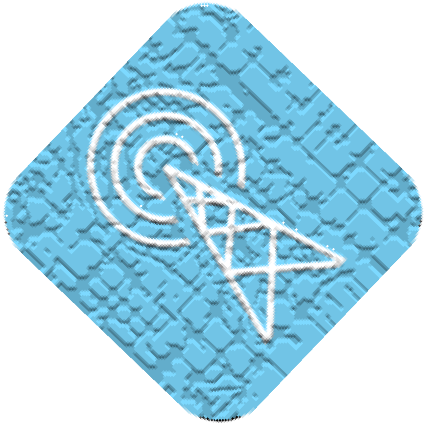
Data Science Concepts

Lesson08–Time Series Concepts

Objective

After completing this lesson you will be able to:



- Describe the application areas of Time Series
- Explain Trend, Seasonality, Cyclic and Noise component in data
- Explain the various models in time series.

Forecasting demand of products and services for long range and short range planning.

Amazon sells more than 13 millions SKUS and has more than 2 million retailers selling them.

Demand of manpower in IT product and services to manage the business.

Netflix opened up a challenge to forecast movie ratings by customer. Can be further used for movie recommendation.

Time Series Data

Data on Response variable (Y) which is a random variable

Data points are collected at regular interval and arranged in chronological order

Univariate time series: data contains observation of a single variable

- warranty claims at time t

Multivariate time series: data contains observation of more than one variable

- warranty claims at time t
- sale of motor vehicles at time t
- Climatic conditions at time t

Time series components

- Trend Component (T_t)
  - Consistent upward or downward movement of data over a period of time
  - Trend can be identified by seasonality window or seasonality index

- Seasonality Component (S_t)
  - Seasonality, measured in terms of seasonality index, is fluctuations from the trend that occurs within a defined time period (seasons, quarters, months, days of the week, time interval within a day etc.)

- Cyclic Component (C_t)
  - Fluctuations around the trend line which happens due to macro economic changes (recession, unemployment etc.)
  - Fluctuations have repetitive pattern and time between repetition is more than a year but this time is random

- Irregular component (I_t)
  - White noise or random uncorrelated changes that follow a normal distribution with a mean of zero and constant variance

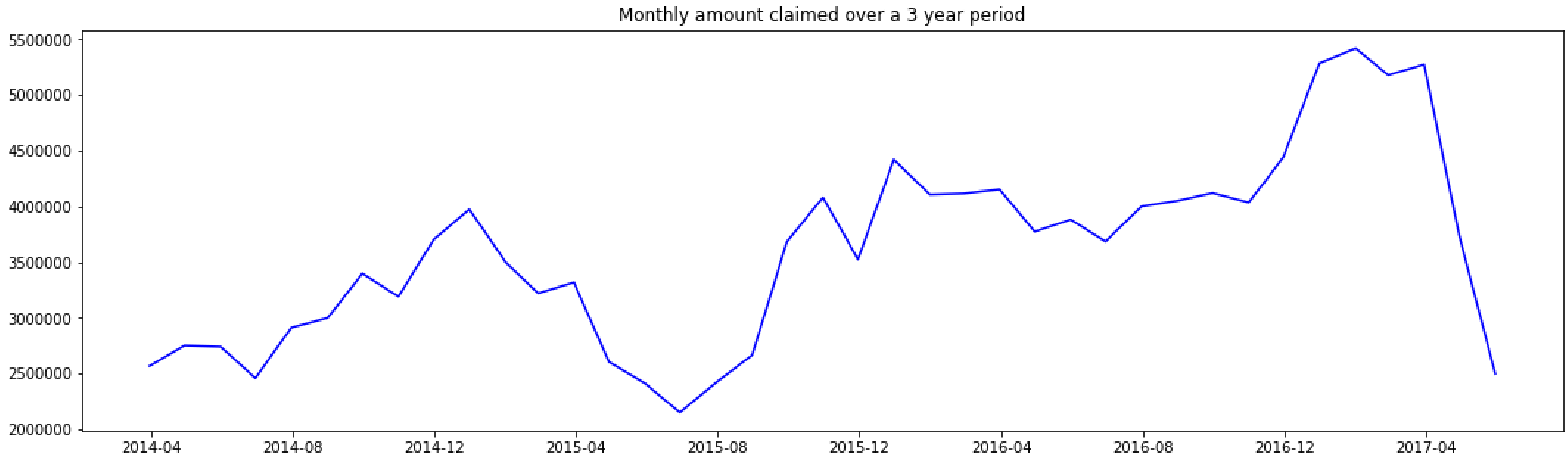
BSE SENSEX
INDEXBOM: SENSEX

39,982.97 −100.57 (0.25%)

6 Jun, 10:03 am IST · Disclaimer

1 day 5 days 1 month 6 months YTD 1 year **5 years** Max

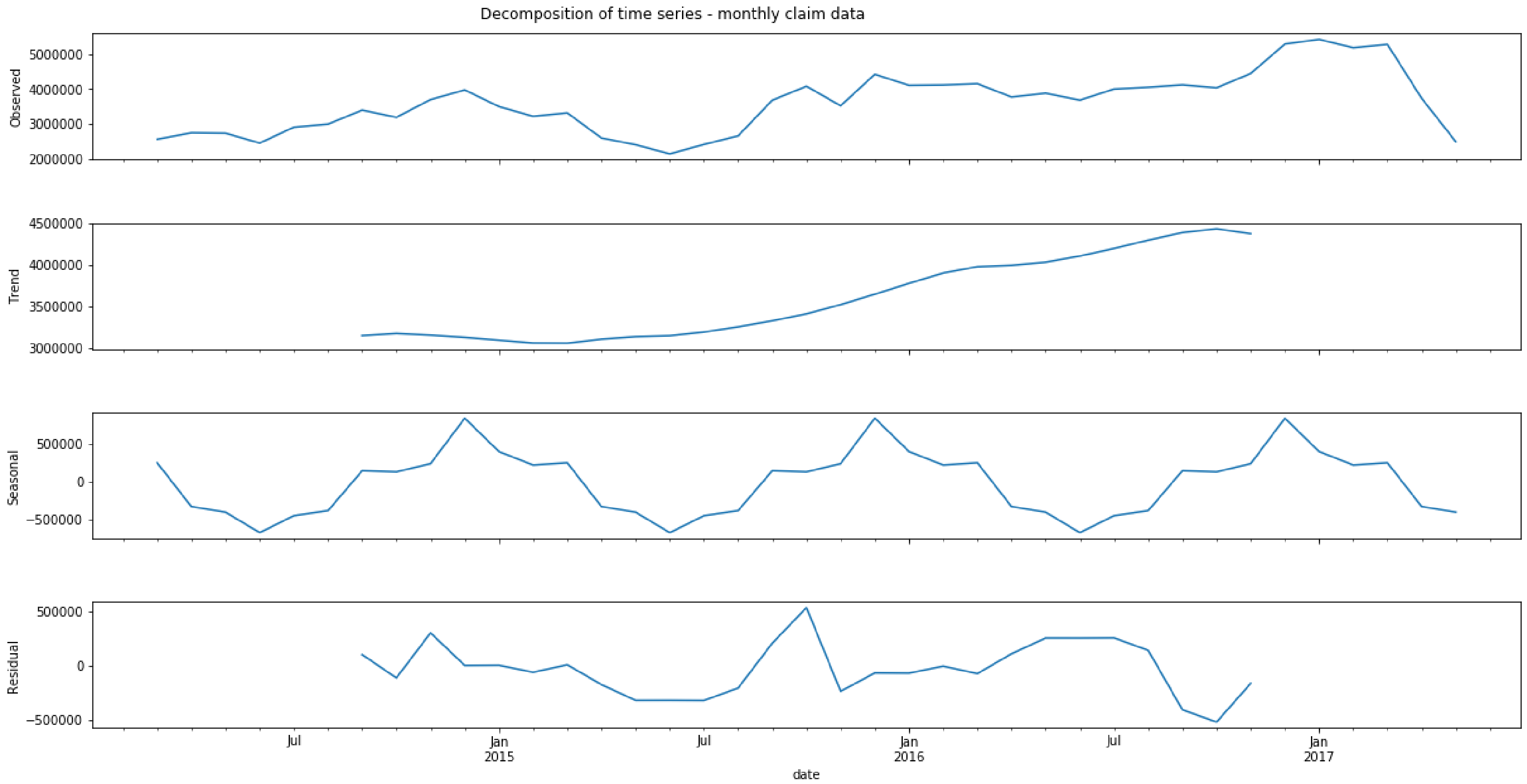
Open	40,136.43	Low	39,966.05
High	40,159.26		



Seasonality window

The idea of finding seasonality window is to find a window size of `s`, for which if the rolling average is calculated for each time point (-s/2<t<s/2), the zigzag motion in the time series data smoothens out.

The rolling average in the window `s` tries to smooth out noise and seasonality and what is observed is a pure trend.



The seasonality window in the claim data was found to be 13

Additive and Multiplicative model

Time series data can be modelled as addition or multiplication of the time series component.

Additive Model	Multiplicative Model
$Y_t = T_t + S_t + C_t + I_t$	$Y_t = T_t * S_t * C_t * I_t$
• Assumes that, seasonal and cyclic component are independent of trend • Useful when seasonal component remain constant around the mean or does not vary with the mean of the series.	• Better fit to most cases • Useful when seasonal component is correlated with the mean of the series

Forecasting Accuracy

- Mean Absolute Error

$$MAE = \sum_{t=1}^{n} \frac{|Y_t - F_t|}{n}$$

- Mean Absolute Percentage Error

$$MAPE = \left(\frac{1}{n} * \sum_{t=1}^{n} \frac{|Y_t - F_t|}{Y_t} \right) * 100$$

- Mean Squared Error

$$MSE = \frac{1}{n} * \sum_{t=1}^{n} (Y_t - F_t)^2$$

- Mean Squared Error

$$RMSE = \sqrt{\frac{1}{n} * \sum_{t=1}^{n} (Y_t - F_t)^2}$$

MAPE is dimensionless and thus can be used to compare different models as well. RMSE is standard deviation of error. Lower RMSE implies better prediction.

Forecasting Techniques

Many forecasting techniques developed based on different logics.

- Moving average, exponential smoothening: Predict future value as a function of past observation.

- Regression based model (AR, MA, ARMA, ARIMA, ARIMAX): Predict future value as a function of past observation. However, are more advanced in terms of handling trend and seasonality effect.

Using complicated method does not guarantee better accuracy in forecasting.

Moving average

Forecast the future value using average or weighted average of past N observations.

$$F_{t+1} = \frac{1}{N} * \sum_{k=t+1-N}^{t} Y_k$$

In case of weighted moving average:

$$F_{t+1} = \sum_{k=t+1-N}^{t} W_k * Y_k$$

W_k is the weight given to past observation. Weight decreases as data becomes older.

$$\sum_{k=t+1-N}^{t} W_k = 1$$

Equation 1 is simple moving average as N past observations have been given equal weight. Assumes that the data does not have significant trend or seasonal component.

Single exponential smoothening

Forecast the future value using differential weights to past observations.

$$F_{t+1} = \alpha * Y_t + (1 - \alpha) * F_t$$

α is called the smoothening constant with a value between 0 to 1.

$$F_{t+1} = \alpha * Y_t + \alpha * (1 - \alpha) * Y_{t-1} + \alpha * (1 - \alpha)^2 * Y_{t-2} \ldots + \alpha * (1 - \alpha)^{t-1} * Y_1 + (1 - \alpha)^2 * F_1$$

Initial value of F_t is taken same as Y_t

Compared to SMA, SES uses the entire history data

Weights assigned declines exponentially with the most recent observation having the highest weight. Model does not perform better in presence of trend.

Double exponential smoothening – Holts Method

Forecast the future value using two equations. One for forecasting level (short term average) and the other for capturing trend

$$L_t = \alpha * Y_t + (1 - \alpha) * F_t$$
$$T_t = \beta * (L_t - L_{t-1}) + (1 - \beta) * T_{t-1}$$

α, β are called the level and trend smoothening constant with a value between 0 to 1.

$$F_{t+1} = L_t + T_t = \{\alpha * Y_t + (1 - \alpha) * F_t\} + \{\beta * (L_t - L_{t-1}) + (1 - \beta) * T_{t-1}\}$$

$$F_{t+n} = L_t + nT_t$$

- L_t is the level representing the smoothed value up to the last data point
- T_t is the slope of the line at period t. n is the number of time periods to forecast in future.
- Initial value of L_t is taken same as Y_t

Weights assigned declines exponentially with the most recent observation having the highest weight. Model does not perform better in presence of seasonal component.

Triple exponential smoothening – Holts Winter Method

Forecast the future value using three equations. One for forecasting level second for capturing trend and third for seasonality

$$L_t = \alpha * \frac{Y_t}{S_{t-c}} + (1-\alpha) * [L_{t-1} + T_{t-1}]$$

$$T_t = \beta * (L_t - L_{t-1}) + (1-\beta) * T_{t-1}$$

$$S_t = \gamma * (\frac{Y_t}{L_t}) + (1-\gamma) * S_{t-c}$$

- α, β, γ are called the level, trend and seasonality smoothening constant with a value between 0 to 1.
- c is the number of seasons. (monthly data c = 12, quarterly data c = 4, daily data c = 7 etc.)

$$F_{t+1} = [L_t + T_t] * S_{t+1-c}$$

Holts Winter Method – Initial value calculation

- Level

$$L_t = Y_t$$

or

$$L_t = \frac{1}{c} * (Y_1 + Y_2 + Y_3 \dots + Y_c)$$

- Trend

$$T_t = \frac{1}{c} * (\frac{Y_t - Y_{t-c}}{12} + \frac{Y_{t-1} - Y_{t-c-1}}{12} + \cdots + \frac{Y_{t-c+1} - Y_{t-2c+1}}{12})$$

- Season: Initial seasonality can be calculated using method of simple averages. Seasonality index can be interpreted as percentage change from the trend line.
    - Apart from simple averages, several other techniques also exists

This technique is very sensitive to initial values of level, trend and seasonal index

Regression based Models

When the data consists of $Y_t \ and \ X_t$, regression based models may be better suited than the previous methods. Y_t may have seasonal variation.

$$F_{d,t+1} = \beta_0 + \beta_1 * x_{1,t} + \beta_2 * x_{2,t} + \cdots + \beta_n * x_{n,t} + \varepsilon$$

- Where $F_{d,t+1}$ represents forecasted demand on de-seasonalized data.

Steps to de-seasonalize data:

1. Estimate seasonality index (S_t) using moving average technique.
2. De-seasonalize data using multiplicative model: $Y_{d,t} = Y_t / S_t$
3. Develop forecasting model on de-seasonalized data ($F_{d,t+1}$)
4. The actual forecast for t+1 is $F_{t+1} = F_{d,t+1} * S_t$

End of Lesson08–Time Series Concepts