Time Series Forecasting

What is Time Series

Time series forecasting is a method used to predict future values based on previously observed values in a time series data. A time series is a sequence of data points collected or recorded at regular time intervals. This kind of data is prevalent in a wide range of fields, including economics, finance, science, and engineering

Characteristics

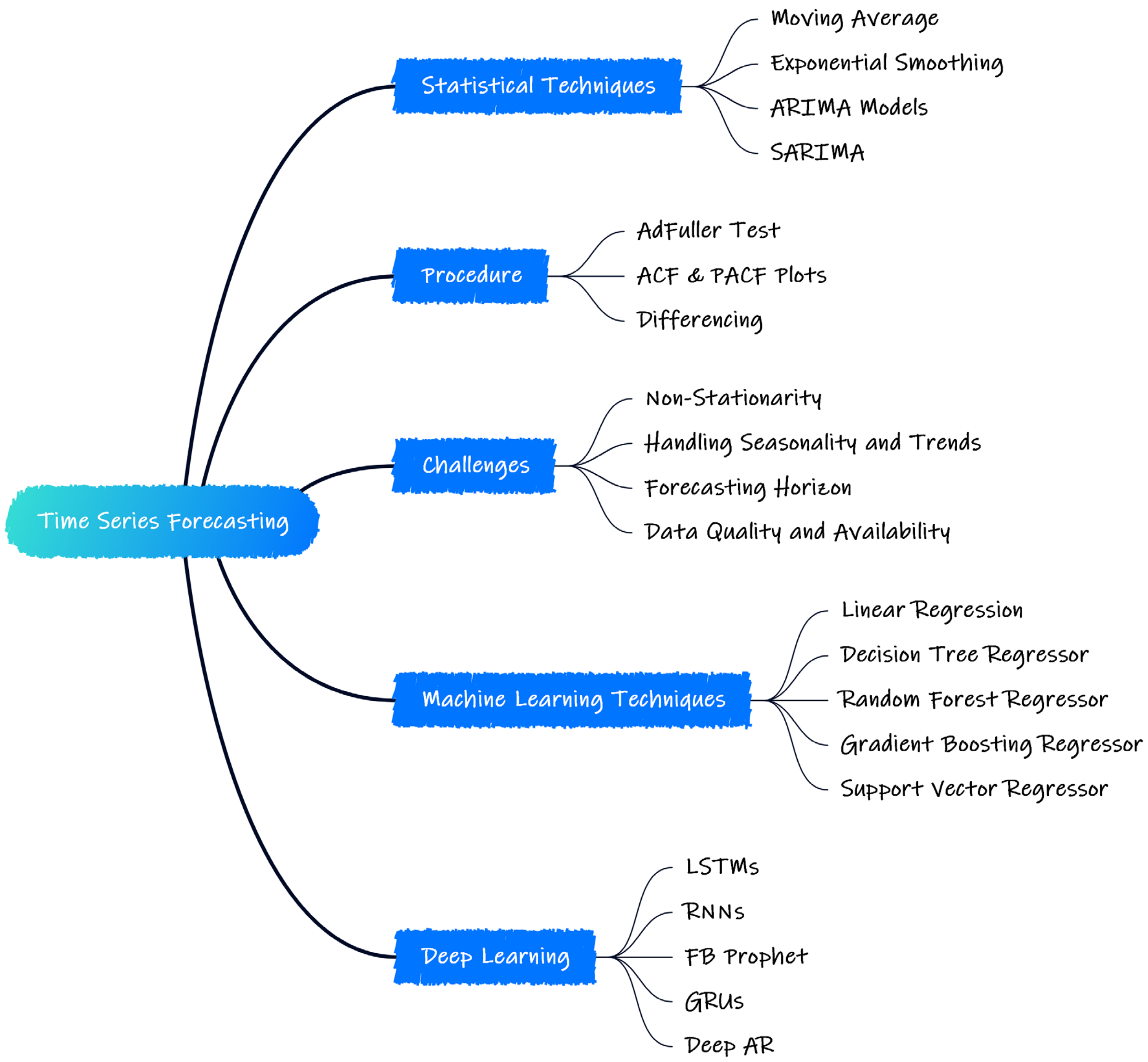
Data points in a time series have a natural temporal ordering. This makes time series data different from cross-sectional data, where observations do not have a time order

Many time series exhibit seasonal variations—a pattern that repeats at regular intervals over time. For example, retail sales often peak during the holiday season.

Time series often have a long-term progression—either up or down. This trend component reflects the overall direction of the series over time

Real-world time series data often contains random variation, or 'noise', which can obscure the underlying signal in the data

Forecasting Techniques



Model Comparison

Model	When to Use	Why to Use	Advantage	Disadvantage
Moving Average (MA)	Simple trend-less data	Easy to implement and interpret. Good for short-term forecasts when data is relatively stable	• Simple to understand and use • Effective in removing noise	Not suitable for complex trends and seasonality
Weighted MA	When recent data is more relevant	To give more importance to recent observations in the forecast	• More weight to recent data • Can be customized based on specific requirements	More complex than MA. Subjective choice of weights
Exponential Smoothing	Data with trends and seasonality	Good for smoothing data and responding to level changes quickly	• Adapts quickly to changes in level • Can handle trend and seasonality	• Not ideal for long-term forecasting • Requires parameter tuning

Model Comparison

Model	When to Use	Why to Use	Advantage	Disadvantage
ARIMA	Non-seasonal or seasonal stationary data	Flexible and widely applicable, especially for stationary time series	• Can model various types of time series data • Good for short-term forecasts	• Complex to understand and use • Not suitable for non-stationary data
SARIMA	Seasonal data with trends	Extends ARIMA to account for seasonality	• Handles both seasonality and non-seasonality • Flexible and can model various types of seasonal time series data	• More complex than ARIMA • Requires careful selection of parameters
Regression	Data with influential external variables	When the forecast depends on external factors or when using multivariate time series	• Can incorporate external variables • Flexible in terms of model choice (linear, polynomial, etc.)	• Assumes a linear relationship (unless using non-linear models) • Requires understanding of underlying factors affecting the series

Model Comparison

Model	When to Use	Why to Use	Advantage	Disadvantage
LSTMs	Complex patterns, long sequences	Good at capturing long-term dependencies in data	• Excellent for learning sequences and patterns • Can handle large datasets	• Requires large amount of data • Computationally intensive
FB Prophet	Seasonal data, daily data, and business forecasting	User-friendly and provides robust forecasts even with missing data or outliers	• Handles outliers well • Intuitive and easy to use	• Less customizable than some models • May not perform well on very irregular time series
Deep AR	Large, multivariate time series	Ideal for probabilistic forecasting and complex temporal dynamics	• Good for large datasets • Can capture complex relationships and patterns in the data	• Requires significant computational resources • More complex to set up and train

ADFuller Test

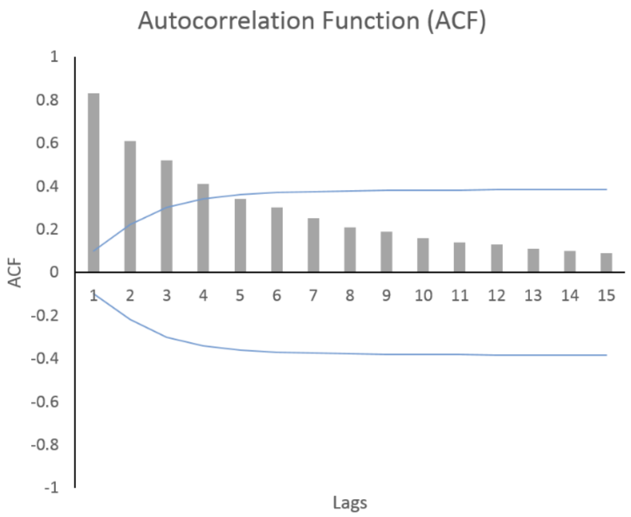
- The Augmented Dickey-Fuller (ADFuller) test is a statistical test used to check for the presence of a unit root in a time series. A unit root indicates non-stationarity, meaning that the mean, variance, and autocorrelation structure of the series can change over time
- Most time series forecasting methods, including ARIMA, assume that the data is stationary. If the test indicates non-stationarity, transformations such as differencing, logging, or deflation may be necessary to stabilize the time series before modeling

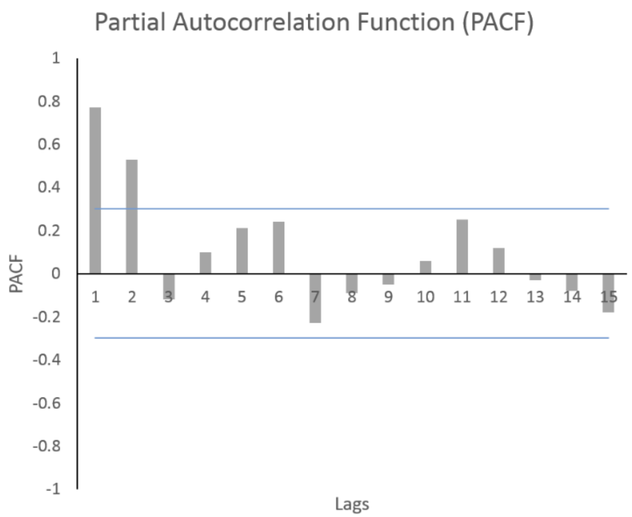
ACF & PACF Plot

- **Auto Correlation Function:** This plot shows the correlation of the series with itself, lagged by x time units. It's useful for identifying the presence of autocorrelation (when current values are correlated with past values).
- Helps in determining the order of the MA (Moving Average) component in ARIMA models
- **Partial Auto Correlation Function:** This plot shows the partial correlation of the series with its own lagged values, controlling for the values of the time series at all shorter lags. It isolates the impact of each lag
- Assists in determining the order of the AR (AutoRegressive) component in ARIMA models

AR Process

- A gradual geometrically declining ACF and a PACF that is significant for only a few lags indicate an AR Process.
- With 2 Significant PACF lags and gradually falling ACF, we can say that the series is an AR(2) Process.
- The lags of AR are determined by the number of significant lags of PACF





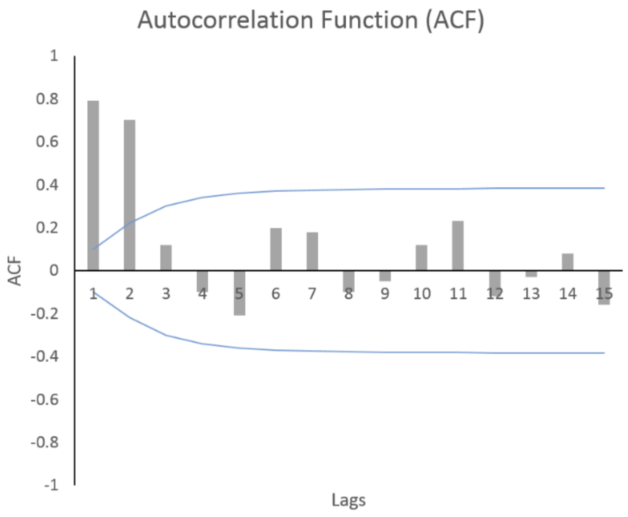
AR Process

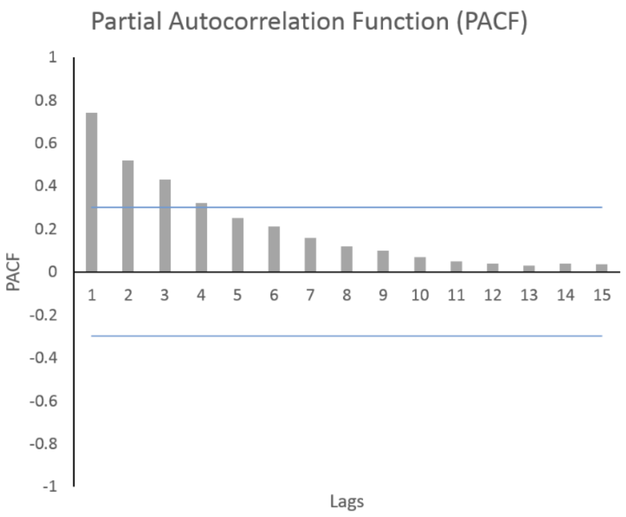
- The AR part of an ARIMA model is typically represented by an equation that captures the dependence of the current value of the series on its previous values. The general form of an AR(p) model, where p is the order of the autoregression, is:
- $Yt=c+\phi 1Yt-1+\phi 2Yt-2+\ldots+\phi pYt-p+\epsilon t$

- Where:
- Yt is the current value of the series.
- c is a constant (intercept).
- $\phi 1,\phi 2,\ldots,\phi p$ are the parameters of the model, representing the impact of each lagged value on the current value.
- $Yt-1,Yt-2,\ldots,Yt-p$ are the lagged values of the series.
- ϵt is white noise (error term)

MA Process

- A gradual geometrically declining PACF and the ACF that is significant for only a few lags indicate an MA Process.
- Here, the PACF is falling geometrically and the ACF has 2 significant lags before dropping.
- This Indicates MA(2) Process, MA lags are determined by the number of significant ACF Lags



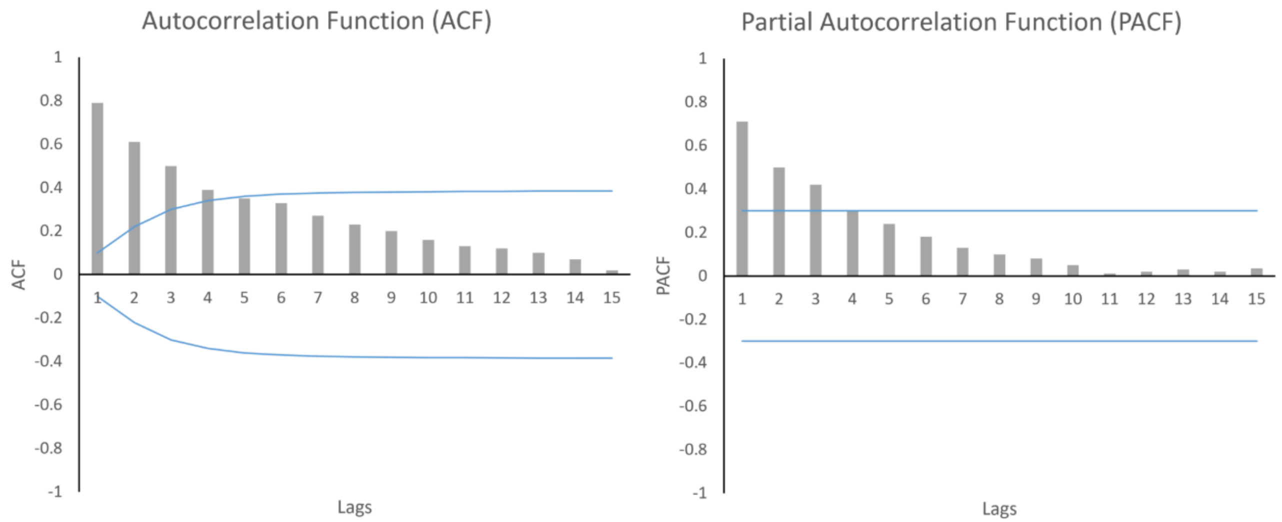


MA Process

- The general form of the MA(q) model, where q is the order of the moving average, is expressed as follows:
- $Y_t = \mu + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \ldots + \theta_q \epsilon_{t-q}$
- Where:
- Y_t is the current value of the series.
- μ is the mean of the series.
- ϵ_t is the white noise (error term) at time t.
- $\theta_1, \theta_2, \ldots, \theta_q$ are the parameters of the model, representing the impact of the error terms at lags 1 through q.
- $\epsilon_{t-1}, \epsilon_{t-2}, \ldots, \epsilon_{t-q}$ are the lagged forecast errors
- The complete ARIMA model is denoted as ARIMA(p, d, q), where p is the number of autoregressive terms, d is the degree of differencing, and q is the number of moving average terms

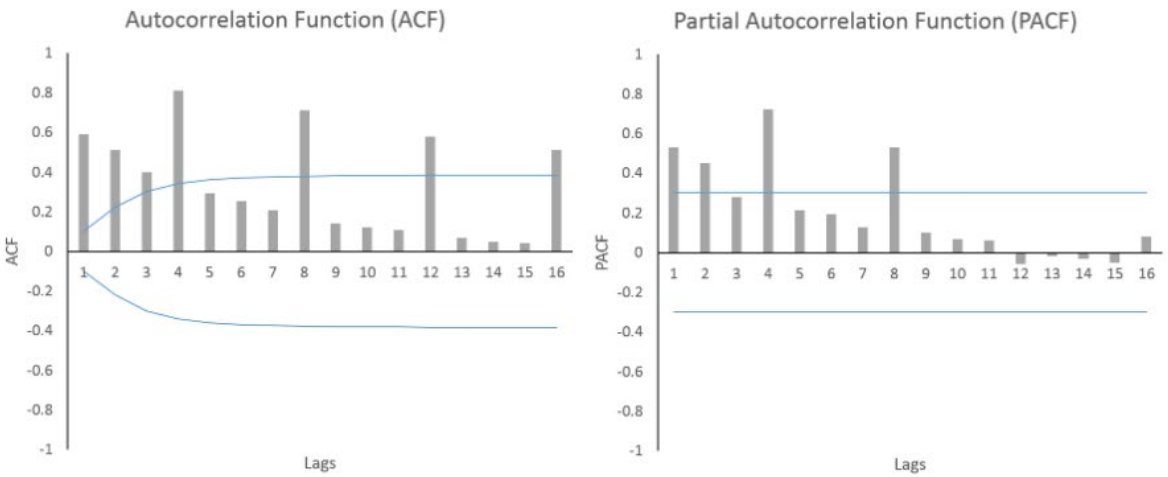
ARMA Process

- An ARMA Process is indicated by geometrically falling ACF and PACF
- The number of AR and MA terms to include in the model can be decided with the help of Information Criteria such as AIC



Seasonal AR

- The ACF is gradually declining with every 4th period and the PACF shows 2 significant seasonal lags.
- This suggest that the series is a seasonal-AR(2) process because the PACF has 2 significant lags.



1^{st} Order Differencing

- This is a transformation that computes the differences between consecutive observations. In mathematical terms, it transforms a series Y_t into $Y_t - Y_{t-1}$.
- **Why It's Important**: It's a common method to remove trend and stabilize the mean of the time series, helping to make it stationary. Stationarity is a key assumption in many time series forecasting models.

2^{nd} Order Differencing

- This involves applying the differencing operation twice, i.e., taking the difference of the difference. Mathematically, it's $(Yt-Yt-1)-(Yt-1-Yt-2)$.
- **Why It's Important**: Sometimes, 1st order differencing isn't enough to achieve stationarity, especially in the presence of a more complex trend or seasonal patterns. 2nd order differencing can help to further stabilize the mean and remove trends.

Lagging

- Lagging in time series forecasting involves shifting the time series data backwards by a certain number of periods. This means creating lagged variables, where each new variable represents the value of the original time series at a previous point in time
- The primary purpose of creating lagged variables is to capture the temporal dependencies within the data. In many time series models, the value at a particular time point is assumed to be related to its previous values (lags).

Deflation

- Deflation, in the context of economic time series data, refers to the process of adjusting the nominal values (values not adjusted for inflation) by a price index (like the Consumer Price Index, CPI) to obtain real values (values adjusted for inflation).
- The aim of deflating a time series is to remove the effects of inflation and to measure the real value or volume of variables over time. This allows for a more accurate analysis of trends and patterns, as it reflects the true underlying changes in the data, excluding price level changes
- To deflate a time series of yearly revenues, you would divide the revenue of each year by the CPI (or another relevant price index) of that year, and then multiply by the CPI of the base year. This adjusts the revenue figures to a constant price level, making year-to-year comparisons meaningful in real terms.

Exponential Smoothing

- Exponential Smoothing is a time series forecasting method that applies exponentially decreasing weights to past observations. This technique is particularly effective for making short-term forecasts. The most basic form of exponential smoothing, known as Simple Exponential Smoothing (SES), is suitable for time series without trend and seasonality
- **Simple Exponential Smoothing (SES)**
- $Y^\wedge t+1=\alpha Yt+(1-\alpha)Y^\wedge t$
- Where:
- $Y^\wedge t+1$ is the forecast for the next period.
- Yt is the actual observation at time t.
- $Y^\wedge t$ is the forecast for period t.
- α is the smoothing factor (a value between 0 and 1).

Exponential Smoothing

1. **Holt’s Linear Trend Method**: This method extends SES by explicitly incorporating the trend component into the model. It involves two equations: one for the level (similar to SES) and one for the trend.
2. **Holt-Winters Seasonal Method**: This method further extends exponential smoothing to handle seasonality. It involves separate smoothing equations for the level, trend, and seasonal components.