<i>Price Relatives of Commodities</i>		
	A	B	C
1934	100	100	100
1935	105	97	121
1936	110	94	125
1937	115	100	130
1938	116	99	128
1939	120	105	130

Compute the link relative, i.e. the price relatives in each year with reference to the previous year as 100.

Solution:

<i>Year</i>	<i>Link Relatives of Commodities</i>		
	A	B	C
1934	100	100	100
1935	105.0	97.0	121.0
1936	104.8	96.9	103.3
1937	104.5	106.4	104.0
1938	100.9	99.0	98.5
1939	103.5	106.1	101.6

Selection of Average:

The choice of an appropriate average to get a single index number for each year. For this purpose, any of the following averages may be used:

1. The arithmetic mean
2. The median
3. The geometric mean

Selection of Appropriate Weights:

The construction of wholesale price index numbers is to decide how to select weights which would indicate the relative importance various commodities in the group. It is evident that all the commodities selected are not equally important.

For example, eggs and coffee cannot be given the same importance as wheat and rice. Wheat is much more important than coffee, it is therefore, desirable that wheat must be given more importance of each commodity; a sample survey should be conducted.

Types:

There are two types of indices.

- ▶ Un weighted indices- in which no specific weights are attached
- ▶ Weighted indices- in which appropriate weights are assigned to various items.

Example:

The following table gives the production of cotton cloth (in million yards) for Pakistan from 1954 to 1963.

<i>Year</i>	1954	1955	1956	1957	1958	1959	1960	1961	1962	1963
<i>Production</i>	282	389	438	470	511	555	564	630	662	681

Find out the index numbers by taking:

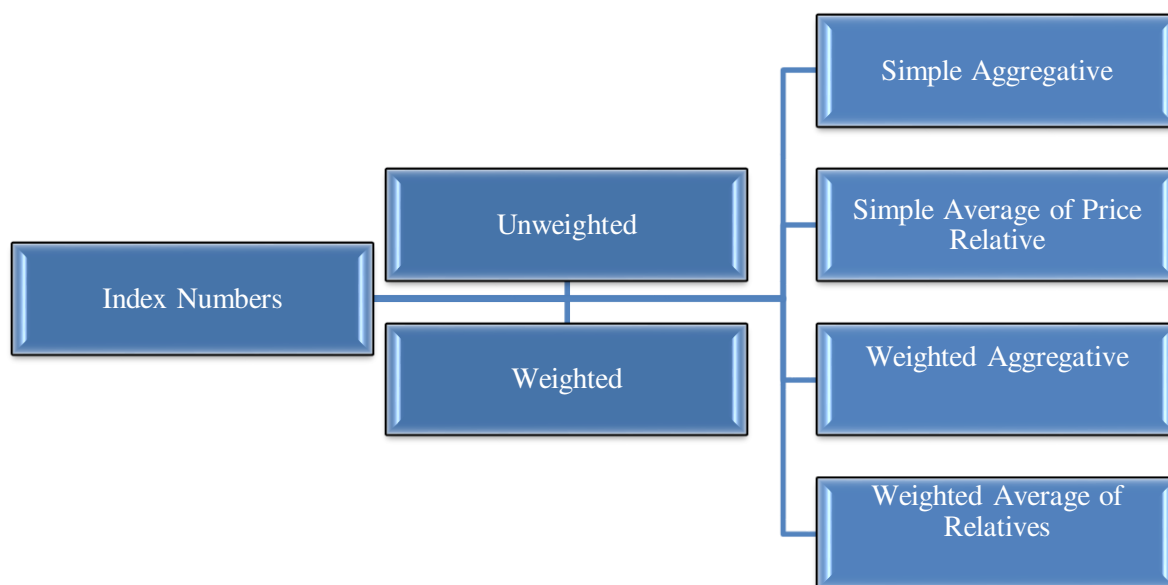
- (i) 1954 as base year
- (ii) Average of 1958, 59, 60 as base period.

Solution:

Year	Production	Price Relatives	
		1954 as base	Average of 1958, 1959 and 1960 as base
1954	282	100	51.90
1955	389	137.94	71.60
1956	438	155.32	80.61
1957	470	166.67	86.50
1958	511	181.21	94.05
1959	555	196.81	102.15
1960	564	200.00	103.80
1961	630	223.40	115.95
1962	662	234.75	121.84
1963	681	241.49	125.34

Lecture: 02

METHOD OF CONSTRUCTION OF INDEX NUMBERS



Un weighted Index Numbers Simple Aggregative Method

It consists in expressing the aggregate price of all commodities in the current year as a percentage of the aggregate price in the base year.

$$P_{0x} = \frac{\sum P_x}{\sum P_0} \times 100$$

Where P_{0n} denotes the price index for the given year n relative to the base year 0 ,

- ♦ $\sum P_n$ denotes the sum of prices for given year, and
- ♦ $\sum P_0$ Denotes the sum of prices for base year.

Example:

The average prices of certain commodities are given below:

Commodity	Unit	Prices (Rs.) in	
		1957	1958
Wheat	Ton	351.00	335.00
Rice	Maund	35.00	32.00
Salt	Maund	10.00	11.00
Sugar	Seer	1.25	1.40
Cloth	Yard	2.25	2.60
Milk	Seer	0.75	0.85
Oil	Gallon	1.25	1.35

Compute a simple aggregative price index number for the year 1958 based on 1957 prices.

Solution:

The simple aggregative price index number is given by:

$$P_{0x} = \frac{\sum P_x}{\sum P_0} \times 100$$

$$P_{0x} = \frac{335.00 + 32.00 + \dots + 1.35}{315.00 + 35.00 + \dots + 1.25} \times 100$$

$$P_{0x} = \frac{384.20}{401.50} \times 100 = 95.7$$

The simple aggregative index for 1958 on the base of 1957 is 95.7. This means that the prices have decreased by 4.3%.

Un weighted Index Numbers Simple Average of Relatives

A simple average of price relatives is an index obtained by taking the average of the price relatives of the given commodities for each year and expressing the result as a percentage. If we take the arithmetic mean, then we have

$$P_{on} = \frac{1}{k} \sum \left(\frac{P_n}{P_0} \right) \times 100$$

Where, k denotes the number of commodities whose price relatives are thus combined.

When geometric mean is used:

$$P_{on} = \left[\pi \left(\frac{P_n}{P_0} \right) \right]^{1/2} \times 100$$

Example:

The following table gives the average wholesale prices in rupees per unit of gold, wheat and cotton during the years 1912 – 1917.

<i>Commodities</i>	<i>Average prices in Rs. Per unit</i>					
	1912	1913	1914	1915	1916	1917
Gold	25.3	30.8	33.4	35.5	35.3	36.0
Wheat	17.3	14.5	4.9	5.7	17.1	11.6
Cotton	7.8	5.4	6.7	5.6	7.2	10.2

Using 1912 as the base period, compute the simple average of relative's price indices for the years 1913 to 1917.

Solution:

Years	Price relatives			Sum of price relatives	Index No. as average of relatives
	Gold	Wheat	Cotton		
1912	100	100	100	300	100
1913	122	84	69	275	91.6
1914	132	28	86	246	82.1
1915	140	33	72	245	81.7
1916	140	100	92	332	110.7
1917	142	67	131	340	113.4

Simple Average of Price Relatives

In the previous lecture, we read the method of simple average of price relatives by using the three different methods:

- Average
- Median
- Geometric Mean

Example:

From the data given below, compute the index number of prices, taking 1962 as base. Use

- (i) Simple average of price relatives.
- (ii) The median of price relatives.
- (iii) The geometric mean of price relatives.

	<i>Commodity (Price in Rs.)</i>			
<i>Year</i>	<i>Firewood</i>	<i>Soft cake</i>	<i>Kerosene oil</i>	<i>Matches</i>
1962	3.25	2.50	0.20	0.06
1963	3.44	2.80	0.22	0.06
1964	3.50	2.00	0.25	0.06
1965	3.75	2.50	0.25	0.06

Solution:

We calculate the price index numbers as in the following table:

	<i>Price Relatives</i>				<i>Total</i>	<i>Index Numbers by</i>	
<i>Year</i>	<i>Firewood</i>	<i>Soft cake</i>	<i>Kerosene oil</i>	<i>Matches</i>		<i>(i) Mean</i>	<i>(ii) Median</i>
1962	100	100	100	100	400	100	100
1963	106	112	110	100	428	107	108
1964	108	80	125	100	413	103	104
1965	115	100	125	100	440	110	108

Continue...

	<i>Price Relatives</i>				$\Sigma \log x$		
<i>Year</i>	<i>Firewood</i>					$\frac{1}{k} \Sigma \log x$	<i>G.M.</i>
1962	100	100	100	100	8	2	100
Log x	2	2	2	2			
1963	106	112	110	100	8.116	2.029	106.90
Log x	2.025	2.049	2.041	2			
1964	108	80	125	100	8.033	2.008	101.86
Log x	2.033	1.903	2.097	2			
1965	115	100	125	100	8.158	2.039	109.39
Log x	2.061	2	2.097	2			

Chain Indices:

Chain indices are the index numbers which are calculated by using link relatives.

Example:

Construct chain indices for the following years, taking 1940 as the base and using the simple average of relatives.

<i>Year</i>	<i>Price in Rs. Per mound</i>		
	<i>Wheat</i>	<i>Rice</i>	<i>Maize</i>
1940	2.80	10.50	2.70
1941	3.40	10.80	3.20
1942	3.60	10.60	3.50
1943	4.00	11.00	3.80
1944	4.20	11.50	4.00

Solution:

First we calculate the link relative by the formula:

$$\text{Link Relative} = \frac{\text{Price of commodity in the given year}}{\text{Price of commodity in the preceding year}} \times 100$$

$$\text{Or } P_{n-1,n} = \frac{p_n}{p_{n-1}} \times 100$$

<i>Year</i>	<i>Price in Rs. Per mound</i>			<i>Sum of relatives</i>	<i>Simple average of relatives</i>	<i>Chain Indices</i>
	<i>Wheat</i>	<i>Rice</i>	<i>Maize</i>			
1940	100	100	100	300	100	100
1941	121	103	119	343	114	114
1942	106	98	109	313	104	118.6
1943	111	104	109	324	108	128.1
1944	105	105	105	315	105	134.5

Example:

Find the chain indices from the following price relatives of four commodities, using the

- Geometric mean of the relatives for each year.

	<i>Commodity</i>			
<i>Year</i>	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>
1951	81	77	119	55
1952	62	104	87	82
1953	104	87	111	100
1954	93	75	154	96
1955	60	43	165	88

Solution:

The link relatives for commodity A (for instance) are computed below:

$$\begin{aligned}
 \text{Link Relative} &= \frac{p_n}{p_{n-1}} \times 100 \\
 \text{Link Relative for 1952} &= \frac{62}{81} \times 100 = 76 \\
 \text{Link Relative for 1953} &= \frac{104}{62} \times 100 = 168 \\
 \text{Link Relative for 1954} &= \frac{93}{104} \times 100 = 89 \\
 \text{Link Relative for 1955} &= \frac{60}{93} \times 100 = 64
 \end{aligned}$$

Continue.....

Calculation of chain indices, using the geometric mean of the relatives.

	<i>Commodity</i>				$\sum \log x$	$\frac{1}{k} \sum \log x$	<i>G.M.</i>	<i>Chain Index</i>
<i>Year</i>	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>				<i>Numbers</i>
1951	81	77	119	55	7.6100	1.9027	79.9	79.9
1952	62	104	87	82	7.9325	1.9831	96.2	76.9
1953	104	87	111	100	8.4580	2.1145	130.2	100.1
1954	93	75	154	96	8.0092	2.0023	100.6	100.7
1955	60	43	165	88	7.5667	1.8917	77.9	78.4

Hence the chain indices by geometric mean are 79.91 76.9, 100.1, 100.7 and 78.4.

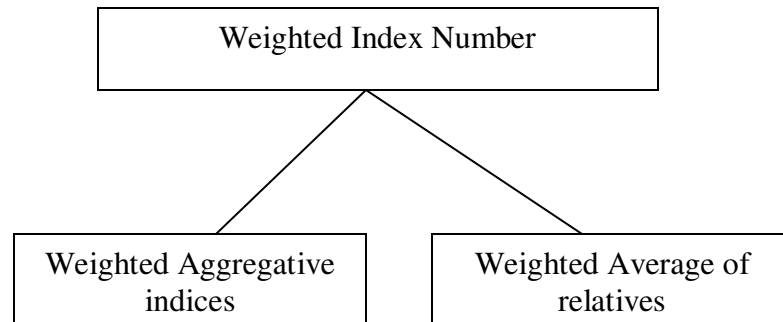
Lecture: 03

In the Current Lecture, we will Study

- ▶ Weighted Index Numbers
- ▶ Weighted Aggregative Price Index Numbers
- ▶ Weighted Average of Relatives Price Index Numbers

Weighted Index Numbers:

An index number that measures the change in the prices of a group of commodities when the relative importance of the commodities (i.e. Weight) has been taken into account is called a weighted price index number. Weighted indices are generally divided into Weighted Aggregative indices and Weighted Average of relative's indices.



Weighted Aggregative Price Index Numbers:

These index numbers are the simple aggregative type with the fundamental difference that weights are assigned to the various items included in the index. There are various kinds of weighted aggregative index numbers are

Laspeyres'
Price
Index

Paasche's
Price
Index

Marshall-
Edgeworth
Price
Index

Fisher's
Ideal Index

Laspeyres' Price Index:

This method was derived by Laspeyres' in 1871. In this method the weights are determined by quantities in the base.

$$P_{0x} = \frac{\sum P_x q_0}{\sum P_0 q_0} \times 100$$

Paasche's Price Index:

This method was derived by German statistician Paasche's in 1874. The weights of current year are used as base year in constructing the Paasche's Index number.

$$P_{0x} = \frac{\sum P_x q_x}{\sum P_0 q_x} \times 100$$

Example:

Construct the following weighted aggregative price index numbers for 1960 by using 1956 as a base period.

- ▶ Laspeyres' index
- ▶ Paasche's index

Commodity	Prices (Rs. Per md.)		Quantities (tons)	
	1956 (base)	1960	1956	1960
A	64	75	270	276
B	40	45	124	118
C	18	21	130	121
D	58	68	185	267

Solution:

Com.	Price		Quantity		p_0q_0	p_1q_0	p_0q_1	p_1q_1
	p_0	p_1	q_0	q_1				
A	64	75	270	276	17280	20250	17664	20700
B	40	45	124	118	4960	5580	720	5310
C	18	21	130	121	2340	2730	2178	2541
D	58	68	185	267	10730	12580	15486	18156
Total					35310	41140	40048	46707

Laspeyres' Price Index:

$$\text{Index for 1960 } (P_{0x}) = \frac{\sum P_x q_0}{\sum P_0 q_0} \times 100 = \frac{41140}{35310} \times 100 = 116.5$$

Paasche's Price Index:

$$\text{Index for 1960 } (P_{0x}) = \frac{\sum P_x q_x}{\sum P_0 q_x} \times 100 = \frac{46707}{40048} \times 100 = 116.5$$

Marshall-Edgeworth Price Index:

In this index the numerator consists of an aggregate of the current year's price multiplied by the weights of both of the base year as well as the current year.

$$P_{0x} = \frac{\sum P_x(q_0 + q_x)}{\sum P_x(q_0 + q_x)} \times 100$$

Fisher's Ideal Index Number:

Fisher's ideal index number is the geometric mean of the Laspeyres' and Paasche's index numbers.

$$P_{0x}(\text{Fisher}) = \sqrt{P_{0x}(\text{Laspeyres}) \times P_{0x}(\text{Paasche})}$$

$$= \sqrt{\frac{\sum p_x q_0}{\sum p_0 q_0} \times \frac{\sum p_x q_x}{\sum p_0 q_x}}$$

Example:

The prices and quantities of three commodities during 1960 and 1964 are given below:

Commodity	Prices (Rs. Per md.)		Quantities Produced (maunds)	
	1960 (p_0)	1964 (p_1)	1960 (q_0)	1964 (q_1)
A	3.95	4.25	9,676	10,436
B	34.80	38.94	78	83
C	61.56	59.70	118	116

Compute the Marshall-Edgeworth index numbers and Fisher Index numbers for 1964, using 1960 as the base period.

Solution:

To calculate the Marshall-Edge worth, first we calculate the necessary products with 1960 as base year. These calculations are shown in the table below:

Computation of Weighted Aggregative Price Indices

Com.	$(p_0 + q_1)$	$p_1(q_0 + q_1)$	$p_0(q_0 + q_1)$
A	20,111	85471.75	79,438.45
B	161	6269.34	5,602.80
C	234	13,969.80	14,40.04
Total		105,710.89	99,446.29

Marshall-Edge worth price index:

$$P_{0x} = \frac{\sum P_x(q_0 + q_x)}{\sum P_x(q_0 + q_x)} \times 100 = \frac{107510.89}{99446.29} = 106.3$$

Com.	Price		Quantity		p_0q_0	p_1q_0	p_0q_1	p_1q_1
	p_0	p_1	q_0	q_1				
A	3.95	4.25	9676	10436	38220.2	41123	41222.2	44353
B	34.80	38.94	78	83	2714.4	3037.32	2888.4	3232.02
C	61.56	59.70	118	116	7264.08	7044.6	7140.96	6925.2
Total					48198.68	51204.92	51251.56	54510.22

Fisher's "Ideal" index numbers:

$$\sqrt{\frac{\sum p_1q_0}{\sum p_0q_0} \times \frac{\sum p_1q_1}{\sum p_0q_1}} \times 100$$

$$\sqrt{\frac{51204.92}{48198.68} \times \frac{54510.22}{51251.56}} \times 100$$

$$\sqrt{1.062 \times 1.064} \times 100 = 106.30$$

Weighted Average of Relatives Price Index Numbers:

It is computed multiplying each price relative by its weight, summing these products and dividing by the sum of the weights. The important types of the weighted average of relative's price indices are given below:

Laspeyres'
Price
Index

Paasche's
Price Index

Laspeyres' Index Numbers:

$$P_{0n} = \frac{\sum (p_n/p_0) p_0 q_0}{\sum p_0 q_0} \times 100$$

Where the price relatives are weighted by the total value of commodities in the base year

Paasche's Index Numbers:

$$P_{0n} = \frac{\sum \left(\frac{p_n}{p_0} \right) p_0 q_n}{\sum p_0 q_n} \times 100$$

Where, the price relatives are weighted by the total value of commodities in the given year at base year.

Example:

Using the data, compute the weighted average of relative's price index numbers for 1964 by

- ▶ Laspeyres method
- ▶ Paasche's method

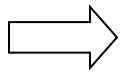
<i>Commodity</i>	<i>Prices (Rs. Per md.)</i>		<i>Quantities Produced (maunds)</i>	
	<i>1960 (p₀)</i>	<i>1964 (p₁)</i>	<i>1960 (q₀)</i>	<i>1964 (q₁)</i>
A	3.95	4.25	9,676	10,436
B	34.80	38.94	78	83
C	61.56	59.70	118	116

Solution:

<i>Com.</i>	<i>Price</i>		<i>Quantity</i>		<i>Price relative</i>	<i>Weights</i>		<i>Weighted Price Relatives</i>	
	<i>p₀ 1960</i>	<i>p₁ 1964</i>	<i>q₀ 1960</i>	<i>q₁ 1964</i>		<i>p₀q₀</i>	<i>p₀q₁</i>		
A	3.95	4.25	9,676	10,436	1.076	38220.2	41222.20	41124.94	44355.09
B	34.80	38.94	78	83	1.119	2714.40	2888.40	3037.41	3232.12
C	61.56	59.70	118	116	0.970	7264.08	7140.95	7046.16	6926.73
Total						48198.68	51251.56	51208.51	54513.94

Laspeyres' index for 1964

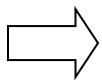
$$= \frac{\sum (P_n / P_0) P_0 Q_0}{\sum P_0 Q_0} \times 100$$



$$\frac{51,208.51}{48,198.68} \times 100 = 106.24,$$

Paasche's index for 1964

$$= \frac{\sum \left(\frac{P_n}{P_0} \right) P_0 Q_n}{\sum P_0 Q_n} \times 100$$



$$\frac{54,513.94}{51,251.56} \times 100 = 106.4.$$

Example:

Compute Fisher's price index numbers for 1976 with 1961 as base from the following data:

Commodity	Quantity (units)		Value (Rs.)	
	1961	1976	1961 ($p_0 q_0$)	1976 ($p_1 q_1$)
A	100	150	600	1200
B	80	100	400	700
C	60	72	180	432
D	30	33	450	363
Total			1630	2695

Solution:

Unit prices		$p_0 q_1$	$p_1 q_0$
p_0	p_1		
6	8	800	900
5	7	560	500
3	6	360	216
15	11	330	495

		$\frac{2050}{\sum p_0 q_1}$	$\frac{2111}{\sum p_1 q_0}$
--	--	-----------------------------	-----------------------------

Fisher's index numbers for 1976:

$$\Rightarrow \sqrt{\frac{\sum p_1 q_0}{\sum p_0 q_0} \times \frac{\sum p_1 q_1}{\sum p_0 q_1}} \times 100$$

$$= \sqrt{\frac{2050}{1630} \times \frac{2695}{2111}} \times 100$$

$$= 126.72.$$

Example:

Calculate the index numbers of Marshall-Edgeworth and Fisher's Ideal type for the following data.

<i>Commodity</i>	<i>Base Year</i>		<i>Current Year</i>	
	<i>Price</i>	<i>Quantity</i>	<i>Price</i>	<i>Quantity</i>
	P_0	q_0	P_1	q_1
A	7	70	5	49
B	5	27	7	28
C	10	35	9	29
D	9	50	4	42
E	3	16	10	25

Solution:

<i>Commodity</i>	<i>Base Year</i>		<i>Current Year</i>		$p_0 q_0$	$p_0 q_1$	$p_1 q_0$	$p_1 q_1$
	<i>Price</i>	<i>Quantity</i>	<i>Price</i>	<i>Quantity</i>				
	P_0	q_0	p_1	q_1				
A	7	70	5	49	490	343	350	245
B	5	27	7	28	135	140	189	196
C	10	35	9	29	350	290	315	261

D	9	50	4	42	450	378	200	168
E	3	16	10	25	48	75	160	250
					1473	1226	1214	1120

Fisher's index numbers:

$$\Rightarrow \sqrt{\frac{\sum p_1 q_0}{\sum p_0 q_0} \times \frac{\sum p_1 q_1}{\sum p_0 q_1}} \times 100$$

$$\Rightarrow \sqrt{\frac{1214}{1473} \times \frac{1120}{1226}} \times 100$$

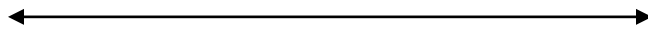
$$\Rightarrow \sqrt{0.8242 \times 0.9135} \times 100$$

$$\Rightarrow = 86.8.$$

Marshall-Edge worth index:

$$\frac{\sum p_1 q_0 + \sum p_1 q_1}{\sum p_0 q_0 + \sum p_0 q_1} \times 100 \quad \Rightarrow \quad \frac{1214 + 1120}{1473 + 1226} \times 100$$

$$\frac{2334}{2699} \times 100 = 86.5.$$



Lecture No. 4

In the Previous Lecture:

- ▶ Method of Construction of Un-Weighted Index Numbers
- ▶ Simple Aggregative Method
- ▶ Simple Average of Relatives

In the Current Lecture,

- ▶ Tests for Index Number
- ▶ Consumer Price Index
- ▶ Construction of Consumer Price Index
- ▶ Sensitive Price Indicator

Test for Index Number,

From a theoretical viewpoint, a “good” index number formula is required to satisfy the following tests proposed by Irvin Fisher (1867 – 1947).

- ▶ Time Reversal Test
- ▶ Factor Reversal Test
- ▶ Circular Test

Time Reversal Test:

Time Reversal Test may be stated as follows:

“If the time subscripts 0 and n are interchanged in a price (or quantity) index number formula wherever they appear, then the resulting price (or quantity) formula should be the reciprocal of the original index formula, ignoring the factor 100.” Symbolically, the test requires that,

$$P_{0n} = \frac{1}{P_{n0}} \text{ or } P_{0n} \times P_{n0} = 1$$

For the purposes of illustration, let us consider some of the important formulae and see whether they or do not satisfy this test.

- The Laspeyres’ price index expressed in ratio rather than percentage form (i.e. omitting the factor 100), is given by:

$$P_{0n} = \frac{\sum P_n q_0}{\sum P_0 q_0}$$

Interchanging the time subscripts, we get

$$P_{n0} = \frac{\sum P_0 q_n}{\sum P_n q_n}$$

But $P_{0n} \times P_{n0}$, i.e. $\frac{\sum p_n q_0}{\sum p_0 q_0} \times \frac{\sum p_0 q_n}{\sum p_n q_n} \neq 1$.

Hence it does not satisfy the time reversal test.

- ii. Again, taking the Paasche's price index formula, i.e. $\frac{\sum p_n q_n}{\sum p_0 q_n}$ and interchanging the time subscripts, we get
- $$P_{n0} = \frac{\sum p_n q_n}{\sum p_0 q_n}$$

But again $P_{0n} \times P_{n0}$ i.e. $\frac{\sum p_n q_n}{\sum p_0 q_n} \times \frac{\sum p_0 q_0}{\sum p_n q_0} \neq 1$, which shows that this formula, too, does not obey the time reversal test.

- iii. Interchanging the time subscripts in the Marshall-Edge worth price index, i.e. in the formula

$$P_{0n} = \frac{\sum p_n (q_0 + q_n)}{\sum p_n (q_0 + q_n)}$$

Multiplying together, we obtain

$$P_{0n} \times P_{n0} = \frac{\sum p_n (q_0 + q_n)}{\sum p_n (q_0 + q_n)} \times \frac{\sum p_0 (q_n + q_0)}{\sum p_0 (q_n + q_0)} = 1.$$

- iv. Interchanging the time subscripts in the Fisher price index, i.e. in the formula

$$P_{0n} = \sqrt{\frac{\sum p_n q_0}{\sum p_0 q_0} \times \frac{\sum p_n q_n}{\sum p_0 q_n}}$$

Multiplying together, we obtain

$$P_{0n} \times P_{n0} = \sqrt{\frac{\sum p_n q_0}{\sum p_0 q_0} \times \frac{\sum p_n q_n}{\sum p_0 q_n}} \times \sqrt{\frac{\sum p_0 q_n}{\sum p_n q_0} \times \frac{\sum p_0 q_0}{\sum p_n q_0}} = 1.$$

Hence the time reversal test is satisfied.

Factor Reversal Test:

Factor Reversal Test may be stated as follows:

“If the factors p's (prices) and q's (quantities) occurring in a price (or quantity) index formula be interchanged (or reversed) so that a quantity (or price) index formula is obtained, then the product of the two index numbers should equal the value index number, i.e. In other words, the factor reversal test requires that

(Price index) (Quantity index) = Value index.

For the purposes of illustration, let us consider some of the important formulae and see whether they or do not satisfy this test.

- i. Interchanging the factors p's and q's in the formula

$$P_{0n} = \frac{\sum p_n q_0}{\sum p_0 q_0} \text{ (Laspeyres' price index)}$$

$$Q_{0n} = \frac{\sum q_n p_0}{\sum q_0 p_0} \text{ (Laspeyres' quantity index)}$$

The product of these two is $\frac{\sum p_n q_0}{\sum p_0 q_0} \times \frac{\sum q_n p_0}{\sum q_0 p_0}$, which is not equal to $\frac{\sum p_n q_0}{\sum p_0 q_0}$, the value index.

ii. Again interchanging the factors p 's and q 's, the formula

$$P_{n0} = \frac{\sum p_n q_n}{\sum p_0 q_n} \text{ (Paasche's price index)}$$

$$\text{Transforms into, } Q_{n0} = \frac{\sum q_n p_n}{\sum q_0 p_n} \text{ (Paasche's quantity index)}$$

$$\text{Thus } P_{0n} \times Q_{0n} = \frac{\sum p_n q_n}{\sum p_0 q_n} \times \frac{\sum q_n p_n}{\sum q_0 p_n} \neq \frac{\sum p_n q_n}{\sum p_0 q_0} \text{ (Value Index)}$$

Hence Paasche's price index does not satisfy the factor reversal test.

iii. The Marshall- Edgeworth's price index is

$$P_{0n} = \frac{\sum p_n (q_0 + q_n)}{\sum p_n (q_0 + q_n)}$$

Interchanging the factors p 's and q 's, the quantity index is obtained as

$$Q_{0n} = \frac{\sum q_n (p_0 + p_n)}{\sum q_n (p_0 + p_n)}$$

But,

$$\frac{\sum p_n (q_0 + q_n)}{\sum p_n (q_0 + q_n)} \times \frac{\sum q_n (p_0 + p_n)}{\sum q_n (p_0 + p_n)} \neq \frac{\sum p_n q_n}{\sum p_0 q_0},$$

i.e. (price index) (quantity index) $\neq$ value index. Thus the factor-reversal test is not satisfied.

iv. By definition, Fisher's price index is given by

$$P_{0n} = \sqrt{\frac{\sum p_n q_0}{\sum p_0 q_0} \times \frac{\sum p_n q_n}{\sum p_0 q_n}}$$

Fisher's quantity index is obtained by interchanging the factors p 's and q 's as

$$Q_{0n} = \sqrt{\frac{\sum q_n p_0}{\sum q_0 p_0} \times \frac{\sum q_n p_n}{\sum q_0 p_n}}$$

Therefore,

$$P_{0n} \times Q_{0n} = \sqrt{\frac{\sum p_n q_0}{\sum p_0 q_0} \times \frac{\sum p_n q_n}{\sum p_0 q_n}} \times \sqrt{\frac{\sum q_n p_0}{\sum q_0 p_0} \times \frac{\sum q_n p_n}{\sum q_0 p_n}} = \frac{\sum p_n q_n}{\sum p_0 q_0}$$

i.e. (price index) (quantity index) = value index. Thus the factor-reversal test is satisfied.

Circular Test:

The circular test is an extension of the time reversal test, which is applicable for two years, i.e. the given year and the base year. But in circular test, three or more years are taken into account. This test may be stated thus: “If P_{ab} is the price index for the year ‘b’ with year ‘a’ as base, P_{bc} is the price index for the year ‘c’ with year ‘b’ as base and P_{ca} is the price index for the year ‘a’ with year ‘c’ as base, then the product of these index numbers should equal 1. In other words, the circular test requires $P_{ab} \times P_{bc} \times P_{ca} = 1$ ”

This test is not obeyed by any of the weighted index numbers unless the weights are constant. For the purposes of illustration, we consider the weighted aggregate price index number with fixed weights, say, q_0 .

Let there be three years denoted by ‘a’, ‘b’, and ‘c’.

Then the weighted price index numbers are

$$P_{ab} = \frac{\sum p_b q_0}{\sum p_a q_0}, P_{bc} = \frac{\sum p_c q_0}{\sum p_b q_0} \text{ and } P_{ca} = \frac{\sum p_a q_0}{\sum p_c q_0}$$

Multiplying them together, we get

$$P_{ab} \times P_{bc} \times P_{ca} = \frac{\sum p_b q_0}{\sum p_a q_0} \times \frac{\sum p_c q_0}{\sum p_b q_0} \times \frac{\sum p_a q_0}{\sum p_c q_0} = 1.$$

Which shows that the circular test is satisfied?

Example:

Show with the help of the following data that the factor-reversal and time-reversal tests are satisfied by Fisher’s Ideal formula for index number construction.

Commodity	Base Year		Current Year	
	Price (In Rs.)	Quantity (in mds.)	Price (in Rs.)	Quantity (in mds.)
A	6	50	10	56
B	4	100	2	120
C	2	60	6	60
D	10	30	12	24
E	8	40	12	36

Solution:

Commodity	Price		Quantity		p_0q_0	p_1q_0	p_0q_1	p_1q_1
	p_0	p_1	q_0	q_1				
A	6	10	50	56	300	500	336	560
B	4	2	100	120	400	200	480	240
C	2	6	60	60	120	360	120	360
D	10	12	30	24	300	360	240	288
E	8	12	40	36	320	480	228	432
Total					1440 $\sum p_0q_0$	1900 $\sum p_1q_0$	1464 $\sum p_0q_1$	1880 $\sum p_1q_1$

Now Fisher's price index for the current year (omitting the factor 100) is given by,

$$P_{01} = \sqrt{\frac{\sum p_1q_0}{\sum p_0q_0} \times \frac{\sum p_1q_1}{\sum p_0q_1}} = \sqrt{\frac{1900}{1440} \times \frac{1880}{1464}}$$

And Fisher's quantity index for the current year (omitting the factor 100) is given by:

$$Q_{01} = \sqrt{\frac{\sum q_1p_0}{\sum q_0p_0} \times \frac{\sum q_1p_1}{\sum q_0p_1}} = \sqrt{\frac{1464}{1440} \times \frac{1880}{1900}}$$

$$\begin{aligned}
 P_{01} \times Q_{01} &= \sqrt{\frac{1900}{1440} \times \frac{1880}{1464}} \times \sqrt{\frac{1464}{1440} \times \frac{1880}{1900}} \\
 &= \sqrt{\frac{1880}{1440} \times \frac{1880}{1400}} = \frac{1880}{1440} = \frac{\sum p_1q_1}{\sum p_0q_0}
 \end{aligned}$$

Value of Index

Hence we see that Fisher's Ideal formula for index numbers satisfies the factor-reversal test.

Again interchanging the time subscripts in Fisher's price index (omitting the factor 100), we get

$$\begin{aligned}
 P_{01} &= \sqrt{\frac{\sum p_0q_1}{\sum p_1q_1} \times \frac{\sum p_0q_0}{\sum p_1q_0}} \\
 P_{01} \times P_{10} &= \sqrt{\frac{1900}{1440} \times \frac{1880}{1464}} \times \sqrt{\frac{1464}{1880} \times \frac{1440}{1900}} = 1
 \end{aligned}$$

Thus,

Hence Fisher's Ideal formula for index numbers also satisfies the time-reversal test.

Consumer Price Index:

- ▶ A consumer price index (CPI) is designed to measure the changes in the composite price of a specified “basket” of goods and services during the given period as compared with the base period.
- ▶ The quantities consumed or the expenditures incurred on various groups are used as weights for the average retail prices prevailing in the locality concerned during the base and given period.
- ▶ A consumer price index is also called a household-budget price index or a retail price index.

Construction of Consumer Price Index:

- ▶ Scope
- ▶ Household Budget Inquiry and Allocation of Weights
- ▶ Price Data
- ▶ Computation of Index
- i. The Aggregate Expenditure
- ii. Household Budget Method

Aggregate Expenditure Method:

- ▶ The quantities consumed by households in the base year are taken as the weights.
- ▶ The quantity $\sum p_0 q_0$ represents the aggregate expenditure acquired on the various items in the base year.
- ▶ The quantity $\sum p_n q_0$ represents the given year with base year quantities.
- ▶ The laspeyres' formula is given by:

$$P_{0n} = \frac{\sum p_n q_0}{\sum p_0 q_0} \times 100$$

- ▶ The given year quantities are not used as weights because they change from year to year.
- ▶ It involves expense, labor and time would be needed every year.

Household Budget Method:

- ▶ The method is called Household Budget Method, because the amounts of money spent by the household concerned are obtained from a household consumption survey.
- ▶ This method is also known as the Weighted Average of Relatives.

$$P_{0n} = \frac{\sum W \left(\frac{p_n}{p_0} \right)}{\sum W} \times 100 = \frac{\sum W \times I}{\sum W}$$

Where W may be $p_0 q_0$ and $I = \frac{p_n}{p_0} \times 100$.

Example:

Compute the consumer price index number for 1940 on the basis of 1939 from the following data, using

- i. Aggregate Expenditure Method
- ii. Household Budget Method.

Commodity	Quantity Consumed In 1939	Unit Of Price	Price in 1939	Price in 1940
<i>Rice</i>	6 mounds	Rs. per md.	5.75	6.00
<i>Wheat</i>	6 mounds	Rs. per md.	5.00	8.00
<i>Gram</i>	1 mound	Rs. per md.	6.00	9.00
<i>Sugar</i>	1 mound	Rs. per md.	20.00	15.00
<i>Arhar</i>	6 mounds	Rs. per md.	8.00	10.00
<i>Ghee</i>	4 seers	Rs. per md.	2.00	1.50

Solution:

- i. Aggregate Expenditure Method:

Commodity	Quantity Consumed In 1939 q_0	Unit Of Price	Price in 1939 P_0	Price in 1940 P_1	$P_0 q_0$	$P_1 q_0$
Rice	6 mounds	Rs. per md.	5.75	6.00	34.50	36.00
Wheat	6 mounds	Rs. per md.	5.00	8.00	30.00	48.00
Gram	1 mound	Rs. per md.	6.00	9.00	6.00	9.00
Sugar	1 mound	Rs. per md.	20.00	15.00	20.00	15.00
Arhar	6 maunds	Rs. per md.	8.00	10.00	48.00	60.00
Ghee	4 seers	Rs. per md.	2.00	1.50	8.00	6.00
Total					146.50	174.00

- i. Aggregate Expenditure Method:

$$\begin{aligned}
 \text{Consumer price index for 1940} &= \frac{\sum P_1 q_0}{\sum P_0 q_0} \times 100 \\
 &= \frac{174.00}{146.50} \times 100 = 118.8
 \end{aligned}$$

ii. Household Budget Method:

Commodity	Quantity Consumed In 1939 q_0	Unit Of Price	Price in 1939 p_0	Price in 1940 p_1	Weight W $p_0 q_0$	Price Relative $I = \frac{p_1}{p_0} \times 100$	$W \times I$
Rice	6 maunds	Rs. per md.	5.75	6.00	34.50	104.30	3598.36
Wheat	6 maunds	Rs. per md.	5.00	8.00	30.00	160.00	4800.00
Gram	1 maund	Rs. per md.	6.00	9.00	6.00	150.00	900.00
Sugar	1 maund	Rs. per md.	20.00	15.00	20.00	75.00	1500.00
Arhar	6 maunds	Rs. per md.	8.00	10.00	48.00	125.00	6000.00
Ghee	4 seers	Rs. per md.	2.00	1.50	8.00	75.00	600.00
Total					146.50	174.00	17398.36

ii. Household Budget Method:

$$\begin{aligned}
 \text{Consumer price index for 1940} &= \frac{\sum W \times I}{\sum W} \\
 &= \frac{17398.36}{146.50} = 118.8
 \end{aligned}$$

Sensitive Price Indicator:

- ▶ The degree to which the price of a product affects the consumers purchasing behaviors.
- ▶ The degree of price sensitivity varies from product to product and from consumer to consumer.
- ▶ In economics, price sensitivity is commonly measured using the price elasticity demand.



Lecture: 05

In the Current Lecture

- ▶ Drawbacks of Consumer Price Index
- ▶ Uses of Index Numbers
- ▶ Limitations of Index Numbers
- ▶ Practical Questions

Drawbacks of Consumer Price Index:

- ▶ It is practically difficult to clearly define one category of people from another.
- ▶ As the construction of consumer price indices involves the sampling of goods and services, the sampling error and biases may affect indices and render them to suspect.
- ▶ In case of certain goods, it is difficult to collect prices actually needed.
- ▶ It is also difficult to eliminate the effect of changes in quality and grade of goods and services purchased by households.
- ▶ During the course of household budget inquiry, the prices of goods and services and their demand may change.
- ▶ The consumer price indices cannot be used for comparing the price changes in consumption goods and services in two localities or in two households in the same locality as no two households can be homogeneous.

Uses of Index Numbers:

- ▶ The price index numbers are used to measure changes in a particular group of prices and help us in comparing the movement in prices of one commodity with another.
- ▶ Index numbers of industrial production provide a measure of change in the level of industrial production in a country.
- ▶ The quantity index numbers show the rise or fall in the volume of production, volume of exports and imports, etc.
- ▶ The import and export price indices are used to measure the changes in the terms of trade of a country.
- ▶ Index numbers are also used to forecast business conditions of a country and to discover seasonal fluctuations and business cycles.
- ▶ The consumer price indices indicate the movements in retail price of consumption goods and services.
- ▶ Index numbers are also used to measure enrolment changes, intelligence quotients and the performance of students.

Limitations of Index Numbers:

- ▶ It is not practicable to price all the goods and services as well as to take into account all changes in quantity or product.
- ▶ Since the construction of almost all index numbers is based on sampling of some kind, therefore the sampling errors creep into their calculations.
- ▶ In price indices the choice of a normal period is difficult as few periods can be regarded as normal for all segments of the economy.
- ▶ The results obtained by different methods of construction may not quite agree.
- ▶ Comparisons of changes in variables over long periods are not reliable.
- ▶ All index numbers are not suitable for all purposes.

Example:

The prices of four commodities quoted at Multan for May 1970, April 1971 and May 1971 are given below:

Commodity	May 1970	April 1971	May 1971
Wheat	17.50	18.37	17.58
Barley	14.58	14.58	16.50
Jawar	14.67	13.94	15.25
Bajra	17.50	13.75	13.42

Compute price index numbers for April 1971 and May 1971 with May 1970 as base, using (i) simple aggregative method, (ii) simple average of price relatives and (iii) geometric mean of price relatives.

Solution:**1. Simple aggregative method:**

$$P_{on} = \frac{\sum P_n}{\sum P_o} \times 100$$

i. Simple aggregative method for April 1971:

$$P_{01} = \frac{18.37 + 14.58 + 13.94 + 13.75}{17.50 + 14.58 + 14.67 + 17.50} \times 100 = 94.38$$

ii. Simple aggregative method for May 1971:

$$P_{02} = \frac{17.58 + 16.50 + 15.25 + 13.42}{17.50 + 14.58 + 14.67 + 17.50} \times 100 = 97.67$$

2. Simple average of price relatives:

	Price relatives				Sum of price relatives	Index No. As average of relatives
Year	Wheat	Barley	Jawar	Bajra		
May 1970	100	100	100	100	400	100
April 1971	105.0	100	95.0	78.6	378.6	94.65
May 1971	100.5	113.2	104.0	76.7	394.4	98.60

3. Geometric mean of price relatives:

	Price relatives				$\sum \log x$	$\sum \log x / n$	Indices
Year	Wheat	Barley	Jawar	Bajra			
May 1970 log	100 2.0000	100 2.0000	100 2.0000	100 2.0000	8.0000	2.0000	100
April 1971 log	105.0 2.0212	100 2.0000	95.0 1.9777	78.6 1.8954	7.8943	1.9738	94.14
May 1971 log	100.5 2.0021	113.2 2.0539	104.0 2.0170	76.7 1.8848	7.9578	1.9894	97.59

Example:

Construct the Chain Indices for the following data for price relatives for 1941 to 1944.

Year	Sugar	Gur	Tea	Coffee
1941	98	75	82	99
1942	100	82	74	100
1943	114	83	78	104
1944	109	84	89	95

Solution:

Year	Link Relatives				Total of L.R	Average	Chain Indices
	Sugar	Gur	Tea	Coffee			
1941	98	75	82	99	354	88.5	88.5
1942	102	109	90	101	402	100.5	88.9
1943	114	101	105	104	424	106.0	94.2
1944	96	101	114	91	402	100.5	94.7

Example:

Given the following, construct Fisher's Ideal index number for

- i. 1964, taking 1957 as base year.
- ii. 1957, taking 1964 as base year.

Commodity	1957		1964	
	Price	Quantity	Price	Quantity
Rice	9.3	100	4.5	99
Wheat	6.4	11	3.7	19
Jawar	5.1	5	2.7	3

Solution:

1. Fisher's Ideal index number for
- i. 1964, taking 1957 as base year.

Commodity	1957 Base Year		1964 Current Year		P_0q_0	P_0q_1	P_1q_0	P_1q_1
	Price	Quantity	Price	Quantity				
Rice	9.3	100	4.5	90	930.0	450.0	837.0	405.0
Wheat	6.4	11	3.7	19	70.4	40.7	121.6	70.3
Jawar	5.1	5	2.7	3	25.5	13.5	15.3	8.1

Fisher's Ideal index number for 1964:

$$\begin{aligned}
 &= \sqrt{\frac{\sum p_1q_0}{\sum p_0q_0} \times \frac{\sum p_1q_1}{\sum p_0q_1}} \times 100 \\
 &= \sqrt{\frac{504.2}{1025.5} \times \frac{483.5}{973.9}} \times 100 = 49.4
 \end{aligned}$$

1. Fisher's Ideal index number for
- ii. 1957, taking 1964 as base year.

Commodity	1957 Current Year		1964 Base Year		P_0q_0	P_0q_1	P_1q_0	P_1q_1
	Price	Quantity	Price	Quantity				
Rice	9.3	100	4.5	90	405	837	450	930
Wheat	6.4	11	3.7	19	70.3	121.6	40.7	70.4
Jawar	5.1	5	2.7	3	8.1	15.3	13.5	25.5
					483.4	973.9	504.2	1025.9

1. Fisher's Ideal index number for 1957:

$$\begin{aligned}
 &= \sqrt{\frac{\sum P_1q_0}{\sum P_0q_0} \times \frac{\sum P_1q_1}{\sum P_0q_1}} \times 100 \\
 &= \sqrt{\frac{973.9}{483.4} \times \frac{1025.9}{504.2}} \times 100 = 202.5
 \end{aligned}$$

Example:

The prices and quantities of three commodities during 1955 and 1965 are given below:

Commodity	1955		1965	
	Price	Quantity	Price	Quantity
A	10	501	12	600
B	40	100	38	194
C	50	76	40	56

Using 1955 as the base period and base period quantities as the weights, compute the weighted aggregative price index and the weighted average of relative price index for 1965.

Solution:

1. Weighted Aggregate Price Index for 1965:

Commodity	1955		1965		P_1q_0	P_0q_0
	Price	Quantity	Price	Quantity		
A	10	501	12	600	6012	5010
B	40	100	38	194	3800	4000
C	50	76	40	56	3040	3800

1. Weighted Aggregate Price Index for 1965:

$$P_{0n} = \frac{\sum p_n q_0}{\sum p_0 q_0} \times 100$$

$$P_{0n} = \frac{\sum p_1 q_0}{\sum p_0 q_0} \times 100$$

$$P_{0n} = \frac{12852}{12810} \times 100 = 100.33$$

2. Weighted Average of Relative Price Index for 1965:

<i>Commodity</i>	<i>1955</i>		<i>1965</i>		$\frac{p_1}{p_0}$	$\left(\frac{p_1}{p_0}\right) p_0 q_0$
	<i>Price</i>	<i>Quantity</i>	<i>Price</i>	<i>Quantity</i>		
<i>A</i>	10	501	12	600	1.20	6012
<i>B</i>	40	100	38	194	0.95	3800
<i>C</i>	50	76	40	56	0.80	3040

2. Weighted Average of Relative Price Index for 1965:

$$P_{0n} = \frac{\sum (P_n / q_0) p_0 q_0}{\sum p_0 q_0} \times 100$$

$$P_{0n} = \frac{\sum (P_1 / q_0) p_0 q_0}{\sum p_0 q_0} \times 100$$

$$P_{0n} = \frac{12852}{12810} \times 100 = 100.33$$



Lecture: 06

Time series and Data Analysis

Structures of Data:

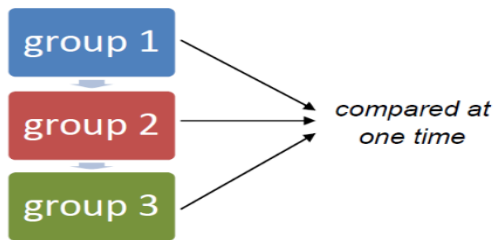
- ▶ Cross-Sectional Data
- ▶ Time series Data
- ▶ Panel Data

Cross Sectional Data:

- ▶ These are data from units observed at the same time or in the same time period.
- ▶ The data may be single observations from a sample survey or from all units in a population.

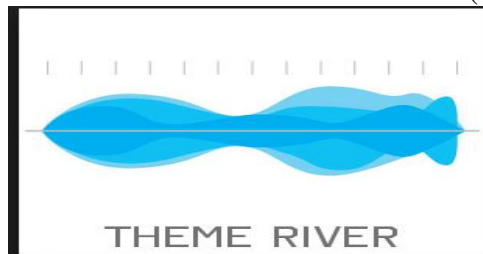
Cross-Sectional Research Study:

different groups



Time Series Data:

These are data from a unit (or a group of units) observed in several successive periods.



Panel Data:

- ▶ Panel data, also called longitudinal data or cross sectional time series data, are data where multiple cases (people, firms, countries etc).
- ▶ were observed at two or more time periods. An example is the National Longitudinal Survey of Youth, where a nationally representative sample of young people were each surveyed repeatedly over multiple years.



Objective of Time Series Analysis:

- ▶ Compact description of data
- ▶ Interpretation
- ▶ Forecasting
- ▶ Control
- ▶ Hypothesis testing
- ▶ Simulation

Lecture content:

- ▶ Time Series
- ▶ Historigram
- ▶ Component of Time Series

Time Series:

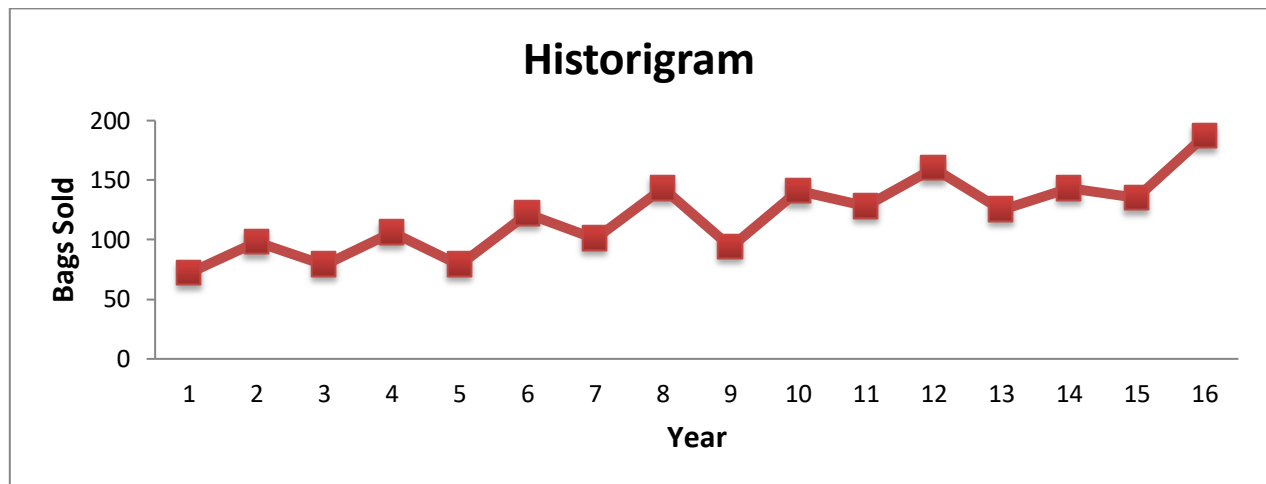
- ▶ A time series consists of numerical data collected, observed or recorded at more or less regular intervals of time each hour, day, month, quarter or year.
- ▶ It is set of data in which observations are arranged in a chronological order.

Example of Time Series:

- ▶ Hourly temperature recorded at a area.
- ▶ Weekly prices of wheat in Lahore.
- ▶ Monthly consumption of electricity in a certain town.
- ▶ Monthly total of passengers carried by rail.
- ▶ Quarterly sales of a certain fertilizer.
- ▶ Annual rainfall at Karachi for a number of years.
- ▶ Enrolment of students in a college over a number of years.

Graph of Time Series:

- ▶ A curve showing changes in the value of one or more items from one period to next is known as the graph of a time series. This curve is called a ***Historigram***.
- ▶ **Historigram** displays the variations in time series dealing with prices, production, imports, population, etc.
- ▶ To construct a ***historigram***, time is taken along X-axis and the values of the variable along Y-axis. Points are plotted and then connected the points by the straight line.



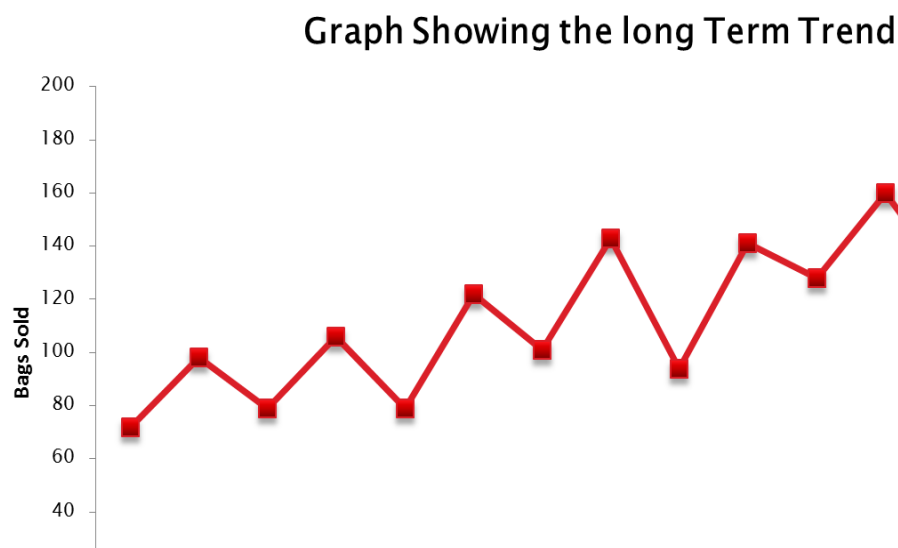
Component of Time Series:

- ▶ Secular Trend
- ▶ Seasonal Variation
- ▶ Cyclic Variation
- ▶ Random Variation

Secular Trend:

- ▶ A secular trend (T) is a long-term movement that persists (Continues) for many years and indicates the general direction of the change of observed values.
- ▶ It refers to a smooth, broad movement of a time series in a same direction, showing a gradual rise or fall within the data.
- ▶ The secular trend generally dominates other variations in the long run and covers a fairly long period of time, not less than 10 years.
- ▶ The long term trend is a peculiar characteristics of most of the economic variables such as sales, prices, industrial production, capital information, etc.
- ▶ It also helps in business planning.

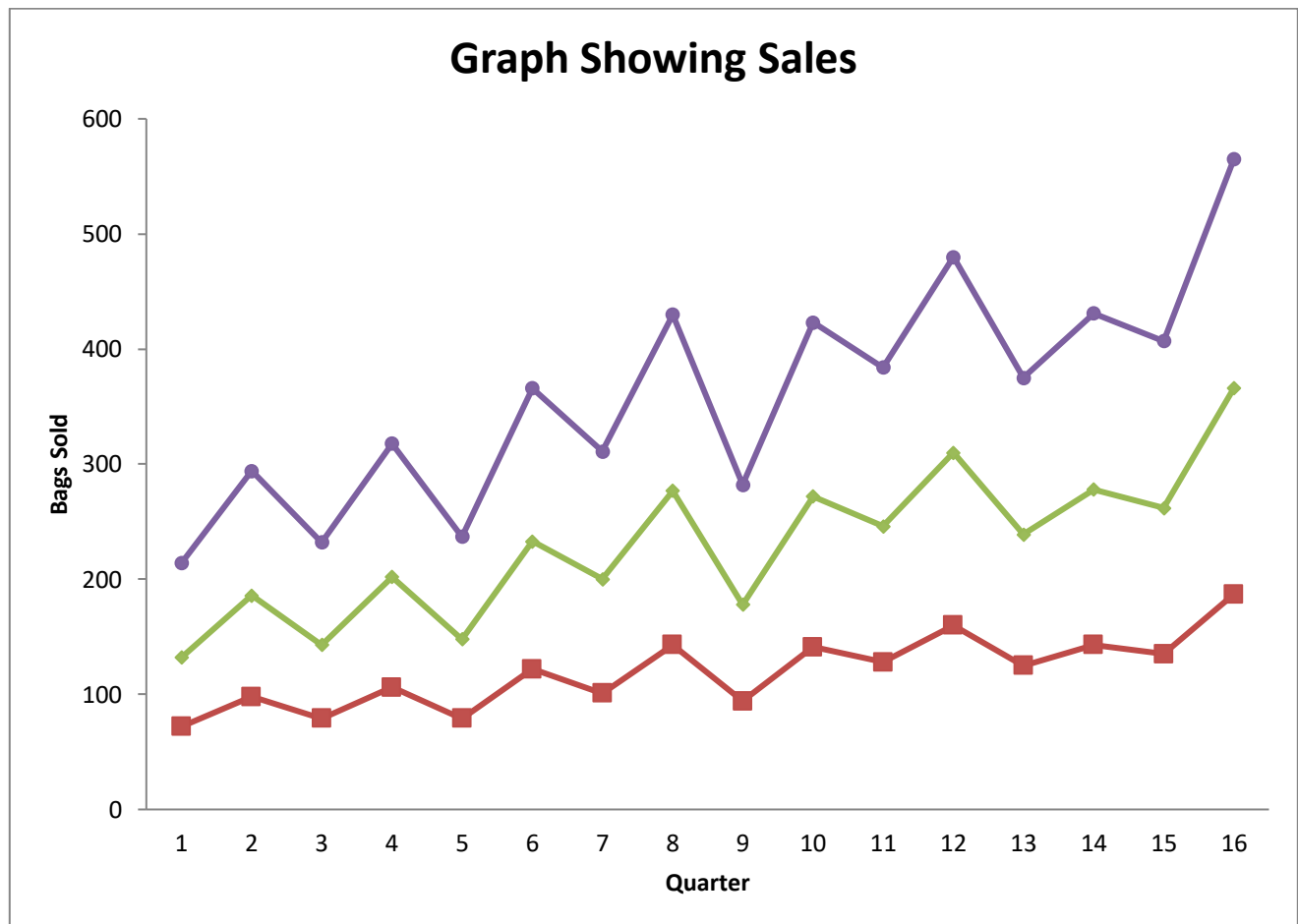
Graph of Secular Trend:



Seasonal Variations:

- ▶ seasonal variations (S), which are mainly caused by the change in seasons, are short-term movements occurring in a periodic manner.
- ▶ The seasonal variations are repeated with more or less the same intensity within a specific period of one year or shorter .
- ▶ The main causes for seasonal variations are the weather conditions, the religious festivals and the social customs.
- ▶ The concept of seasonal variation is occurring within a day, a week, a month, a quarter and so forth.
- ▶ Examples of such variations are the daily variations in temperature or the monthly variations in Bank deposits.

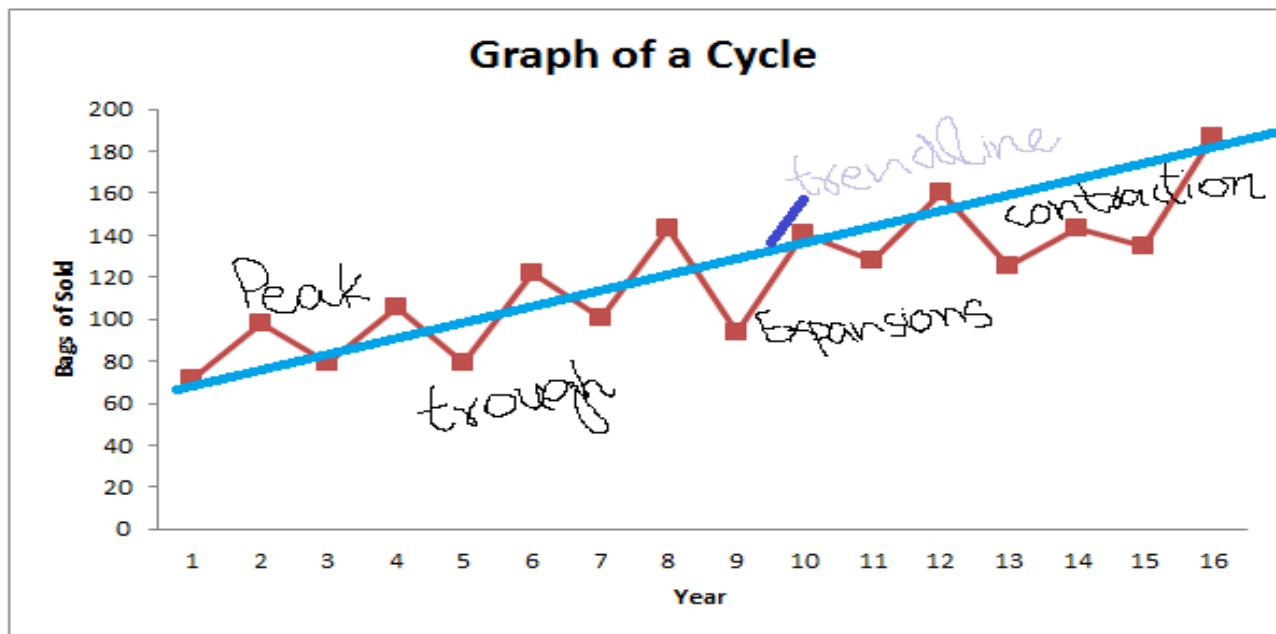
Graph of Seasonal Trend:



Cyclical Fluctuations:

- ▶ Cyclical fluctuations are the long-period oscillations about the long-term trend, which trend to occur in a more or less regular pattern over a period of certain number of years.
- ▶ It is also called business cycles which represent alternating period of prosperity and depression.
- ▶ A cycle, as it is known, is said to be completed when beginning with a peak (a peak is a value which is greater than the two neighboring values), the declining curve reaches a low point, called a trough (a trough is a value which is lower than the two neighboring values) and then rising again reaches the next peak.
- ▶ The period from peak to peak or from trough to trough, is referred to as the duration of a cycle.
- ▶ Cycles have a duration of anywhere from two to ten years or even a longer period.
- ▶ In general, a complete business cycle has the following four phases:
 - 1) Period of prosperity
 - 2) Period of contraction
 - 3) Period of recession or depression
 - 4) Period of recovery or expansion

Graph of Cyclical Variations:



Random Variations:

- ▶ These variations are irregular and unsystematic in nature.
- ▶ These variations occur in a completely unpredictable manner as they are caused by some unusual events such as floods, droughts, strikes, fires, earthquakes, wars, locusts, political events, etc.
- ▶ These variations are also known as the accidental, residual or erratic variations.
- ▶ It is difficult to make a study of such non-recurring variations, though they can be easily identified.

Practice Questions:

Question 01:

Associate the following phenomena of business with the components of time series they belong to:

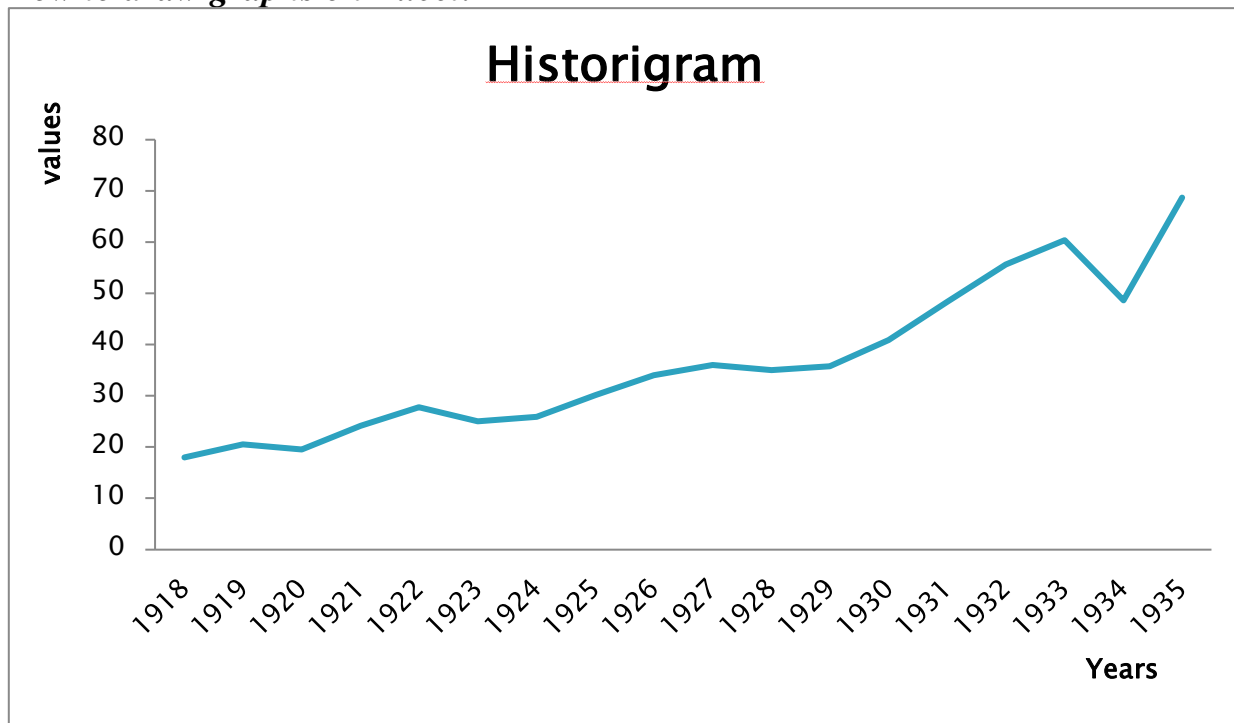
- ▶ The prosperity in a business.
- ▶ The production of sugar recorded for 1956,1957,....,1962.
- ▶ The weekly statement of the sale of pens.
- ▶ The festival sale.
- ▶ The fire in a factory.

Question 02:

With which characteristic movement of a time series would you mainly associate each of the following?

- ▶ A recession
- ▶ An increase in employment during summer months
- ▶ The decline in the death rate due to advance in science
- ▶ A steel strike
- ▶ A continually increasing demand for smaller automobiles.

How to draw graphs on Excel:



Lecture: 07

In the Current Lecture

- ▶ Time Series Decomposition
- ▶ Analysis of Secular Trend
- ▶ Moving Average

Time Series Decomposition:

- ▶ A time series analysis is mainly concerned with the decomposition of the observed series data into its components so as to estimate their separate effects.
- ▶ To do this, we must make assumptions about the relationship existing among the various components. Such as:
 - Multiplicative Model
 - Additive Model

Multiplicative Model:

- ▶ In multiplicative model, we assume that each observed value at any time can be determined by the product of the measures of all four components T, C, S and I. Symbolically,

$$Y_t = T \times S \times C \times I = TSCI$$

- ▶ Where observed values and T-values are stated in original units but the other components S, C and I are expressed in percentages.

Additive Model:

- ▶ In additive model, each value that we observed, is thought of as being a sum of all the four components.

$$Y_t = T + S + C + I$$

- ▶ Where the components T, S, C and I are assumed to be mutually independent.
- ▶ Before analyzing a time series, it is often desirable to adjust the data for calendar variations, for holidays, for price changes and so forth.

Analyzing the Secular Trend:

- ▶ The analysis of a trend component involves its measurement and/or elimination from an observed time series data.
- ▶ To measure a trend which can be represented as a straight line or some type of smooth curve, the following methods are used.
 - Method of freehand curve
 - Method of Semi-averages
 - Method of moving averages
 - Method of least-squares

Method of Freehand Curve:

- ▶ Plot the given data on a graph paper and join the plotted points by segments of straight line.
- ▶ Observe the up and down movements on the graph and draw a smooth curve or a straight line freehand passing through the plotted points that show the direction of change.
- ▶ This line may also be drawn by a transparent ruler.
- ▶ This method is simple and quick.

Disadvantages of Method of Freehand Curve:

- ▶ The freehand curve is a rough and subjective method. As the drawing of trends depends on individual judgment and experience, different persons will draw the graphical trend at different positions with different slopes.

Method of Semi-Averages:

- ▶ Divide the values in the series into two equal parts, including the middle value in each half or omitting it altogether when the number of values is odd.
- ▶ Find the average of the values in each part and place the average values against the respective midpoints of the two parts.
- ▶ Plot these two average values on the graph of the original values, draw a straight line connecting the two points and extend the line to cover the whole series.
- ▶ This method is simple and quick. It gives an entirely objective result when the trend is a straight line.

Disadvantages of Method of Semi-Averages:

- ▶ The arithmetic mean is used to average the observed values is greatly affected by abnormally small or large values.
- ▶ The method is only suitable when the trend is linear or nearly linear.

Example:

Compute and insert on graphs the semi-average trends for the following series:

a) Annual profits in thousands of rupees in a certain business.

<i>Year</i>	<i>1973</i>	<i>1974</i>	<i>1975</i>	<i>1976</i>	<i>1977</i>	<i>1978</i>	<i>1979</i>	<i>1980</i>
<i>Profit</i>	85	97	100	90	83	105	112	120

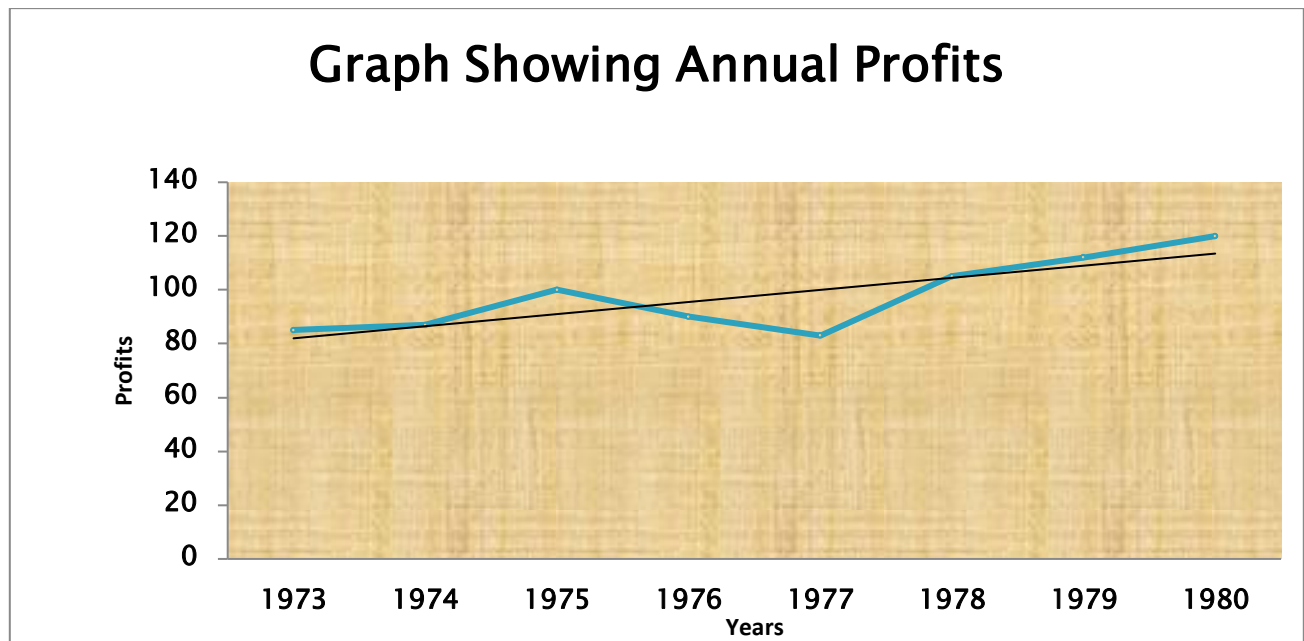
b) Indoor patients ('000) treated in Hospitals in the Punjab.

<i>Year</i>	<i>1967</i>	<i>1968</i>	<i>1969</i>	<i>1970</i>	<i>1971</i>	<i>1972</i>	<i>1973</i>	<i>1974</i>	<i>1975</i>
<i>Patients</i>	276	270	260	286	302	321	351	348	346

Solution:

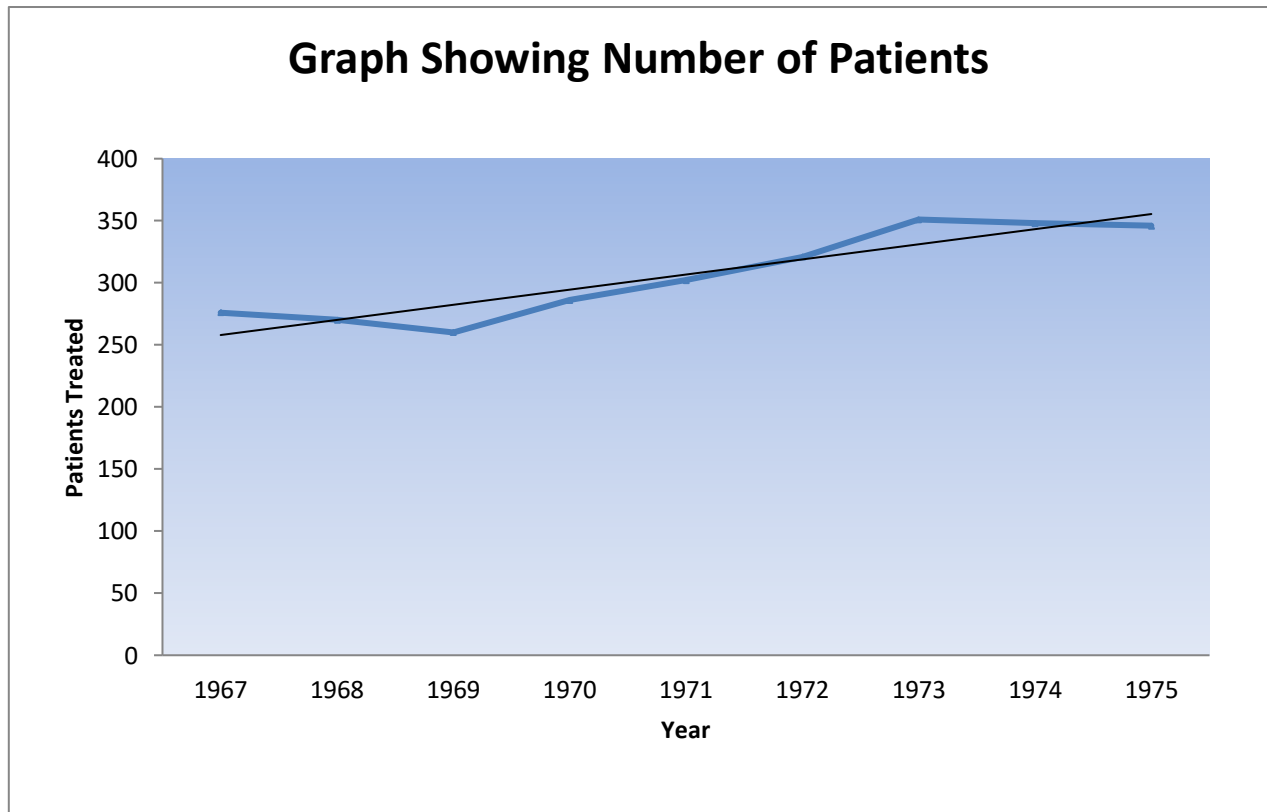
a)

<i>Year</i>	<i>Profit</i>	<i>Total</i>	<i>Averages</i>
1973	85	372	93
1974	97		
1975	100		
1976	90		
1977	83	420	105
1978	105		
1979	112		
1980	120		



b)

<i>Year</i>	<i>Patients</i>	<i>Total</i>	<i>Averages</i>
1967	276	1092	273
1968	270		
1969	260		
1970	286		
1971	302		
1972	321	1366	341.5
1973	351		
1974	348		
1975	346		



Example:

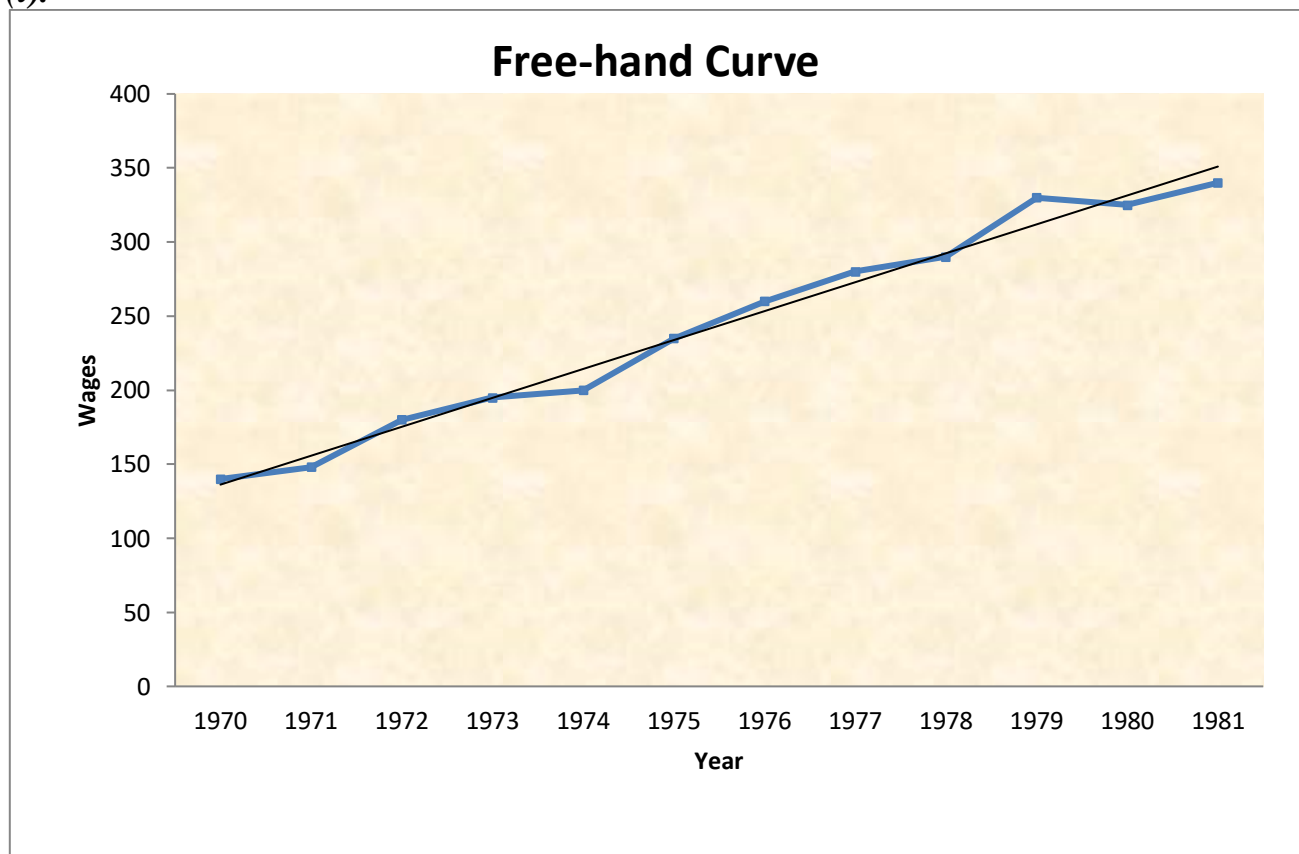
Plot the following data showing average wages in rupees of some workers during 1970-1981. Compute the method of

- i. Free-hand curve
- ii. Semi-averages

Year	1970	1971	1972	1973	1974	1974	1975	1976	1977	1978	1979	1980
Wag-es	140	148	180	195	200	235	260	280	290	330	325	340

Solution:

(i).



ii. *Semi-averages*

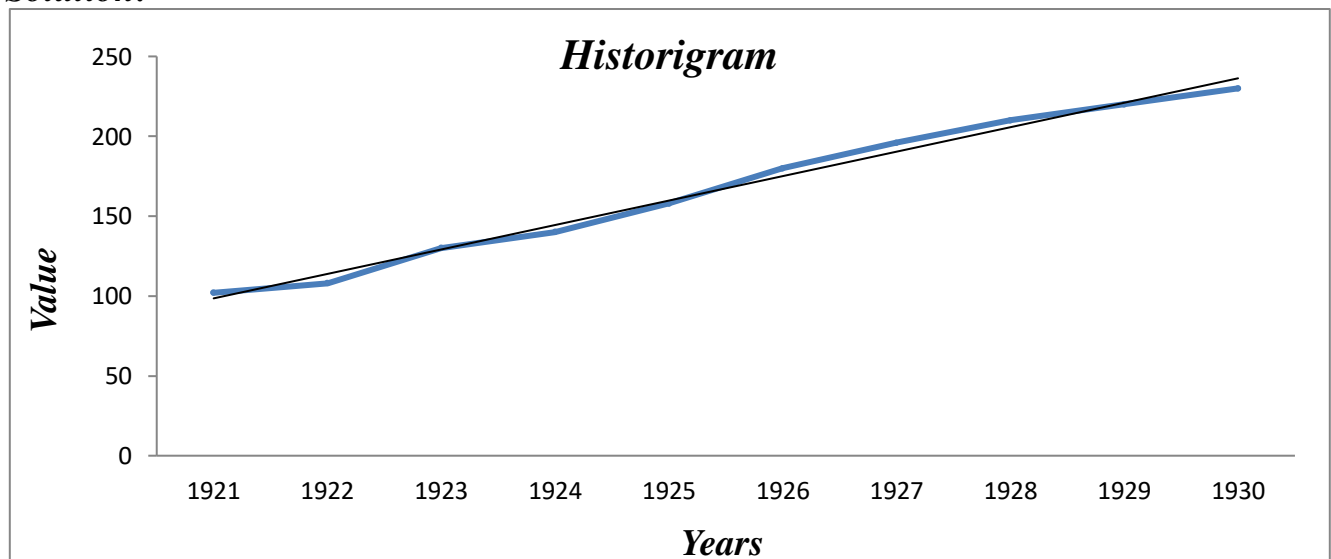
<i>Year</i>	<i>Wages</i>	<i>Total</i>	<i>Averages</i>
1970	140	1098	183
1971	148		
1972	180		
1973	195		
1974	200		
1975	235	1825	304.17
1976	260		
1977	280		
1978	290		
1979	330		
1980	325		
1981	340		

Example:

Compute and insert on graphs the semi-average trends for the following series:

<i>Year</i>	<i>1921</i>	<i>1922</i>	<i>1923</i>	<i>1924</i>	<i>1925</i>	<i>1926</i>	<i>1927</i>	<i>1928</i>	<i>1929</i>	<i>1930</i>
<i>Value</i>	102	108	130	140	158	180	196	210	220	230

Solution:



ii. Semi-averages

Year	Wages	Total	Averages
1921	140	863	172.6
1922	148		
1923	180		
1924	195		
1925	200		
1926	235	1395	279
1927	260		
1928	280		
1929	290		
1930	330		

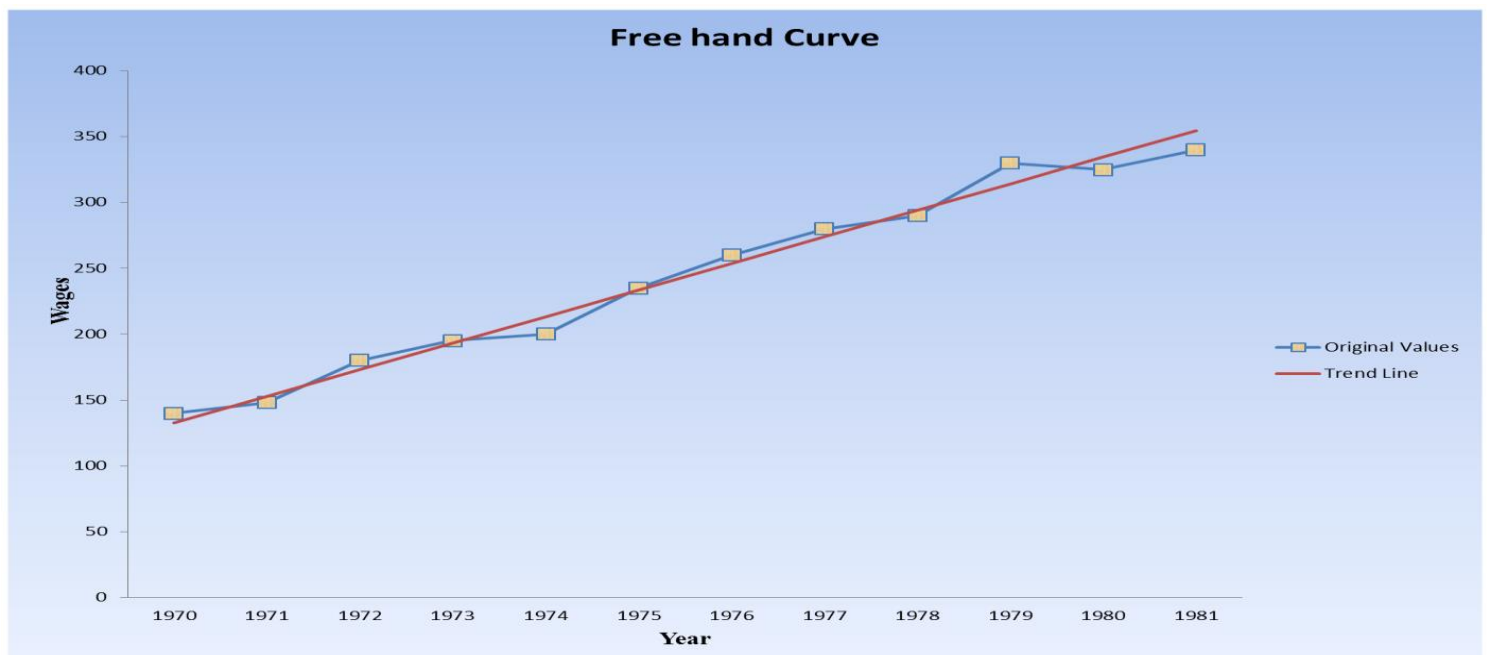
Example:

Plot the following data showing average wages in rupees of some workers during 1970-1981 and find the trend by the method of

- Free-hand curve
- Semi-averages

Year	1970	1971	1972	1973	1974	1974	1975	1976	1977	1978	1979	1980
Wag-es	140	148	180	195	200	235	260	280	290	330	325	340

i. Free-hand curve



ii. Semi-averages

Year	Wages	Total	Averages	Trend Values
1970	140	1098	183	132.5
1971	148			152.7
1972	180			172.9
1973	195			193.1
1974	200			213.3
1975	235			233.5
1976	260	1825	304	253.7
1977	280			273.9
1978	290			294.1
1979	330			314.3
1980	325			334.5
1981	340			354.7

Method of Moving Average:

- ▶ The k-period moving averages are defined as the averages calculated by using the k consecutive values of the observed series, then repeating the operation by dropping one value at the beginning and including the first value after the preceding total, and so on, moving on one value to calculate each successive average.
- ▶ Each k-period moving average is placed against the middle of its time period.
- ▶ The average will directly to the observed value in the series when k is odd and when k is even, the moving average will be placed at points which will be located between two periods.
- ▶ For the purposes of illustration, let us first choose k to be odd, say k=3 years, then we compute the 3-year moving averages as

$$a_1 = \frac{1}{3} (Y_1 + Y_2 + Y_3)$$

$$a_2 = \frac{1}{3} (Y_2 + Y_3 + Y_4)$$

$$a_3 = \frac{1}{3} (Y_3 + Y_4 + Y_5)$$

and so on. The averages so obtained are placed opposite the middle year of each group, i.e. opposite the 2nd year, 3rd year, 4th year and so on.

- ▶ Next, we choose k to be even, say k=4 years. Then the 4-years moving averages are computed as

$$a_1 = \frac{1}{4}(Y_1 + Y_2 + Y_3 + Y_4)$$

$$a_2 = \frac{1}{4}(Y_2 + Y_3 + Y_4 + Y_5)$$

$$a_3 = \frac{1}{4}(Y_3 + Y_4 + Y_5 + Y_6)$$

and so on.

- ▶ A 4-year centered moving average

$$\begin{aligned} a'_1 &= \frac{1}{2} \left[\frac{1}{4}(Y_1 + Y_2 + Y_3 + Y_4) + \frac{1}{4}(Y_2 + Y_3 + Y_4 + Y_5) \right] \\ &= \frac{1}{8} [Y_1 + 2Y_2 + 2Y_3 + 2Y_4 + Y_5] \end{aligned}$$

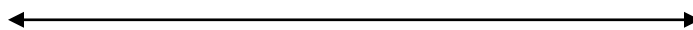
That is, a 4-year centered moving average. These averages are plotted on the graph of original values and the line connecting these points is the moving average trend.

Advantages of Moving Average:

- ▶ The method of moving averages is easy and simple.
- ▶ The moving averages of appropriate period estimate the combined effects of trend and cyclical components and give a smooth version of the series by removing the seasonal and other effects.

Disadvantages of Moving Average:

- ▶ The moving average trend does not provide values at the end or the beginning of the original series by half the period.
- ▶ The simple moving average are affected by large Y-values. It may be reduced by using the geometric mean.
- ▶ In the absence of an appropriate period, moving averages may make the resulting series more cyclical than the observed series.
- ▶ The method does not provide the mathematical expression for the trend.



Lecture: 08

In the Current Lecture:

- ▶ Analysis of Moving Average
- ▶ Method of Least Square

In the Previous Lecture:

- ▶ Time Series Decomposition
- ▶ Analysis of Secular Trend
- ▶ Moving Average

Example:

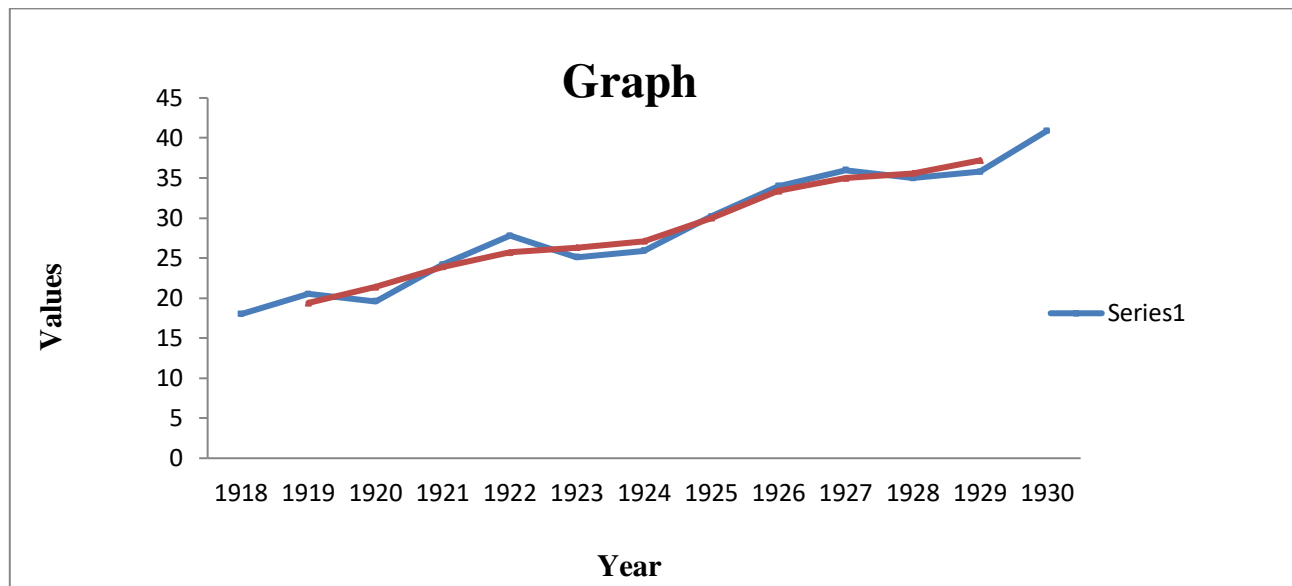
Compute “3-year moving averages” for the following data and show the moving average trends on the graph.

<i>Year</i>	<i>1918</i>	<i>1919</i>	<i>1920</i>	<i>1921</i>	<i>1922</i>	<i>1923</i>	<i>1924</i>	<i>1925</i>	<i>1926</i>	<i>1927</i>	<i>1928</i>	<i>1929</i>	<i>1930</i>
<i>Value</i>	18.0	20.5	19.6	24.2	27.8	25.1	25.9	30.2	34.0	36.0	35.0	35.8	40.9

Solution:

3-year Moving Average

<i>Year</i>	<i>Values</i>	<i>Total</i>	<i>Average</i>
<i>1918</i>	18.0		
<i>1919</i>	20.5	58.1	19.4
<i>1920</i>	19.6	64.3	21.4
<i>1921</i>	24.2	71.6	23.9
<i>1922</i>	27.8	77.1	25.7
<i>1923</i>	25.1	78.8	26.3
<i>1924</i>	25.9	81.2	27.1
<i>1925</i>	30.2	90.1	30.0
<i>1926</i>	34.0	100.2	33.4
<i>1927</i>	36.0	105.0	35.0
<i>1928</i>	35.0	106.8	35.6
<i>1929</i>	35.8	111.7	37.2
<i>1930</i>	40.9		



Example:

Compute (i) 5-year moving averages and (ii) 7-year moving averages for the following data and show the moving average trends on the graph.

Year	1918	1919	1920	1921	1922	1923	1924	1925	1926	1927	1928	1929	1930
Value	18.0	20.5	19.6	24.2	27.8	25.1	25.9	30.2	34.0	36.0	35.0	35.8	40.9

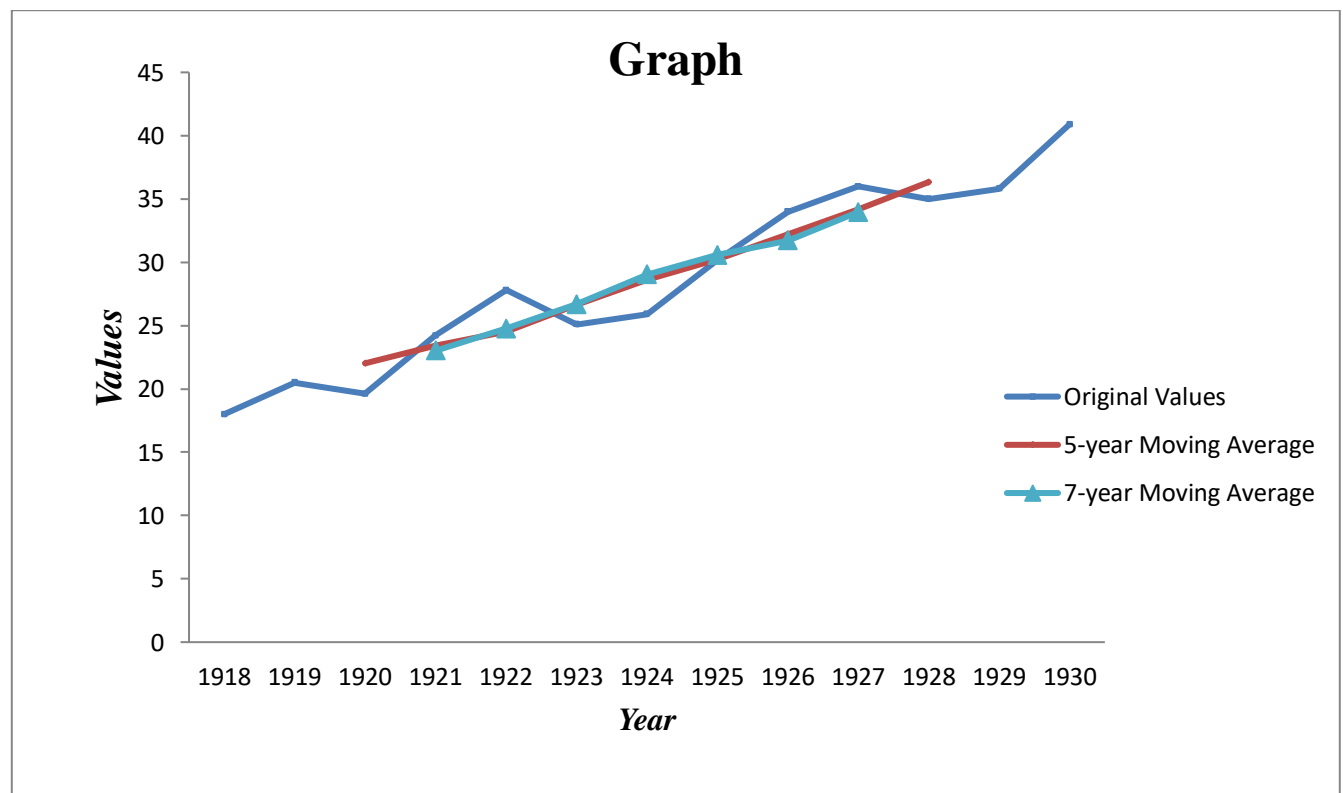
Sloution:

1. 5-year Moving Average

Year	Values	Total	Average
1918	18.0		
1919	20.5		
1920	19.6	110.1	22.0
1921	24.2	117.2	23.4
1922	27.8	122.6	24.5
1923	25.1	133.2	26.6
1924	25.9	143.0	28.6
1925	30.2	151.2	30.2
1926	34.0	161.1	32.2
1927	36.0	171.0	34.2
1928	35.0	181.7	36.3
1929	35.8		
1930	40.9		

2. 7-year Moving Average

<i>Year</i>	<i>Values</i>	<i>Total</i>	<i>Average</i>
1918	18.0		
1919	20.5		
1920	19.6		
1921	24.2	161.1	23.0
1922	27.8	173.3	24.8
1923	25.1	186.8	26.7
1924	25.9	203.2	29.0
1925	30.2	214.0	30.6
1926	34.0	222.2	31.7
1927	36.0	237.8	34.0
1928	35.0		
1929	35.8		
1930	40.9		



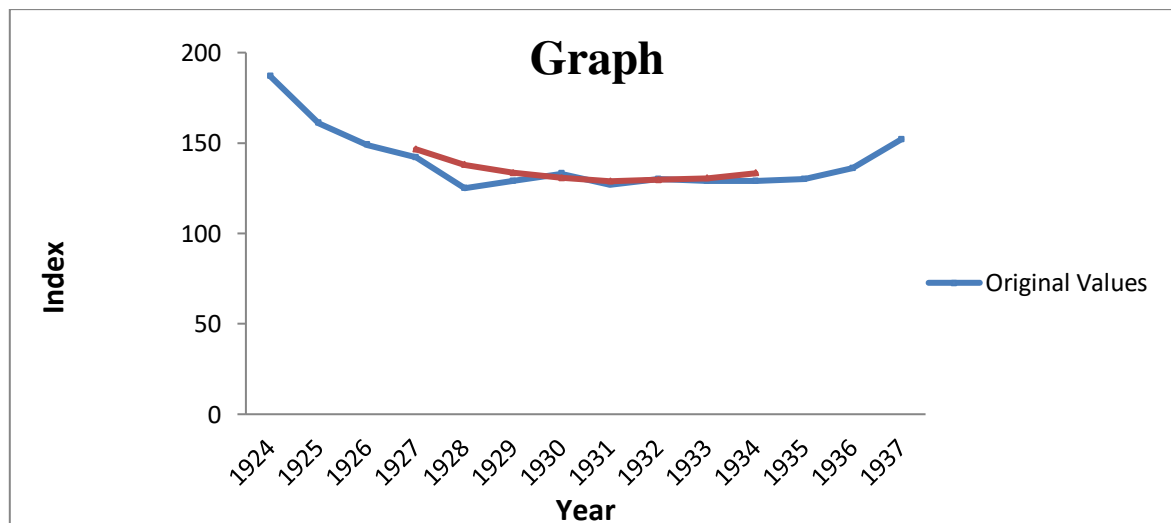
Example:

Determine a line in the following series of values by the use of a simple average of seven consecutive terms.

Year	1924	1925	1926	1927	1928	1929	1930	1931	1932	1933	1934	1935	1936	1937
Index	187	161	149	142	125	129	133	127	130	129	129	130	136	152

Solution:**7-year Moving Average**

Year	Values	Total	Average
1924	187		
1925	161		
1926	149		
1927	142	1026	146.6
1928	125	966	138.0
1929	129	935	133.6
1930	133	915	130.7
1931	127	902	128.9
1932	130	907	129.6
1933	129	914	130.6
1934	129	933	133.3
1935	130		
1936	136		
1937	152		



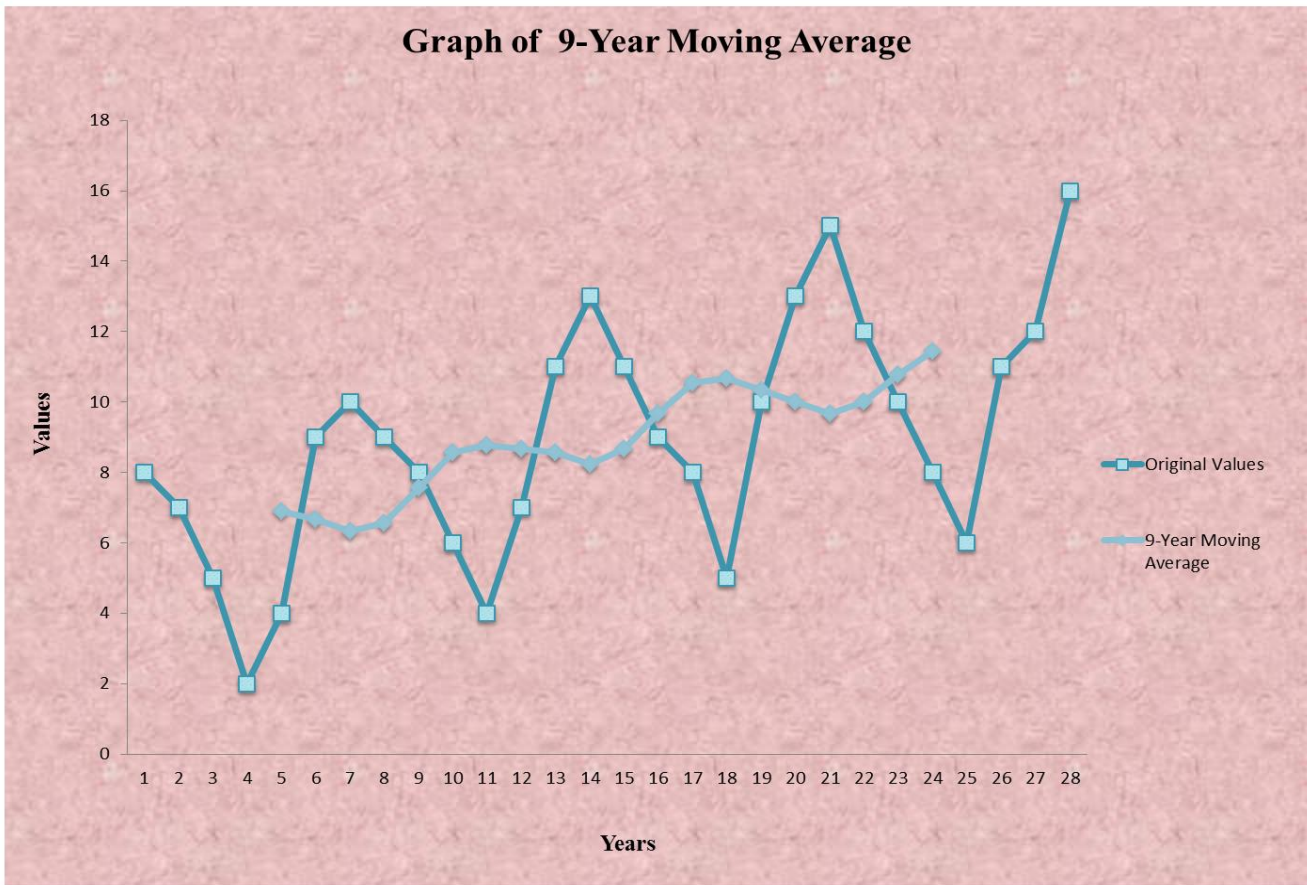
Example:

Find out and plot the nine-year moving average for the following series:

8, 7, 5, 2, 4, 9, 10, 9, 8, 6, 4, 7, 11, 13, 11, 9, 8, 5, 10, 13, 15, 12, 10, 8, 6, 11, 12, 16.

Solution:

Year	Value	9-Year Moving	
		<i>Total</i>	<i>Average</i>
1	8		6.9
2	7	62	6.7
3	5	60	6.3
4	2	57	6.6
5	4	59	7.6
6	9	68	8.6
7	10	77	8.8
8	9	79	8.7
9	8	78	8.6
10	6	77	8.2
11	4	74	8.7
12	7	78	9.7
13	11	87	10.6
14	13	95	10.7
15	11	96	10.3
16	9	93	10.0
17	8	90	9.7
18	5	87	10.0
19	10	90	10.8
20	13	97	11.4
21	15	103	
22	12		
23	10		
24	8		
25	6		
26	11		
27	12		
28	16		



Example:
 Compute 4-month centered moving average from the following:
 23, 26, 28, 30, 31, 35, 37, 32, 34, 38.

Solution:

Month	Values	4-Month Moving Total	4-Month Centered Moving Totals	4-Month Centered Moving Averages
1	23	107	222	27.75
2	26	115	239	29.88
3	28	124	257	32.12
4	30	133	268	33.50
5	31	135	273	34.12
6	35	138	279	34.88
7	37	141		
8	32			
9	34			
10	38			

Example:

The estimated number of visitors ('00s) at a holiday resort were as follows:

Year	1971	1972	1973	1974	1975	1976	1977	1978	1979	1978
Visitor	31	49	74	62	65	73	70	84	86	79

Calculate the 2-year centered moving average for the above following table:

Solution:

Year	Visitors	2-year Moving totals	2-year centered moving totals	2-year centered moving average
1971	31	80	203	50.75
1972	49	123	259	64.75
1973	74	136	263	65.75
1974	62	127	265	66.25
1975	65	138	281	70.25
1976	73	143	297	74.25
1977	70	154	324	81.00
1978	84	170	335	83.75
1979	86	165		
1980	79			

Method of Least Squares:

- ▶ A trend line, which can be described by a mathematical equation of the form of a straight line, a parabola, or an exponential, can be measured using the method of least squares, by letting time to be the independent variable.
- ▶ Let us suppose that the long-term trend appears to be linear. Then the equation of the least squares linear trend would be

$$\hat{Y}_i = a + bt$$

the values of a and b, the two constants, are determined by solving the following two normal equations simultaneously:

$$\sum Y = na + b \sum t,$$

$$\sum ty = a \sum t + b \sum t^2$$

- ▶ The trend values are computed from this equation by substituting the values of t and are plotted on the graph of original values.
- ▶ To simplify computations, the time variable (t) is coded by taking the time deviations from the middle point of the periods and the coded time period is denoted by X.
- ▶ When the number of years is odd, the middle year is taken as 0 or its origin.
- ▶ When the number of years is even, X=0 is taken at the middle of two middle years.

Odd number of years		Even number of years	
Year	Coded year	Year	Coded year
1978	-1	1978	-1
1979	0	1979	-0.5
1980	1	1980	0.5
		1981	1

- ▶ Sometimes, the numbers 1, 2, 3, ..., may also be assigned to X. The coded form of the linear trend would be $Y_t = a + bX$ and the two-normal equations would reduce to :

$$\sum Y = na \text{ and } \sum XY = b \sum X^2$$
- ▶ When the long-term trend of a time series graph appears to be a curve, we fit a parabola curve of the second or third degree or a curve of some other type such as exponential, modified exponential, Gompertz, logistic, etc. It should be kept in mind that a trend has the smaller $\sum (Y - Y_t)^2$ is the better fit.

Advantages of Least Squares Method:

- ▶ The method of least square is mathematically sound and is therefore useful for prediction purposes.
- ▶ It is also an objective method.

Disadvantages of Least Squares Method:

- ▶ It is sometime difficult to choose a proper mathematical equation.
- ▶ In case of non-linear trends, the method may involve heavy calculations.

Example:

For the following time series, determine the trend line by using the method of least squares for fitting a straight line. Also plot the actual values and trend values on the graph.

Year	1970	1971	1972	1973	1974	1975	1976	1977	1978
Value	12	23	37	48	41	37	48	61	70

Solution:

Two Normal Equations:

$$a = \frac{\sum Y}{n}$$

$$b = \frac{\sum XY}{\sum X^2}$$

Year (t)	Coded year (X)	Y	XY		Trend
1970	-4	12	-48	16	18.09
1971	-3	23	-69	9	24.04
1972	-2	37	-74	4	29.99
1973	-1	48	-48	1	35.94
1974	0	41	0	0	41.89
1975	1	37	37	1	47.84
1976	2	48	96	4	53.79
1977	3	61	183	9	59.74
1978	4	70	280	16	65.69
Total	0	377	357	60	377.01

Two Normal Equations

$$a = \frac{\sum Y}{n} = \frac{377}{9} = 41.89$$

$$b = \frac{\sum XY}{\sum X^2} = \frac{357}{60} = 5.95 \quad Y_t = 41.89 + 5.95X$$

Graph

