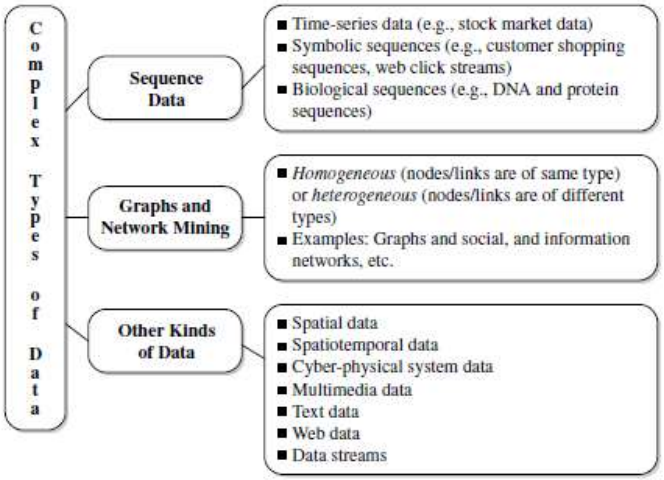
Topic :

1. Forecasting Time Series
2. Introduction to time series

Description:



Complex data types for mining.

In time-series data, sequence data consist of long sequences of numeric data, recorded at equal time intervals (e.g., per minute, per hour, or per day). Time-series data can be generated by many natural and economic processes such as stock markets, and scientific, medical, or natural observations.

Regression and Trend Analysis in Time-Series Data

Regression analysis of time-series data has been studied substantially in the fields of statistics and signal analysis. However, one may often need to go beyond pure regression analysis and perform trend analysis for many practical applications. Trend analysis builds an integrated model using the following four major components or movements to characterize time-series data:

1. Trend or long-term movements: These indicate the general direction in which a time-series graph is moving over time, for example, using *weighted moving average* and the *least squares* methods to find *trend curves.*

2. Cyclic movements: These are the long-term oscillations about a trend line or curve.

3. Seasonal variations: These are nearly identical patterns that a time series appears to follow during corresponding seasons of successive years such as holiday shopping seasons. For effective trend analysis, the data often need to be "deseasonalized" based on a **seasonal index** computed by autocorrelation.

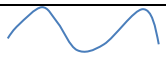
4. Random movements: These characterize sporadic changes due to chance events such as labour disputes or announced personnel changes within companies.

Trend analysis can also be used for time-series forecasting, that is, finding a mathematical function that will approximately generate the historic patterns in a time series, and using it to make long-term or short-term predictions of future values. ARIMA (auto-regressive integrated moving average), long-memory time-series modeling, and autoregression are popular methods for such analysis. Time series forecasting is performed in nearly every organization that works with quantifiable data.

- Retail stores use it to forecast sales.
- Energy companies use it to forecast reserves, production, demand, and prices.
- Educational institutions use it to forecast enrolment.
- Governments use it to forecast tax receipts and spending.
- International financial organizations such as the World Bank and International Monetary Fund use it to forecast inflation and economic activity.
- Transportation companies use time series forecasting to forecast future travel.
- Banks and lending institutions use it (sometimes badly!) to forecast new home purchases and
- Venture capital firms use it to forecast market potential and to evaluate business plans.

A time series is a sequence of observations over a certain period. The simplest example of a time series that all of us come across on a day to day basis is the change in temperature throughout the day or week or month or year. The analysis of temporal data is capable of giving us useful insights on how a variable changes over time.

Components of Time series:

Type	Representation
Trend T	
Seasonality S	
Cycle C	
Randomness R	

Analysis of Time Series:
Moving Average Method
A moving average is defined as an average of fixed number of items in the time series which move through the series by dropping the top items of the previous averaged group and adding the next in each successive average.

Let $(t_1, y_1), (t_2, y_2), \ldots, (t_n, y_n)$ denote given time series $y_1, y_2, \ldots, y_n$ are the values of the variable y; corresponding to time periods $t_1, t_2, \ldots, t_n$, respectively.

The moving averages of order m are defined as

$$\frac{y1+y2+\cdots+ym}{m}; \qquad \frac{y2+y3+\cdots+ym+1}{m+1}$$

Here $y_1+y_2+\ldots+y_m$, $y_2+y_3+\ldots+y_{m+1}$,… are called moving totals of m.
In using moving averages in estimating the trend, we shall have to decide as what should be the order of the moving averages. The order of the moving average should be equal to the length of the cycles in the time series. In case the order of the moving averages is given in the problem itself, then we shall use that order for computing the moving average. The order of the moving averages may

either be odd or even.

The moving averages of order 3 are

$\frac{y1+y2+y3}{3}; \frac{y2+y3+y4}{3}; \frac{yn-2+yn-1+yn}{3}$

These moving averages are called the trend values. They are considered to correspond 2nd, 3rd, …, (n−1)th years, respectively.

Merits

1. The moving average method eliminates the short-term fluctuations.
2. It reduces the effect of extreme values.
3. As the free-hand method, this method is not subject to personal prejudice and bias of the estimator.
4. This method is a flexible method.

Demerits

1. Moving average method is not fully mathematical.
2. If the series given is a very large one, then the calculation of moving average is cumbersome.
3. The choice of the period of moving average needs a great amount of care. If an inappropriate period is selected, a true picture of the trend cannot be obtained.
4. It is very much affected by extreme values.

Example: Estimate the trend values using the data given below by taking a 3-yearly moving averages.

Year	1980	1981	1982	1983	1984	1985	1986	1987	1988
Sales (in lakhs of units)	65	95	80	115	105	135	125	150	140

Solution: Trend by 3-yearly moving averages

Year	**Sales (in lakhs of units)**	**3-year moving totals**	**3-year moving averages**
1980	65	–	–
1981	95	65+95+80=240	240/3=80
1982	80	95+80+115=290	290/3=96.67
1983	115	80+115+105=300	300/3=100
1984	105	115+105+135=355	355/3=118
1985	135	105+135+125=365	365/3=121.67
1986	125	135+125+150=410	410/3=136.67
1987	150	125+150+140=415	415/3=138.33
1988	140	–	–

Estimate the trend values using the data given below by taking a 4-yearly moving averages:

Year	Value	4-yearly moving totals	4-yearly moving averages	4-yearly centered moving averages
1969	4			
1970	7	46	46/4=11.5	14.75
1971	20	72	18.0	20.625
1972	15	93	23.25	23.75
1973	30	97	24.25	26.875
1974	28	118	29.5	28.875
1975	24	113	28.25	29.75
1976	36	125	31.25	33.50
1977	25	143	35.75	36.875
1978	40	152	38	
1979	42			
1980	45			

Courtesy : https://www.sciencedirect.com/topics/engineering/moving-average#:~:text=A%20moving%20average%20is%20defined,next%20in%20each%20successive%20average.