TIME SERIES

"A time series may be defined as a collection of readings belonging to different time periods of some economic variable or composite of variables."

Mathematically, a time series is defined by the functional relationship

$$y_t = f(t)$$

Where y_t is the value of the phenomenon (or variable) under consideration at time t. For example,

(i) The population (y_t) of a country or a place in different years (t),
(ii) The number of births and deaths (y_t) in different months (t) of the year,
(iii) The sale (y_t) of a departmental store in different months (t) of the year,
(iv) The temperature (y_t) of a place on different days (t) of the week, and so on,

constitute time series. Thus, if the values of a phenomenon or variable at times $t_1, t_2, \ldots, t_n$ are $y_1, y_2, \ldots, y_n$ respectively, then the series

t:	t_1	t_2	t_3	……	t_n
y_t:	y_1	y_2	y_3	……	y_n

constitute a time series.

Components of time series

The various forces at work, affecting the values of a phenomenon in a time series, can be broadly classified into the following four categories, commonly known as the components of a time series, some or all of which are present (in a given time series) in verifying degrees.

(a) Secular Trend or Long-term Movement.
(b) Periodic Changes for Short-term Fluctuations.
 (i) Seasonal variations, and (ii) cyclic variations.
(c) Random or Irregular Movements.

(A) **Trend**. Secular trend or simply trend we mean the general tendency of the data to increase or decrease during a long period of time. This is true of most of series of business and economics statistics. For example, and upward tendency would be seen in data pertaining (संबंधित) to population, agricultural production, currency in circulation etc., While a downward a tendency will be noticed in data of births and deaths, epidemics etc.,

It may be clearly noted that trend is the general, smooth, long term average tendency.

Linear and nonlinear (curvi – linear) trend. if the time-series values plotted on graph cluster more, or less, around a straight line then the trend exhibited by the time series is termed as linear otherwise nonlinear (curvi – linear). In a straight-line trend, the time series values increase or decrease more or less by a constant absolute amount, I.e., the rate of growth (or decline) is constant. In an economic and business phenomenon, the rate of growth or decline is not of constant nature through but varies considerably in different sectors of time.

The long term 'long period of time' is a relative term and cannot be defined exactly. In some cases, a period as small as a week may be fairly long while in some cases, a period as long as two years may not be enough.

(B) **Periodic changes**. It would be observed that in many social and economic phenomena (घटना), apart from the growth factor in a time series there are forces at work which prevent the smooth flow of the series in a particular direction and tend to repeat themselves over a period of time. These factors do not act continuously but operate in a regular spasmodic (मरोड़-संबंधी) manner. The resultant effect of such process may be classified as:

(i) seasonal variations, and (ii) Cyclic variations

1.. Seasonal variations. These variations in a time series are due to the rhythmic forces which operate in a regular and periodic manner over a span of less than a year, i.e., During a period of 12 months and have the same or almost same pattern year after year. Thus, seasonal variations in a time series will be there if the data are recorded quarterly (every three months), monthly, weekly, daily, hourly, and so on. Thus, in a time series data where only annual figures are given, there are no seasonal variations. These seasonal variations may be attributed to the following two causes:

(i) **Those resulting from natural forces.** As the name suggests, the various seasons or weather conditions and climatic changes play an important role in seasonal movement. For instance, the sale of umbrellas pick up very fast in rainy season, the demand for electric fans goes up in summer season, the sale of ice and ice cream increases very much in summer, the sale of woollens go up in winter all being affected by natural forces.

(ii) **Those resulting from man-made conventions**. These variations in a time series within a period of 12 months are due to habits, fashions, customs (रिवाज) and conventions (सम्मेलन, परंपरा) of the people in the society. For instance, the sale of jewellery and ornaments goes up in marriages.

2.. Cyclic variations. The oscillatory movements in a time series with period of oscillation more than one year are termed as cyclic fluctuations. One complete period is called a 'cycle'. The cyclic movements in a time series are generally attributed to the so-called 'Business Cycle', which may also be referred to as the 'Four-Phase Cycle' composed of prosperity (समृद्धि) (period of boom), recession (व्यापारिक मंदी), depression (न्यूनता) and recovery (उगाही).

Most of the economic and commercial series, for example, series relating to prices, production and wages etc., Are affected by business cycles. Cyclic fluctuations, through more or less regular are not periodic.

(C) **Irregular (or random) Component.** Apart from the regular variations, almost all the series contains another factor called the random or irregular or residual fluctuations, which are not accounted for by secular trend and seasonal and cyclic variations. These fluctuations are purely random, erratic (अनियमित), unforseen (पहले से न सोचा हुआ), unpredictable and are due to numerous non – recurring and irregular circumstances which are beyond the control of human hand but at the same time are a part of our system such as earthquakes, wars, floods, Revolutions, epidemics, etc.

Analysis of time series

The main problems in the analysis of time series are:

(i) To identify the forces or components at work, the net effect of whose interaction is exhibited by the movement of a time series, and

(ii) To isolate, study, analyses and measure them independently, i.e., by holding and other things constant.

Mathematical models for time series. The following are the two models commonly used for the decomposition of a time series into its components.

1. **Decomposition by additive hypothesis (or additive model).** According to the additive model, A time series can be expressed as:

$$y_t = T_t + S_t + C_t + R_t$$

Where y_t is the time series value at time t, T_t represents the trend value, S_t, C_t and R_t represent the seasonal, cyclic and random fluctuations at time t.

Obviously, the term S_t will not appear in a series of annual data. The additive model implicitly implies that seasonal forces (in different years), cyclic forces (in different cycles) and regular forces (in different long term period) operate with equal absolute effect irrespective of the trend value. As such C_t (and S_t) will have positive or negative values, according as whether we are in an above normal or below normal phase of the cycle (and year) and the total of positive and negative values for any cycle (and any year) will be zero. R_t will also have positive or negative value and in the long-term ($\sum R_t$) close will be zero. Occasionally, there may be a few isolated occurrences of extreme R_t of episodic nature.

The additive model assumes that all the four components of the time series operate independently of each other so that none of these components has any effect on the remaining three.

2. **Decomposition by multiplicative hypothesis (or multiplicative model).** On the other hand, if we have reasons to assume that the various components in a time series operate proportionately to the general level of the series, the traditional or classical multiplicative model is appropriate. According to the multiplicative model,

$$y_t = T_t \times S_t \times C_t \times R_t$$

where S_t, C_t and R_t, instead of assuming positive and negative value, are indices fluctuating above or below unity and the geometric means of S_t in a year, C_t in a cycle and R_t in a long term period are unity. In a time series with both positive and negative values, the multiplicative model cannot be applied unless the time series is translated by adding a suitable positive value. It may be pointed out that the multiplicative decomposition of a time series is same as the additive decomposition of logarithmic values of the original time series i.e.,

$$\log y_t = \log T_t + \log S_t + \log C_t + \log R_t$$

In practice, most of the series related to economic data confirm to multiplicative model.

Mixed model. in addition to the additive and multiplicative models discussed above, the components in a time series may be combined in a large number of other ways. The different models, defined under different assumptions will yield different results. Some of the mixed models resulting from different combinations of additive and multiplicative models are given below:

$$y_t = T_t C_t + S_t R_t$$

$$y_t = T_t + S_t C_t R_t$$

$$y_t = T_t + S_t + C_t R_t$$

Uses of time series. The time series analysis is of greater importance not only to businessman or economist but also to people working in various disciplines in natural, social and physical sciences. Some of its uses are enumerated below:

1. It enables us to study the past behaviour of the phenomenon under consideration. i.e., To determine the type and nature of the variations in the data.
2. The segregation and study of the various components of is of paramount importance to a businessman in the planning of future operations and in the formulation of executive and policy decisions.
3. It helps to compare the actual current performance of accomplishments (उपलब्धि) with the expected ones (on the basis of the past performances) and analyse the cause of such variations, if any.
4. IT enables us to predict or estimate or forecast the behaviour of the phenomenon in future which is very essential for business planning.
5. It helps us to compare the changes in the values of different phenomenon at different time or places, etc.

Measurement of Trend

trend can be studies and/or measured by the following methods:

(i) Graphic (or Free Hand Curve Fitting) Method
(ii) Method of Semi Averages,
(iii) Method of Curve Fitting by Principle of Least Squares, and
(iv) Method of Moving Averages.

We shall now discuss each of the three methods in detail.

1. **Graphic method**. A free-hand smooth curve obtained on plotting the values y_t against t, enables us to form an idea about the general trend of

the series. Smoothly smoothing of the curve eliminates other components, such as regular and irregular fluctuations.

Merits.

This method does not involve any complex mathematical technique and can be used to describe all types of trend linear and non-linear. Thus, simplicity and flexibility are strong points of this method.

Demerits.

1. In its main drawbacks are the method is very subjective i.e., The bias of the person handling the data plays a very important role and as such different trend curves will be obtained by different persons for the same set of data. As such 'trend by inspection' should be attempt only by skilled and experienced statisticians and this limits the utility and popularity of this method.
2. It does not enable us to measure trend.

2. **Method of semi averages**. In this method, the whole data is divided into two parts with respect to time, for example, if we are given y_t for t from 1991 – 2000, i.e., over a period of 12 years, the two equal parts will be the data from 1991 to 1996 and 1997 to 2002. In case of odd number of years the two parts are obtained by omitting the value corresponding to the middle year, for example, the data from 1991 – 2001 the value corresponding to the middle year, such as 1996 being omitted. Next we compute the arithmetic mean for each part and plot these two averages (means) close against the mid value of the respective time periods covered by each part. The line obtained by joining these two points is the required trend line and may be extended both ways to estimate intermediate of future values.

Merits

1. As compared with graphic method, the way the obvious advantage of this method is its objectivity in the sense that everyone who applies it would get the same result. Moreover, we can also estimate the trend values.
2. It is really comprehensible as compared to the method of 'least squares' or the 'moving average method'.

Limitations.

1. This method assumes linear relationship between the plotted points which may not exist. Moreover, the limitations of arithmetic mean as an average also stand in its way.

Que 1. Fit a trend line to the following data by the method of semi averages:

Year	Bank Clearance (Rs. Crores)	Year	Bank Clearance (Rs. Crores)
1992	53	1999	87
1993	79	2000	79
1994	76	2001	104
1995	66	2002	97
1996	69	2003	92
1997	94	2004	101
1998	105		

Solution. Here, since n = 13 (odd)

The two parts consists 1992 to 1997 and 1999 to 2004

1998 is omitted.

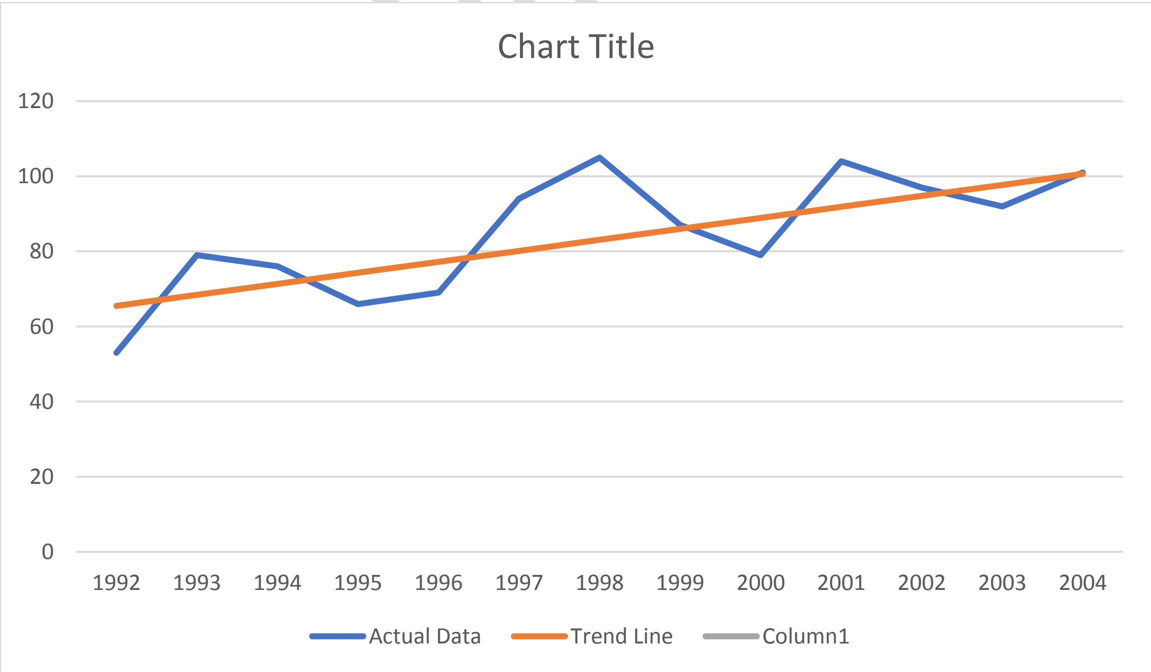
$\bar{X}_1 = average\ sales\ of\ first\ part$

$\bar{X}_1 = \frac{437}{6} = 72.83$ (Rs. Crores)

And $\bar{X}_2 = average\ sales\ of\ second\ part$

$\bar{X}_2 = \frac{560}{6} = 93.33$ (Rs. Crores)

$\bar{X}_1$ and $\bar{X}_2$ is plotted in the following diagram



3.Method of Curve Fitting by Principal of Least Squares. The principal of least squares is the most popular and widely used method of fitting mathematical functions to a given set of data. An examination of the plotted data often provides an adequate basis for deciding upon the type of trend to use. The various types of curves that may be used to describe the given data in practice are:

(if y_t is the value of the variable corresponding to time t)

(i) A straight line: $y_t = a + bt$

(ii) Second degree parabola: $y_t = a + bt + ct^2$

(iii) K[th] degree parabola: $y_t = a_0 + a_1 t + a_2 t^2 + \ldots\ldots + a_k t^k$

(iv) Exponential Curves: $y_t = ab^t$

$$\Rightarrow \quad \log y_t = \log a + t \log b$$

$$\Rightarrow \quad Y = A + B\, t$$

(v) Second degree curve fitted to logarithms: $y_t = a\, b^t\, c^{t^2}$

$$\Rightarrow \quad \log y_t = \log a + t \log b + t^2 \log c$$

$$\Rightarrow \quad Y = A + B\, t + C\, t^2$$

Fitting of Straight Line by Least Squares Method. let the straight line trend between the given time series values (y_t) and time t be given by the equation:

$$y_t = a + bt$$

Principal of least square consists in minimising the sum of squares of the deviations between the given values of y_t and their estimates given by above equation. In other words, we have to find a and b such that for given values of y_t corresponding to n different values of t,

$$E = \sum_t (y_t - a - b\, t)^2$$

Is minimum. For a maxima or minima of E, for variation in a and b, we should have

$$\left.\begin{aligned} \frac{\partial E}{\partial a} &= 0 = -2 \sum (y_t - a - b\, t) \\ \frac{\partial E}{\partial b} &= 0 = -2 \sum t\,(y_t - a - b\, t) \end{aligned}\right\} \Rightarrow \begin{aligned} \sum y_t &= na + b \sum t \\ \sum t\; y_t &= a \sum t + b \sum t^2 \end{aligned}$$

Which are the normal equations foe estimating a and b.

Fitting of Second Degree (Parabola) Trend. Let the second degree parabolic trend curve be:

$$y_t = a + bt + ct^2$$

Proceeding similarly as in the case of a straight line, the normal equations for estimating a, b and c are given by:

$$\sum y_t = na + b \sum t + c \sum t^2$$
$$\sum t\, y_t = a \sum t + b \sum t^2 + c \sum t^3$$
$$\sum t^2\, y_t = a \sum t^2 + b \sum t^3 + c \sum t^4$$

The summation being taken over the values of the time series.

Fitting of Exponential Curve: $y_t = ab^t$

$$\Rightarrow \quad \log y_t = \log a + t \log b$$
$$\Rightarrow \quad Y = A + B\, t \qquad \ldots (1)$$

Where, $Y = \log y_t$, $A = \log a$ and $B = \log b$

Equation (1) is a straight line in t and Y and thus the normal equations for estimating A and B are

$$\sum Y = n\, A + B \sum t$$
$$\sum t\; Y = A\, t + B \sum t^2$$

On solving these two equations, we get the values of A and B, and

$$a = antilog\ A, \quad and\ \ b = antilog\ B$$

Merits and Drawbacks of Trend fitting by the Principal of Least Squares.

Merits. The method of least squares is the most popular and widely used method of fitting mathematical functions to a given set of observations. It has the following advantages:

1. Because of its mathematical or analytical character, this method completely eliminates the element of subjective judgement or personal bias on the part of the investigator.

2. Unlike the method of moving averages, this method enables us to compute the trend values for all the given time periods in the series.
3. The trend equation can be used to estimate or predict the values of the variable for any period t in future or even in the intermediate periods of the given series and the forecast values are also quite reliable.

Drawbacks.

1. The method is quite tedious and time consuming as compared with other methods. It is rather difficult for a non – mathematical person to understand and use.
2. The addition of even a single new observation necessitates all calculations to be done afresh.
3. The most serious limitation of this method is the determination of the type of trend curve to be fitted, foe example, we should fit a linear or a parabolic trend or some other more complicated trend curve.

Que 2. In a certain industry, the production of a certain commodity (in '000 units) during the years 1994 – 2004 is given in the adjoining table:

Year	Production (in '000 units)	Year	Production (in '000 units)
1994	66.6	2000	93.2
1995	84.9	2001	111.6
1996	88.6	2002	88.3
1997	78.0	2003	117.0
1998	96.8	2004	115.2
1999	109.9		

(i) Graph the data.

(ii) Obtain the least square line fitting the data and construct the graph the trend line.

(iii) Compute the trend values for the year 1994 – 2004 and estimate the production of commodity during the years 2005 and 2006, if the present trend continues.

Solution. Let we shift the origin to 1999

$$x = t - 1999$$

The normal equation for fitting of trend line $y_t = a + b\,x$ are given by:

$$\sum y_t = na + b \sum x$$
$$\sum x\ y_t = a \sum x + b \sum x^2$$

Year (t)	Production ('000 units) (y_t)	x	$x\ y_t$	x^2
1994	66.6	-5	-333.0	25
1995	84.9	-4	-339.6	16
1996	88.6	-3	-265.8	9
1997	78.0	-2	-156.0	4
1998	96.8	-1	-96.8	1
1999	109.9	0	0	0
2000	93.2	1	93.2	1
2001	111.6	2	223.2	4
2002	88.3	3	264.9	9
2003	117.0	4	468.0	16
2004	115.5	5	576.0	25
Total	**1045.7**	**0**	**434.1**	**110**

Now the normal equations can be re – written as

$$1045.7 = 11a + b\,.0$$
$$434.1 = 0\,.\ a + 110\,b$$

On solving these two equation we get

$$a = 95.06 \qquad and \qquad b = \ 3.94$$

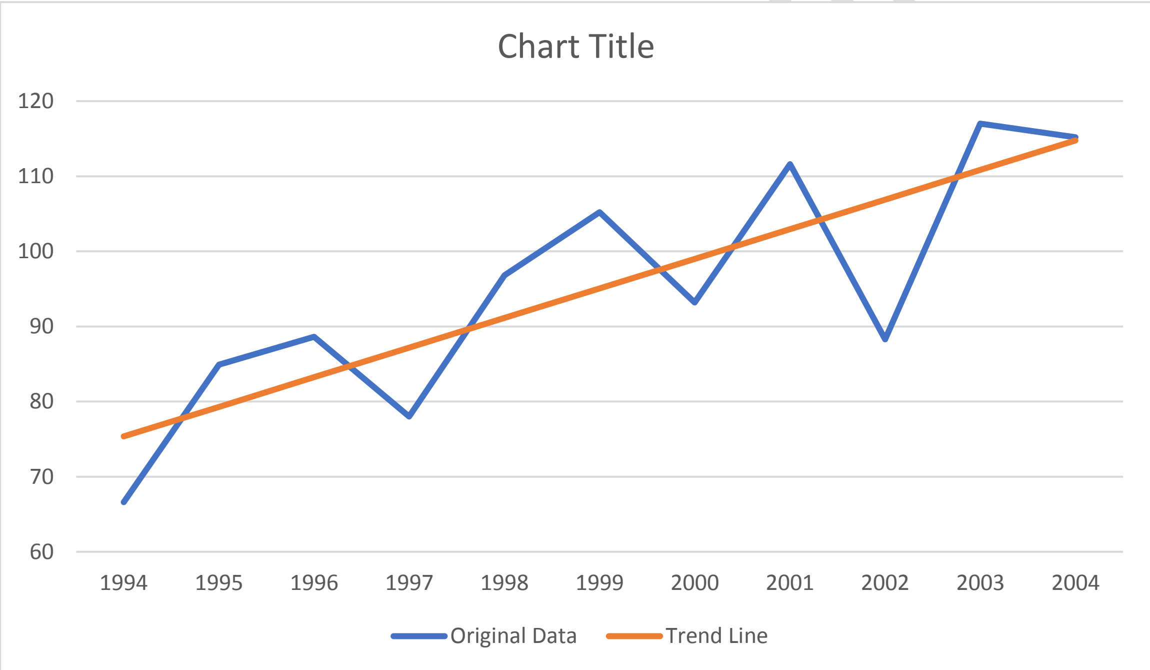
The equation of trend line is given by

$$y_t = 95.06 + 3.94\ x$$

Now trend values are given by

Year (t)	Production ('000 units) (y_t)	x	Trend values $y_t = 95.06 + 3.94\ x$
1994	66.6	-5	75.36
1995	84.9	-4	79.30

1996	88.6	-3	83.24
1997	78.0	-2	87.18
1998	96.8	-1	91.12
1999	109.9	0	95.06
2000	93.2	1	99.00
2001	111.6	2	102.94
2002	88.3	3	106.88
2003	117.0	4	110.82
2004	115.5	5	114.76
2005		6	118.70
2006		7	122.64



Que 3. Fit a straight line trend by the method of least squares to the following data relating to the scales of a leading department store. Assuming that the same rate of change continues, what would be predicted earnings for the year 2006?

Year:	**1997**	**1998**	**1999**	**2000**	**2001**	**2002**	**2003**	**2004**
Scales (Crores Rs.)	**76**	**80**	**130**	**144**	**138**	**120**	**174**	**190**

Solution: Trend line $y_t = 124.17 + 14.66\,x$,

$$\hat{y}_e(2006) = 212.13\ (crores\ Rs.)$$

Que 4. The following figures are the production data of a certain factory manufacturing air conditioners:

Year:	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
Production ('000 unit)	17	20	19	26	24	40	35	55	51	74	79

Fit a second degree parabolic trend curve to the above data and obtain the trend values.

Solution. Let we shift the origin to 1995

$$x = t - 1995$$

Let the second degree parabolic curve be $y_t = a + bx + cx^2 \quad \dots\dots(1)$

The normal equations for fitting curve (1) are given by

$$\sum y_t = na + b\sum x + c\sum x^2$$
$$\sum x\,y_t = a\sum x + b\sum x^2 + c\sum x^3$$
$$\sum x^2\,y_t = a\sum x^2 + b\sum x^3 + c\sum x^4$$

Year (t)	Production ('000 units) y_t	x	x^2	x^3	x^4	xy_t	$x^2\,y_t$
1990	17	-5	25	-125	625	-85	425
1991	20	-4	16	-64	256	-80	320
1992	19	-3	9	-27	81	-57	171
1993	26	-2	4	-8	16	-52	104
1994	24	-1	1	-1	1	-24	24
1995	40	0	0	0	0	0	0
1996	35	1	1	1	1	35	35
1997	55	2	4	8	16	110	220
1998	51	3	9	27	81	153	459
1999	74	4	16	64	256	296	1184
2000	79	5	25	125	625	395	1975

Total	**440**	**0**	**110**	**0**	**1958**	**691**	**4917**

Now the normal equations can be re – written as

$$440 = 11\,a + b\,.0\ + 110\,c$$
$$691\ = 0\,.a\ + 110\,.b\ + 0.\ c$$
$$4917 = 110\,.a\ + 0\,.b\ + 1958\,.c$$

On solving these equations we get

$$a = \ 34 \quad ,\quad b = \ 6.28 \quad and \quad c = 0.60$$

Substituting the values of a, b and c in equation (1) the trend line represented by

$$y_t = 34 + 6.28\,x + 0.60\,x^2$$

Year (t)	Production ('000 units) y_t	x	x^2	Trend values $y_t = 34 + 6.28\,x + 0.60\,x^2$
1990	17	-5	25	17.6
1991	20	-4	16	18.48
1992	19	-3	9	20.56
1993	26	-2	4	23.90
1994	24	-1	1	28.32
1995	40	0	0	34.00
1996	35	1	1	40.88
1997	55	2	4	48.96
1998	51	3	9	58.24
1999	74	4	16	68.72
2000	79	5	25	80.40

Que 5. Fit a parabolic curve of second degree to the data given below and estimate the value for 2008 and comment on it:

Year:	**2002**	**2003**	**2004**	**2005**	**2006**
Sales ('000 Rs.)	**10**	**12**	**13**	**10**	**8**

Solution: parabolic trend $y_t = 12.314 - 0.6\,x - 0.857\,x^2$

$$\hat{y}_e(2008) = -3.798$$

Growth Curves and Their Fitting. The various growth curves, viz. the modified exponential, Gompertz and logistic curves cannot be determined by the principle of least squares. Special techniques have been devised for fitting these curves to the given set of data.

Modified Exponential Curve and its Fitting. As already pointed out modified exponential curve is given by

$$y_t = a + bc^t, \quad a > 0$$

Where y_t *represents* the time series value at the time t and a, b, and c are constants, called its parameters.

Taking the first difference $\Delta y_t = y_{t+h} - y_t$

$$= a + bc^{t+h} - a - bc^t$$

$$\Delta y_t = bc^t\,(c^h - 1)$$

Where h is the interval of differencing.

Similarly, $\Delta y_{t-h} = y_t - y_{t-h}$

$$= a + bc^t - a - bc^{t-h}$$

$$\Delta y_{t-h} = bc^{t-h}\,(c^h - 1)$$

Now,

$$\frac{\Delta y_t}{\Delta y_{t-h}} = \frac{bc^t\,(c^h - 1)}{bc^{t-h}\,(c^h - 1)} = c^h\,(constant)$$

Thus, the most striking feature of the modified exponential curve is that the first differences of the consecutive values of y_t corresponding to equivalent values of t change by a constant ratio.

Method of three Selected Points. We take three ordinates y_1, y_2 and y_3 corresponding to three equidistant values of t , (say) t_1, t_2 and t_3 respectively such that

$$t_2 - t_1 = t_3 - t_2$$

Now substituting in the curve

$$y_1 = a + bc^{t_1}, \quad y_2 = a + bc^{t_2} \quad and \quad y_3 = a + bc^{t_3}$$

$$\Rightarrow \quad y_2 - y_1 = a + bc^{t_2} - a + bc^{t_1} = b(c^{t_2} - c^{t_1}) = bc^{t_1}\,(c^{t_2 - t_1} - 1)$$

Similarly

$$y_3 - y_2 = bc^{t_2}\,(c^{t_3 - t_2} - 1)$$

On dividing

$$\frac{y_3 - y_2}{y_2 - y_1} = \frac{bc^{t_2}\,(c^{t_3 - t_2} - 1)}{bc^{t_1}\,(c^{t_2 - t_1} - 1)}$$

$$\frac{y_3 - y_2}{y_2 - y_1} = c^{t_2 - t_1}$$

$$c = \left(\frac{y_3 - y_2}{y_2 - y_1}\right)^{\frac{1}{t_2 - t_1}}$$

Substituting the value of c in $y_2 - y_1$

$$y_2 - y_1 = bc^{t_1}\,(c^{t_2 - t_1} - 1)$$

$$y_2 - y_1 = b\left(\frac{y_3 - y_2}{y_2 - y_1}\right)^{\frac{t_1}{t_2 - t_1}} \left(\frac{y_3 - y_2}{y_2 - y_1} - 1\right)$$

$$y_2 - y_1 = b\left(\frac{y_3 - y_2}{y_2 - y_1}\right)^{\frac{t_1}{t_2 - t_1}} \left(\frac{y_3 - 2y_2 + y_1}{y_2 - y_1}\right)$$

$$\frac{(y_2 - y_1)^2}{y_3 - 2y_2 + y_1} = b\left(\frac{y_3 - y_2}{y_2 - y_1}\right)^{\frac{t_1}{t_2 - t_1}}$$

$$b = \frac{(y_2 - y_1)^2}{y_3 - 2y_2 + y_1} . \left(\frac{y_2 - y_1}{y_3 - y_2}\right)^{\frac{t_1}{t_2 - t_1}}$$

Now we have

$$y_1 = a + bc^{t_1}$$

$$\Rightarrow \quad a = y_1 - bc^{t_1}$$

Substituting the value of b and c.

$$\Rightarrow \quad a = y_1 - \left[\frac{(y_2 - y_1)^2}{y_3 - 2y_2 + y_1} . \left(\frac{y_2 - y_1}{y_3 - y_2}\right)^{\frac{t_1}{t_2 - t_1}}\right] \left(\frac{y_3 - y_2}{y_2 - y_1}\right)^{\frac{t_1}{t_2 - t_1}}$$

$$\Rightarrow \qquad a = y_1 - \frac{(y_2 - y_1)^2}{y_3 - 2y_2 + y_1}$$

$$\Rightarrow \qquad a = \frac{y_1y_3 - 2y_1y_2 + y_1^2 - y_2^2 - y_1^2 + 2y_1y_2}{y_3 - 2y_2 + y_1}$$

$$\Rightarrow \qquad a = \frac{y_1y_3 - y_2^2}{y_3 - 2y_2 + y_1}$$

Substituting the value of a, b and c in $y_t = a + bc^t, we$ get our fitting of modified exponential curve fitting.

Moving average method.

It consists in measurement of trend by smoothing out the fluctuations of the data by means of a moving average. Moving average of extent (or period) m is a series of successive averages (arithmetic means) of m terms at a time, starting with 1st , 2nd , 3rd term etc. Thus, the first average is the mean of the 1st m terms, the 2nd is the mean of the m terms from 2nd to $(m + 1)th$ term, the 3rd is the mean of the m terms from 3rd to $(m + 2)th$ term and so on.

If m is odd $= (2k + 1)$ say, moving average is placed against the mid value of the time interval it covers, i.e., Against $t = k + 1$ and if m is even $= 2k$ (say) it is placed between the two middle values of the time interval it covers, i.e., Between $t = k$ and $t = k + 1$.

In this method, the main problem which is of paramount importance lies in determining the extent (or the period) of the moving average which will completely eliminate the oscillatory movements affecting the series. It has been established mathematically that if the fluctuations are regular and periodic then the moving average completely eliminates the oscillatory movements provided:

1. The extent of moving average is exactly equal to (or a multiple of) the period of oscillation, and

2.The trend is linear.

Merits: Moving average method is very flexible in the sense that the addition of a few more figures to the data simply results in some more trend values the previous calculation asked calculations are not affected at all.

Drawbacks.

1. It does not provide trend values for all the terms, e.g., for a moving average of extent $2k + 1$ we have to forego (छोड़ देना) the trend values for the first k and the last k terms of the series.

2. It cannot be used for forecasting or protecting 10 predicting future trend which is the main objective of trend analysis.

Que 6. A study of demand (d_t) for the past 12 years $(t = 1, 2, \dots, 12)$ has indicated the following:

$$d_i = \begin{cases} 100 & : \quad t = 1, 2, \dots, 5 \\ 20 & : \quad t = 6 \\ 100 & : \quad t = 7, 8, .., 12 \end{cases}$$

Compute a 5 year moving average.

Que 7. Calculate 5 – yearly and 7 – yearly moving averages of the data given below to obtain trend values and give their graphic representation:

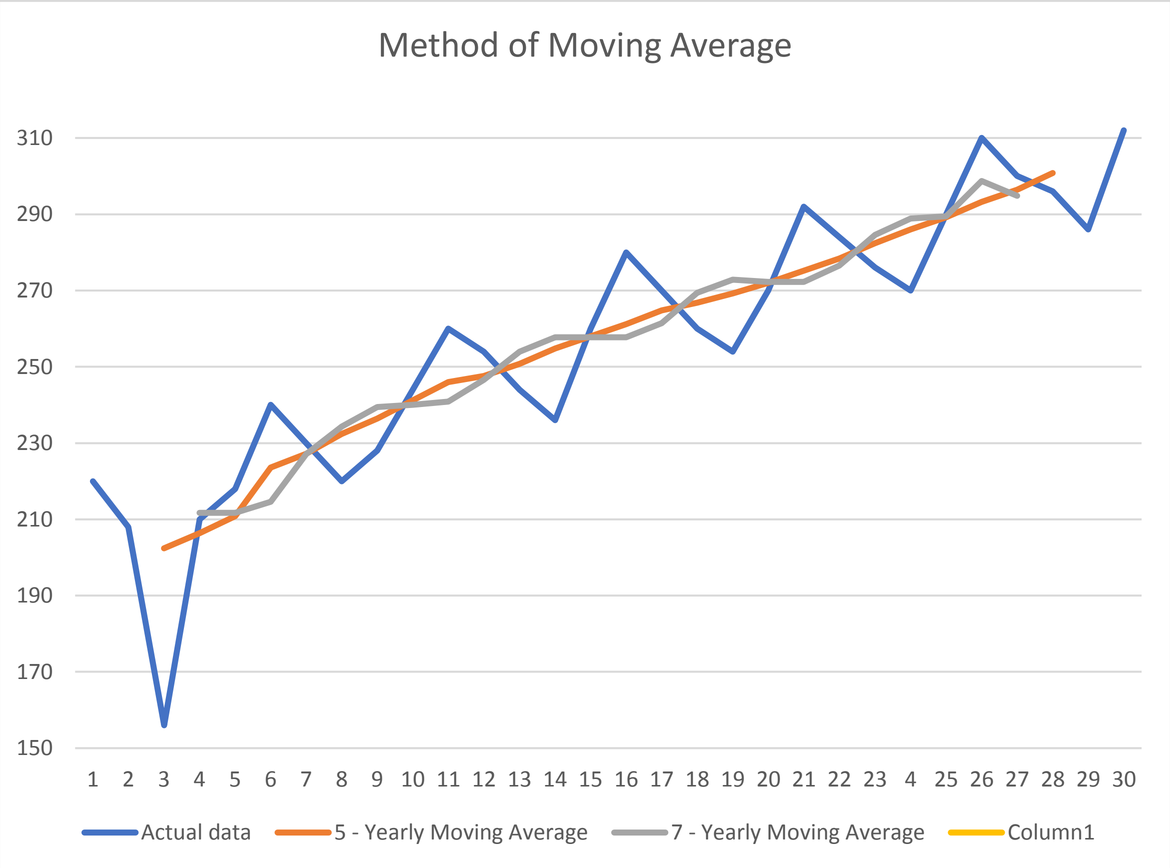
Year:	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Value:	**220**	**208**	**156**	**210**	**218**	**240**	**230**	**220**	**228**	**244**
Year:	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
Value:	**260**	**254**	**244**	**236**	**260**	**280**	**270**	**260**	**254**	**270**
Year:	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**
Value:	**292**	**284**	**276**	**270**	**290**	**310**	**300**	**296**	**286**	**312**

Solution:

Calculation of 5 and 7 yearly moving averages

Year	Value	5 – yearly Moving Total	5 – yearly Moving Averages	7 – yearly Moving Total	7 – yearly Moving Averages
1	220				
2	208				
3	156	1012	202.4		
4	210	1032	206.4	1482	211.71
5	218	1054	210.8	1482	211.71
6	240	1118	223.6	1502	214.57
7	230	1136	227.2	1590	227.14
8	220	1162	232.4	1640	234.28
9	228	1182	236.4	1676	239.42
10	244	1206	241.2	1680	240.00
11	260	1230	246.0	1686	240.85

12	254	1238	247.6	1726	246.57
13	244	1254	250.8	1778	254.00
14	236	1274	254.8	1804	257.71
15	260	1290	258.0	1804	257.71
16	280	1306	261.2	1804	257.71
17	270	1324	264.8	1830	261.42
18	260	1334	266.8	1886	269.42
19	254	1346	269.2	1910	272.85
20	270	1360	272.0	1906	272.28
21	292	1376	275.2	1906	272.28
22	284	1392	278.4	1936	276.57
23	276	1412	282.4	1992	284.57
24	270	1430	286.0	2022	288.85
25	290	1446	289.2	2026	289.42
26	310	1466	293.2	2028	289.71
27	300	1482	296.4	2064	294.85
28	296	1504	300.8		
29	286				
30	312				



MEASUREMENT OF SEASONAL VARIATIONS.

It has already pointed out that one of the types of fluctuations found in time series data is the seasonal component. Many economic and business series have distinct seasonal patterns that are pronounced enough to predict future behaviour of the series. The objectives for studying seasonal patterns in a time series are necessitated by the following reasons:

(i) To isolate the seasonal variations, i.e., to determine the effect of seasonal swing on the value of the given phenomenon, and

(ii) To eliminate them, i.e., to determine the value of the phenomenon if there were no seasonal ups and downs in the series. This is known as de – seasonalising the given data and is necessary for the study of cyclic variations.

The determination of seasonal effects is of paramount importance in planning (i) business efficiency, or (ii) a production programme. For example, the head of a departmental store would be interested to study the variations in the demands of different articles for different months in order to plan his future stocks to carter to the public demands due to seasonal swings. Moreover, the isolation and elimination of seasonal factor from the data is necessary to study the effect of cycles. Obviously, for the study of seasonal variations, the data must be given for 'parts' of the year, viz, monthly or quarterly, weekly, daily or hourly. Different methods for measuring seasonal variations are discussed below.

1.Method of Simple Averages. This is the simplest method of measuring seasonal ariations in a time series and involves the following steps:

1. Arrange the data by years and months (or quarters if quarterly data are given).
2. Compute the average $\bar{x}_i, (i = 1,2,..,12)$ for the ith month for all the years. [ith month, $i = 1,2,..,12$ represents January, February, …., December respectively.]
3. Compute the average $\bar{x}$ of the monthly averages, i.e., $\bar{x} = \frac{1}{12}\sum_{i=1}^{12}\bar{x}_i$.
4. Seasonal indices for different months are obtained by expressing monthly averages as percentage of $\bar{x}$. Thus.

$$Seasonal\ Index\ for\ ith\ month = \frac{\bar{x}_i}{\bar{x}} \times 100$$

Merits and Demerits. This method is very simple and easy to understand and calculate even for a non abstract person.

This method is based on the basic assumption that the data do not contain ant trend and cyclic components and consists in eliminating irregular components by averaging the monthly (or quarterly) values over different years. Since most of the economic time series have trends, these assumptions are not in general true and as such this method, through simple, is not of much practical utility.

Que 8. Use the method of monthly averages to determine the monthly indices for the following data of production of a commodity for the years 2002, 2003, 2004:

Month	Production (In lakh tonnes)			Month	Production (In lakh tonnes)		
	2002	2003	2004		2002	2003	2004
January	12	15	16	July	16	17	16
February	11	14	15	August	13	12	13
March	10	13	14	September	11	13	10
April	14	16	16	October	10	12	10
May	15	16	15	November	12	13	11
June	15	15	17	December	15	14	15

Solution: computation of seasonal indices

Month	Production (In lakh tonnes)			Total	Monthly Average	Seasonal Index
	2002	2003	2004			
January	12	15	16	43	14.33	104.90
February	11	14	15	40	13.33	97.58
March	10	13	14	37	12.33	90.26
April	14	16	16	46	15.33	112.22
May	15	16	15	46	15.33	112.22
June	15	15	17	47	15.66	114.64
July	16	17	16	49	16.33	119.54
August	13	12	13	38	12.66	92.67
September	11	13	10	34	11.33	82.94
October	10	12	10	32	10.66	78.03
November	12	13	11	36	13.00	95.16
December	15	14	15	44	14.66	107.32
Total				**492**		
Average				**13.66**		

2. Ratio to Moving Average Method. Moving average eliminate periodic movement in the extent (period of moving average) is equal to the period of the oscillatory movements sought to be eliminated. Thus, for a monthly data, a 12-month moving average should completely eliminate the seasonal movements if they are of constant pattern and intensity. The method of getting seasonal indices by moving average involved the following steps:

(i) Calculate the 12-month moving average of the data. These moving average values will give estimates of the combined effects of trend and cyclic variation.

(ii) Express the original data (except for 6 months in the beginning and 6 months at the end) as percentages of the centred moving average values. Using multiplicative model, these percentages would then represent the seasonal and irregular components.

(iii) The preliminary seasonal indices are now obtained by eliminating the irregular or random component by averaging these percentage.

(**iv)** The sum of these indices = S (say) will not, in general, be 1200 (for 400) for monthly (for quarterly) data. Finally, an adjustment is done to make the sum of the Indices 1200 (for 400) by multiplying through by a constant factor = 1200/S (or 400/S).

Merits and Demerits. Of all the methods of measuring seasonal variations, the ratio to the moving average method in the most satisfactory, flexible and widely used method for eliminating the seasonal fluctuations in a time series because it irons out both trend and cyclical components from the indices of seasonal variations.

However, an obvious drawback of this method is that there is loss of some trend values in the beginning and at the end and accordingly seasonal Indices for first six months (for 2 quarters) of the first year and last six month (or 2 quarters) of the last year cannot be determined.

Que 9. Apply ratio to moving average method to ascertain seasonal indices from the following data:

Year and Month 2002	No. of persons visiting a place of interest	Year and Month 2003	No. of persons visiting a place of interest	Year and Month 2004	No. of persons visiting a place of interest
January	**90**	**January**	**100**	**January**	**110**
February	**85**	**February**	**89**	**February**	**93**
March	**70**	**March**	**74**	**March**	**78**
April	**60**	**April**	**62**	**April**	**66**
May	**55**	**May**	**55**	**May**	**56**
June	**45**	**June**	**47**	**June**	**40**
July	**30**	**July**	**30**	**July**	**35**
August	**40**	**August**	**43**	**August**	**45**
September	**70**	**September**	**65**	**September**	**72**
October	**120**	**October**	**127**	**October**	**130**
November	**115**	**November**	**118**	**November**	**118**
December	**118**	**December**	**120**	**December**	**124**

Solution:

Year	Month	No. of persons visiting a place of interest	12 points moving totals	12 points moving average	12 point moving average (centred)	Ratio to moving average
2002	January	90				
	February	85				
	March	70				
	April	60				
	May	55				
	June	45				
			898	74.83		
	July	30			75.24	39.87
			908	75.66		
	August	40			75.83	52.74
			912	76.00		
	September	70			76.16	91.91
			916	76.33		

	October	120			76.41	157.04
			918	76.50		
	November	115			76.50	150.32
			918	76.50		
	December	118			76.58	154.08
			920	76.66		
2003	January	100			76.66	130.44
			920	76.66		
	February	89			76.78	115.91
			923	76.91		
	March	74			76.70	96.47
			918	76.50		
	April	62			76.79	80.73
			925	77.08		
	May	55			77.20	71.24
			928	77.33		
	June	47			77.41	60.71
			930	77.50		
	July	30			77.91	38.50
			940	78.33		
	August	43			78.49	54.78
			944	78.66		
	September	65			78.83	82.45
			948	79.00		
	October	127			79.16	160.43
			952	79.33		
	November	118			79.45	148.52
			955	79.58		
	December	120			79.29	151.34
			948	79.00		
2004	January	110			79.20	138.88
			953	79.41		
	February	93			79.49	116.99
			955	79.58		
	March	78			79.87	97.65
			962	80.16		
	April	66			80.28	82.21
			965	80.41		
	May	58			80.41	72.13

			965	80.41		
	June	40			80.58	49.64
			969	80.75		
	July	35				
	August	45				
	September	72				
	October	130				
	November	118				
	December	124				

Here the correction factor for obtaining adjusted seasonal indices

$$k = \frac{1200}{1197.47} = 1.002$$

Computation of adjusted seasonal indices

Month	**2002**	**2003**	**2004**	**seasonal indices (arithmetic average)**	**Adjusted seasonal indices (S.i.×C.F.)**
January		130.44	138.88	134.66	134.92
February		115.91	116.99	116.45	116.68
March		96.47	97.65	97.06	97.25
April		80.73	82.21	81.47	81.63
May		71.24	72.13	71.68	71.82
June		60.71	49.64	55.17	55.28
July	39.87	38.50		39.18	39.25
August	52.74	54.78		53.76	53.86
September	91.91	82.45		87.18	87.35
October	157.04	160.43		158.73	159.04
November	150.32	148.52		149.42	149.71
December	154.08	151.34		152.71	153.01
Total				**1197.47**	**1199.8**

Que 10. Calculate seasonal indices by the ratio to moving average method from the following data:

Quarter	**Year**			
	2014	**2015**	**2016**	**2017**
Q1	**75**	**86**	**90**	**100**
Q2	**60**	**65**	**72**	**78**

Q3	**54**	**63**	**66**	**72**
Q4	**59**	**80**	**85**	**93**

Answer: 122.36,   92.42,   84.69, 100.50

3.Ratio to Trend Method. This method is an improvement over the simple averages method and is based on the assumption that seasonal variation for any given month is constant factor of the trend. The measurement of seasonal variation by this method consists of the following steps:

(i) Compute the trend values by the principle of least squares by fitting an appropriate mathematical curve (straight line, second degree parabolic curve for exponential curve, etc.).

(ii) Express the original data as the percentage of the trend values. Assuming the multiplicative model, these percentage will, therefore contain the seasonal, cyclic and irregular components.

(iii) The cyclic and irregular components are then wiped out by averaging the percentage for different months (quarters) If the data are monthly (quarterly), thus leaving us with indices of seasonal variations. Either Arithmetic mean or median can be used for averaging, but median is preferred to Arithmetic mean.

(iv) Finally, these indices, obtained in step 3, are adjusted to a total of 1200 for monthly data of 400 for quarterly data by multiplying them through by a constant k given by

$$k = \frac{1200}{total\ of\ the\ indices} \qquad or \qquad k = \frac{400}{total\ of\ the\ indices}$$

for monthly and quarterly data respectively.

Merits and Demerits. Since this method attempts at ignoring out the cyclical or aur irregular components by the process of averaging, the purpose will be accomplished only if the cyclical variations are known to be absent or they are not so pronounced (निश्रित) even if present.

The obvious advantage of this method over the moving average method lies in the fact that ratio to trend can be obtained for each month for which the data are available and as such, Unlike the ratio to moving average method there is no loss of data.

Que 11. Using ratio to trend method, determine the quarterly seasonal Indices for the adjoining data

Year	I Qrt.	II Qrt.	III Qrt.	IV Qrt.
1995	**30**	**40**	**36**	**34**
1996	**34**	**52**	**50**	**44**
1997	**40**	**58**	**54**	**48**
1998	**54**	**76**	**68**	**62**
1999	**80**	**92**	**86**	**82**

Que 12. Using the ratio to trend method comedy term in the quarterly seasonal indices:

Year	Quarter I	Quarter II	Quarter III	Quarter IV
1	**65**	**60**	**61**	**63**
2	**70**	**58**	**56**	**60**
3	**68**	**63**	**68**	**67**
4	**65**	**59**	**56**	**62**
5	**60**	**55**	**51**	**58**

Solution: first of all, we determine the trend values for the quarterly averages by fitting a linear trend by the method of least squares. Let we shift origin to 3.

$$x = t - 3$$

The normal equation for fitting of trend line $y = a + b\,x$ are given by:

$$\sum y = na + b \sum x$$
$$\sum x\ y = a \sum t + b \sum x^2$$

Year	Total of quarterly values	Average of quarterly values (y)	$x = t - 3$	x^2	xy	Trend values
1	249	62.25	-2	4	-124.5	63.85
2	244	61	-1	1	-61	62.55
3	266	66.5	0	0	0	61.25
4	242	60.5	1	1	60.5	59.95
5	224	56	2	4	112	58.65
Total		**306.25**	**0**	**10**	**-13**	

Normal equations become

$$306.25 = 5a + b\,.0 \quad \Rightarrow \quad 306.25 = 5a \quad \Rightarrow \quad a = 61.25$$
$$-13 = a\,.0 + b\,.10 \quad \Rightarrow \quad -13 = 10\,b \quad \Rightarrow \quad b = -1.3$$

The trend line given by

$$y = 61.25 - 1.3\,x$$

The negative value of b implies that the data has a decreasing trend. Now we determine the quarterly trend values

Yearly decrease in trend values = 1.3

Quarterly decrease in trend values= $\frac{1.3}{4} = 0.325$

Year	Trend values				Trend eliminated values (given values as % of trend values)			
	Qtr. I	Qtr. II	Qtr. III	Qtr. IV	Qtr. I	Qtr. II	Qtr. III	Qtr. IV
1	64.33	64.01	63.68	63.36	101.04	93.73	95.79	99.43
2	63.03	62.71	62.38	62.06	111.05	92.48	89.77	96.68
3	61.72	61.41	61.08	60.76	110.17	102.58	111.32	110.26
4	60.43	61.11	59.78	59.46	107.56	96.54	93.67	104.27
5	59.13	58.81	58.48	58.16	101.47	93.52	87.20	99.72
Total					**531.29**	**478.85**	**477.75**	**510.36**
	Average seasonal indices				**106.25**	**95.77**	**95.55**	**102.07**
	Adjusted seasonal indices				**106.34**	**95.85**	**95.63**	**102.16**

Now correction factor

$$k = \frac{400}{sum\ of\ indices} = \frac{400}{106.25 + 95.77 + 95.55 + 102.07} = \frac{400}{399.64}$$
$$= 1.0009$$