CONCEPTUAL QUESTIONS

For

TIME SERIES FORECASTING

Website: www.analytixlabs.co.in
Email: info@analytixlabs.co.in

Q. What Is the Time series?

Time Series: A time series is a data series consisting of several values over a time interval, e.g. daily BSE Sensex closing point, weekly sales and monthly profit of a company etc.

Typically, in a time series it is assumed that value at any given point of time is a result of its historical values. This assumption is the basis of performing a time series analysis.

Example: Suppose Mr. X starts his job in year 2010 and his starting salary was $5,000 per month. Every year he is appraised and salary reached to a level of $20,000 per month in year 2014. His annual salary can be considered a time series and it is clear that every year's salary is function of previous year's salary (here function is appraisal rating).

Time Series Modeling involves working on time (years, days, hours, minutes) based data, to derive hidden insights to make informed decision making.

Time series models are very useful models when you have serially correlated data. Most of business houses work on time series data to analyze sales number for the next year, website traffic, competition position and much more. However, it is also one of the areas, which many analysts do not understand.

Q. What are the Components of a Time Series?

1. **Trend:** Series could be **constantly increasing or decreasing** or **first decreasing for a considerable time period and then decreasing**. This trend is identified and then removed from the time series in ARIMA forecasting process.

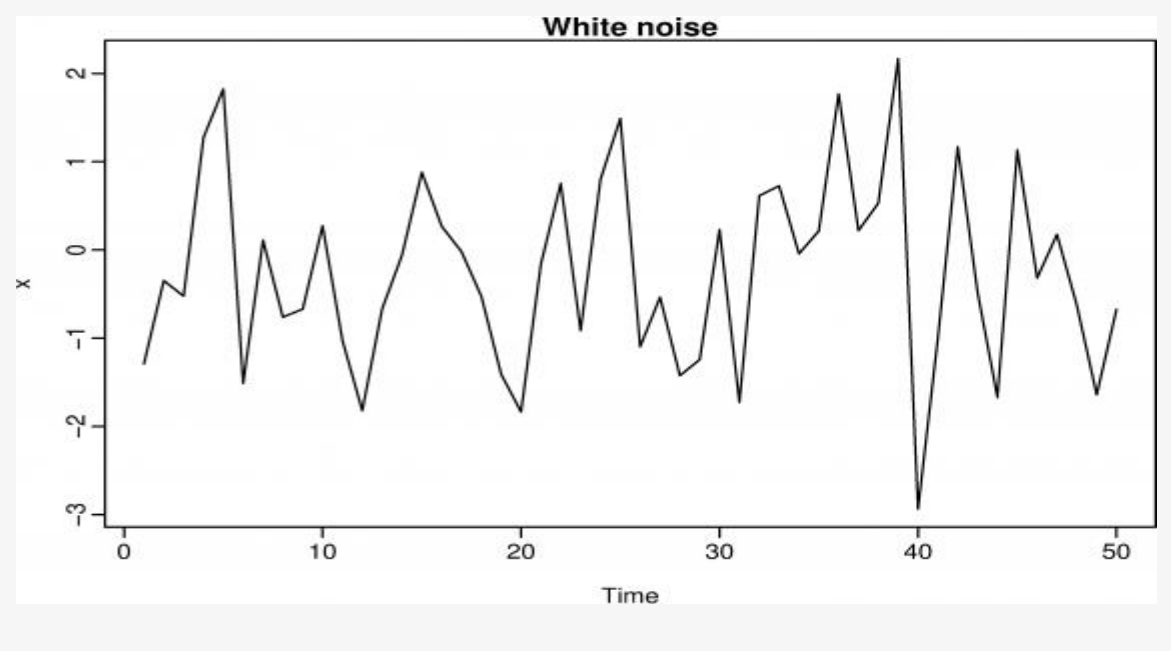
2. **Seasonality:** Repeating pattern with fixed period.
Example - Sales in festive seasons. Sales of Candies and sales of Chocolates peaks in every October Month and December month respectively every year in US. It is because of Halloween and Christmas falling in those months. The time-series should be de-seasonalized in ARIMA forecasting process.

3. **Random Variation (Irregular Component)**
This is the unexplained variation in the time-series which is totally random. Erratic movements that are not predictable because they do not follow a pattern.
Example - Earthquake

Q. What are the terminologies related to Time Series?

1. Stationary Series: A stationary series should have mean and variance of the series is constant over time.The series has to be stationary before building a time series with ARIMA. Most of the time series are non-stationary. If series is non-stationary, we need to make it stationary with detrending, differencing etc.

2. White Noise: A white noise process is one with a constant mean of zero, a constant variance and no correlation between its values at different times. White noise series exhibit a very erratic, jumpy, unpredictable behavior. Since values are uncorrelated, previous values do not help us to forecast future values.

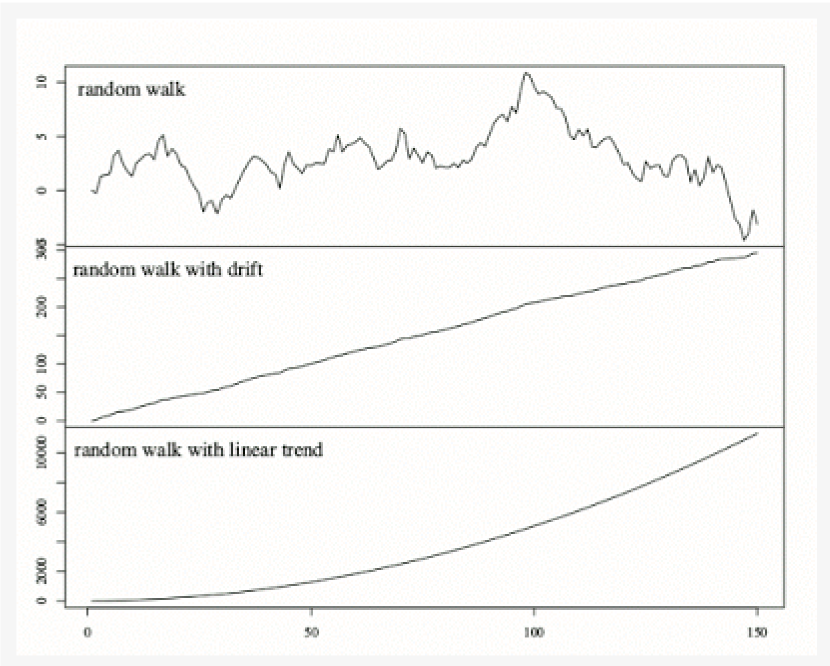


3. Autocorrelation: Autocorrelation refers to the correlation of a time series with its own past and future values. Autocorrelation is also sometimes called "lagged correlation" or "serial correlation".

4. Random Walk: In layman's term, it means past data provides no information about the direction of future movements.
It is called **random-walk-without-drift model**: it assumes that, at each point in time, the series merely takes a random step away from its last recorded position, with steps whose mean value is zero.

If the mean step size is some non-zero value α, the process is said to be a **random-walk-with - drift** (slow steady change) whose prediction equation is **$\hat{Y}t = Yt\text{-}1 + \alpha$**

A random walk process is non-stationary as its mean and variance increases with t.



Time Series : Random Walk

Q. Why Stationary?

To calculate the expected value, we generally take a mean across time intervals. The mean across many time intervals makes sense only when the expected value is the same across those time periods. If the mean and population variance can vary, there is no point estimating by taking an average across time.

Q. What is ARIMA (Box-Jenkins Approach) process?

ARIMA stands for Auto-Regressive Integrated Moving Average. It is also known as **Box-Jenkins approach**. It is one of the most popular techniques used for time series analysis and forecasting purpose.

We would cover ARIMA in a series of blogs starting from introduction, theory and finally the process of performing ARIMA.

Well, coming back to ARIMA, as its full form indicates that it involves two components:

1. **Auto-regressive component**
2. **Moving average component**

We would first understand these components one by one.

Q. what is an Auto-regressive Component?

It implies relationship of a value of a series at a point of time with its own previous values. Such relationship can exist with any order of lag.

Lag: Lag is basically value at a previous point of time. It can have various orders as shown in the table below. It hints toward a pointed relationship.

Month	Sales	Lag(Sales)	Lag2(Sales)
Jan-15	100	-	-
Feb-15	103	100	-
Mar-15	98	103	100
Apr-15	116	98	103
May-15	120	116	98
Jun-15	100	120	116
Jul-15	130	100	120
Aug-15	133	130	100
Sep-15	104	133	130
Oct-15	137	104	133
Nov-15	143	137	104
Dec-15	105	143	137

Time Series : Lag

Q. What is Moving average components?

It implies the current deviation from mean depends on previous deviations. Such relationship can exist with any number of lags which decides the order of moving average.

Moving Average -

Moving Average is average of consecutive values at various time periods. It can have various orders as shown in the table below. It hints toward a distributed relationship as moving itself is derivative of various lags.

Month	Sales	Movave2(Sales)	Movave3(Sales)
Jan-15	100	100	100
Feb-15	103	101.5	101.5
Mar-15	98	100.5	100.33
Apr-15	116	107	105.67
May-15	120	118	111.33
Jun-15	100	110	112.00
Jul-15	130	115	116.67
Aug-15	133	131.5	121.00
Sep-15	104	118.5	122.33
Oct-15	137	120.5	124.67
Nov-15	143	140	128.00
Dec-15	105	124	128.33

Moving Average Explanation

Moving average is itself considered as one of the most rudimentary methods of forecasting. So if you drag the average formula in excel further (beyond Dec-15), it would give you forecast for next month. Both Auto-regressive (lag based) and moving average components in conjunction are used by ARIMA technique for forecasting a time series.

Q. What is autoregression (AR)?

Autoregression is a model that predicts a variable based on its own past values. It assumes that the relationship between the variable and its past values can be represented linearly.

Q. What is autoregressive integrated moving average (ARIMA)?

ARIMA is a model that combines autoregressive (AR), moving average (MA), and differencing (I) components to capture patterns and trends in time series data.

Q. What is the difference between ARMA and ARIMA models?

ARMA models do not involve differencing, while ARIMA models include differencing to make the data stationary.

Q. What is the order of an ARIMA model?

The order of an ARIMA model is represented as (p, d, q), where p is the autoregressive order, d is the differencing order, and q is the moving average order.

Q. How can you determine the order of an ARIMA model?

The order of an ARIMA model can be determined by analyzing the autocorrelation function (ACF) and partial autocorrelation function (PACF) plots of the differenced data.

Q. What is seasonal ARIMA (SARIMA)?

SARIMA is an extension of ARIMA that incorporates seasonal components to model and forecast time series data with seasonal patterns.

Q. What is exponential smoothing?

Exponential smoothing is a time series forecasting method that assigns exponentially decreasing weights to past observations, giving more importance to recent values.

Q. What is the difference between simple exponential smoothing and Holt's linear exponential smoothing?

Simple exponential smoothing uses a single smoothing factor, while Holt's linear exponential smoothing uses two smoothing factors to model the level and trend of the data.

Q. What is the difference between additive and multiplicative seasonality?

In additive seasonality, the seasonal pattern has a constant amplitude, while in multiplicative seasonality, the amplitude of the seasonal pattern varies with the level of the data.

Q. What is a Fourier transform?

Fourier transform is a mathematical technique that decomposes a time series into its constituent frequencies, revealing the presence of different periodic components.

Q. What is a periodogram?

A periodogram is a plot that displays the power spectrum of a time series, providing information about the dominant frequencies.

Q. What is the Box-Jenkins methodology?

The Box-Jenkins methodology is an approach for identifying, estimating, and diagnosing ARIMA models to make time series forecasts.

Q. What is cross-validation in time series analysis?

Cross-validation in time series analysis involves splitting the data into multiple training and testing sets, considering the temporal order of the data, to evaluate the performance of a model.

Q. What is the difference between in-sample and out-of-sample forecasting?

In-sample forecasting refers to predicting the values of the time series within the range of the observed data, while out-of-sample forecasting refers to predicting values beyond the observed data.

Q. What are some popular Python libraries for time series analysis?

Some popular Python libraries for time series analysis are pandas, NumPy, statsmodels, scikit-learn, and Prophet.

Q. What is the difference between MAE, MSE, and RMSE?

MAE (Mean Absolute Error), MSE (Mean Squared Error), and RMSE (Root Mean Squared Error) are evaluation metrics used to measure the accuracy of time series forecasts. MAE and MSE are in the original scale of the data, while RMSE is in the same scale as the data.

Q. What is seasonal decomposition of time series (STL)?

STL (Seasonal and Trend decomposition using Loess) is a method that decomposes a time series into its seasonal, trend, and residual components.

Q. How can you handle missing values in time series data?

Missing values in time series data can be handled by interpolation, forward or backward filling, or using advanced techniques like linear regression or time series imputation methods.

Q. What is the difference between interpolation and extrapolation in time series analysis?

Interpolation involves estimating missing values within the observed range of the time series, while extrapolation involves estimating values beyond the observed range.

Q. What is the purpose of detrending in time series analysis?

Detrending is done to remove the trend component from the time series, allowing for a better analysis of the remaining components.

Q. What is the purpose of deseasonalizing in time series analysis?

Deseasonalizing is done to remove the seasonal component from the time series, allowing for a better analysis of the remaining components.

Q. What is a lag plot?

A lag plot is a graphical tool used to visualize the relationship between an observation and its lagged values.

Q. What is autocorrelation?

Autocorrelation measures the correlation between a time series and its lagged values at different time lags.

Q. What is the Durbin-Watson statistic?

The Durbin-Watson statistic is used to test for the presence of autocorrelation in the residuals of a time series model.

Q. What is the difference between univariate and multivariate time series analysis?

Univariate time series analysis involves modeling and forecasting a single variable, while multivariate time series analysis involves modeling and forecasting multiple variables simultaneously.

Q. What are the advantages of using neural networks for time series analysis?

Neural networks can capture complex nonlinear relationships and adaptively learn from data, making them suitable for modelling and forecasting time series data.

Q. What is long short-term memory (LSTM)?

LSTM is a type of recurrent neural network (RNN) architecture designed to model and predict sequential data, such as time series, by addressing the vanishing gradient problem.

Q. What is a sliding window approach in time series forecasting?

The sliding window approach involves training a time series model on a fixed-size window of past observations and using it to predict the next future value. The window slides forward in time, incorporating new observations for each prediction.

Q. What is a rolling mean/median in time series analysis?

The rolling mean/median is a technique that calculates the mean/median of a fixed-size window of past observations, providing a smoothed representation of the time series.

Q. What is the difference between interpolation and extrapolation in time series analysis?

Interpolation involves estimating missing values within the observed range of the time series, while extrapolation involves estimating values beyond the observed range.

Q. What is the difference between a naive forecast and an exponential smoothing forecast?

A naive forecast simply assumes that the future values of a time series will be the same as the most recent observed value, while exponential smoothing assigns exponentially decreasing weights to past observations.

Q. What is the Box-Cox transformation?

The Box-Cox transformation is a technique used to stabilize the variance of a time series by applying a power transformation. It can be useful when dealing with heteroscedasticity in the data.

Q. How can you handle outliers in time series analysis?

Outliers in time series data can be identified using techniques like the Z-score or the modified Z-score method and can be handled by smoothing, replacing with interpolated values, or removing them from the analysis.

Q. What are some popular time series forecasting algorithms?

Some popular time series forecasting algorithms include ARIMA, SARIMA, exponential smoothing methods (such as Holt-Winters), Prophet, and various machine learning algorithms like random forest, support vector machines, and neural networks.

Q. can you explain ARIMA Model Steps?

The ARIMA (Autoregressive Integrated Moving Average) model is a popular time series forecasting technique. It combines the autoregressive (AR), differencing (I), and moving average (MA) components to capture patterns and trends in time series data. Here's a step-by-step process to build an ARIMA model:

Step 1: Understand the Data
- Examine the time series data you want to model.
- Identify any patterns, trends, or seasonality present in the data.
- Check for missing values or outliers that may need to be addressed.

Step 2: Stationarity
- ARIMA models require the time series data to be stationary.
- Stationarity means that the statistical properties of the data, such as mean and variance, remain constant over time.
- Perform tests like the Augmented Dickey-Fuller (ADF) test to check for stationarity.
- If the data is non-stationary, apply differencing to make it stationary. Differencing involves subtracting the previous observation from the current observation.

Step 3: Determine Order (p, d, q)
- The order of the ARIMA model is denoted as (p, d, q), where:

- p represents the autoregressive (AR) order, which captures the relationship between an observation and a number of lagged observations.
- d represents the differencing order, which is the number of times differencing was applied to make the data stationary.
- q represents the moving average (MA) order, which captures the dependency between an observation and a residual error from a moving average model.
- Determine the order values by analyzing the autocorrelation function (ACF) and partial autocorrelation function (PACF) plots of the differenced data.
- ACF indicates the correlation between the series and its lagged values, while PACF shows the correlation between the series and its lagged values after removing the effect of other lagged values.

Step 4: Fit the ARIMA Model
- Using the determined order (p, d, q), fit the ARIMA model to the data.
- Split the data into training and testing sets, keeping the most recent observations for testing.
- Use a package or library that supports ARIMA modeling, such as statsmodels in Python.
- Fit the ARIMA model to the training data.

Step 5: Model Evaluation
- Validate the ARIMA model by comparing its predictions against the actual values in the test set.
- Calculate appropriate evaluation metrics like mean squared error (MSE), root mean squared error (RMSE), or mean absolute error (MAE).
- Adjust the model parameters if the performance is not satisfactory.

Step 6: Forecasting
- Once you are satisfied with the model's performance, use it to make forecasts on new, unseen data.
- Provide the necessary input data and use the fitted ARIMA model to generate predictions.

Step 7: Model Refinement
- Refine the ARIMA model by iterating through steps 3 to 6, adjusting the order and model parameters as needed.
- Compare the performance of different models using evaluation metrics and choose the one that provides the best results.

Remember that building an accurate ARIMA model often requires experimentation and iterative refinement. Additionally, there are variations and extensions of ARIMA, such as seasonal ARIMA (SARIMA), which consider seasonal patterns in the data.

Certainly! Here are additional time series-related questions and answers:

Q. What is the difference between point forecasting and interval forecasting?

Point forecasting provides a single predicted value for a future time point, while interval forecasting provides a range of possible values within a certain confidence level.

Q. What is the difference between a time series and a cross-sectional data?

A time series data is collected over a sequence of time periods, while cross-sectional data is collected at a specific point in time for different entities or individuals.

Q. What is the partial autocorrelation function (PACF)?

The partial autocorrelation function (PACF) measures the correlation between a time series and its lagged values, while removing the effects of the intermediate lags.

Q. What is the purpose of differencing in time series analysis?

Differencing is used to make a time series stationary by subtracting the previous observation from the current observation, removing trends or seasonality.

Q. What is a unit root in time series analysis?

A unit root is a characteristic of non-stationary time series, indicating that the series has a stochastic trend and its statistical properties do not remain constant over time.

Q. What is the purpose of seasonal differencing?

Seasonal differencing is used to remove the seasonal component from a time series by subtracting the observation from the same point in the previous season.

Q. What is the concept of white noise in time series analysis?

White noise refers to a time series with uncorrelated and identically distributed random variables, having a mean of zero and constant variance.

Q. What is the difference between exogenous and endogenous variables in time series analysis?

Exogenous variables are independent variables that are not affected by the time series itself, while endogenous variables are influenced by the time series.

Q. What is a forecast horizon in time series forecasting?

The forecast horizon refers to the number of time periods into the future for which predictions are made.

Q. How can you handle non-linear relationships in time series analysis?

Non-linear relationships in time series analysis can be handled by using non-linear models like polynomial regression, non-linear regression, or non-linear machine learning algorithms.

Q. What is autoregressive conditional heteroscedasticity (ARCH)?

ARCH is a modeling technique used to capture volatility clustering in time series data, where the variance of the error term is time-varying.

Q. What is a dynamic regression model in time series analysis?

A dynamic regression model incorporates exogenous variables that can affect the time series, allowing for more accurate forecasting.

Q. What is a stationary ARMA process?

A stationary ARMA process is a time series model where both the autoregressive (AR) and moving average (MA) components are stationary.

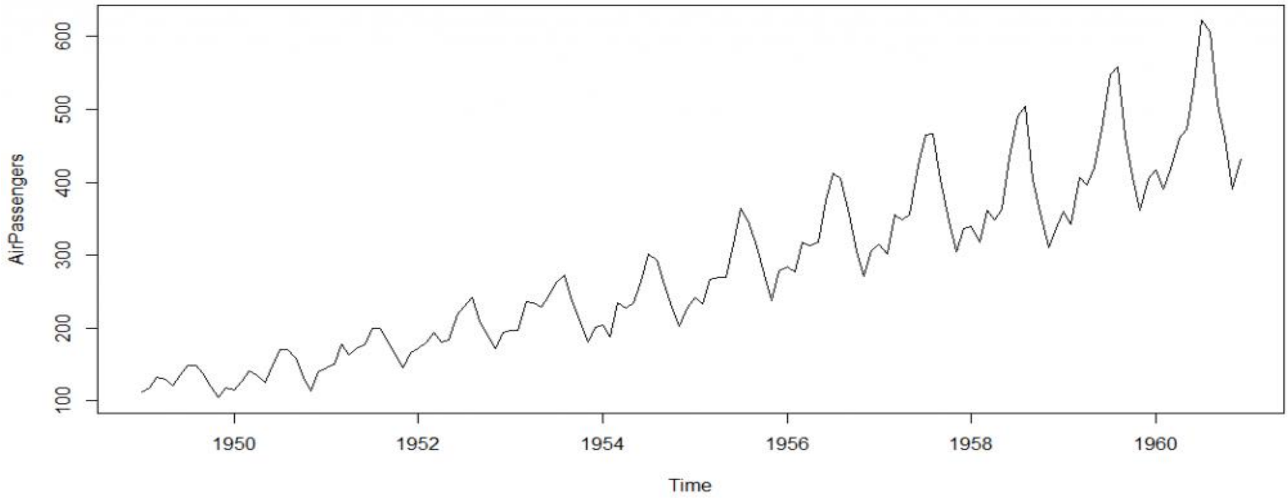
Q. What is the concept of impulse response in time series analysis?

Impulse response measures the impact of a shock or an innovation in a time series on the subsequent values of the series.

Q. What is a seasonality test in time series analysis?

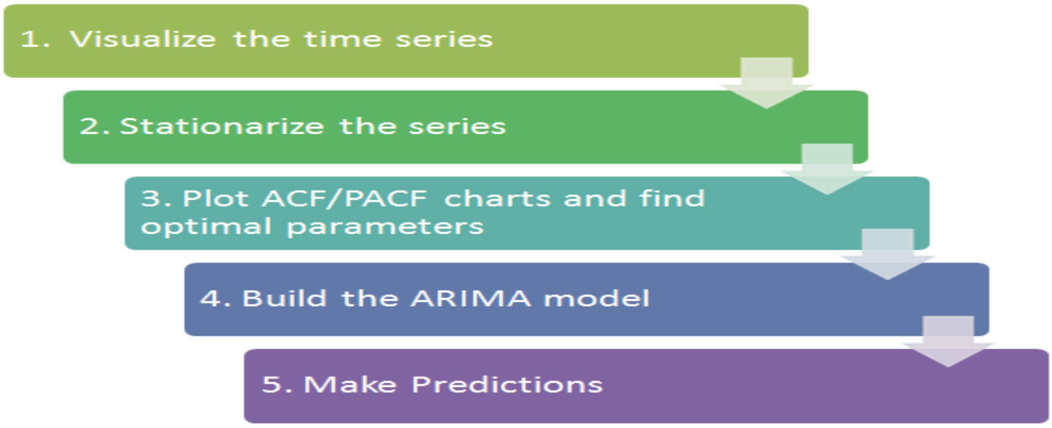
A seasonality test examines the presence of seasonal patterns in a time series using statistical tests or visual inspection of the data.

Q. What are the observations you can make from the below plot?
Ans:



Here are my observations:
1. There is a trend component which grows the passenger year by year.
2. There looks to be a seasonal component which has a cycle less than 12 months.
3. The variance in the data keeps on increasing with time.
We know that we need to address two issues before we test stationary series. One, we need to remove unequal variances. We do this using log of the series. Two, we need to address the trend component. We do this by taking difference of the series.

Q: Can you provide quick overview of ARIMA Time Series Modelling
Ans: Step by step approach on '**How to do a Time Series Analysis**:



Step 1: Visualize the Time Series
It is essential to analyze the trends prior to building any kind of time series model. The details we are interested in pertains to any kind of trend, seasonality or random behaviour in the series. We have covered this part in the second part of this series.

Step 2: Stationarize the Series
Once we know the patterns, trends, cycles and seasonality, we can check if the series is stationary or not. Dickey – Fuller is one of the popular test to check the same. We have covered this test in the first part of this article series. This doesn't ends here! What if the series is found to be non-stationary?
There are three commonly used technique to make a time series stationary:

1. **Detrending**: Here, we simply remove the trend component from the time series. For instance, the equation of my time series is:
x(t) = (mean + trend * t) + error

We'll simply remove the part in the parentheses and build model for the rest.

2. **Differencing**: This is the commonly used technique to remove non-stationarity. Here we try to model the differences of the terms and not the actual term. For instance,
x(t) – x(t-1) = ARMA (p , q)

This differencing is called as the Integration part in AR(I)MA. Now, we have three parameters
p : AR
d : I
q : MA

3. **Seasonality**: Seasonality can easily be incorporated in the ARIMA model directly. More on this has been discussed in the applications part below.

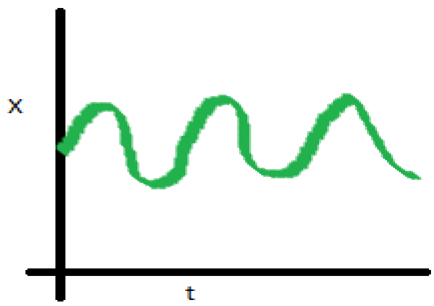
Step 3: Find Optimal Parameters
The parameters p,d,q can be found using ACF and PACF plots. An addition to this approach is can be, if both ACF and PACF decrease gradually, it indicates that we need to make the time series stationary and introduce a value to "d".

Step 4: Build ARIMA Model
With the parameters in hand, we can now try to build ARIMA model. The value found in the previous step might be an approximate estimate and we need to explore more (p,d,q) combinations. The one with the lowest BIC and AIC should be our choice. We can also try some models with a seasonal component. Just in case, we notice any seasonality in ACF/PACF plots.

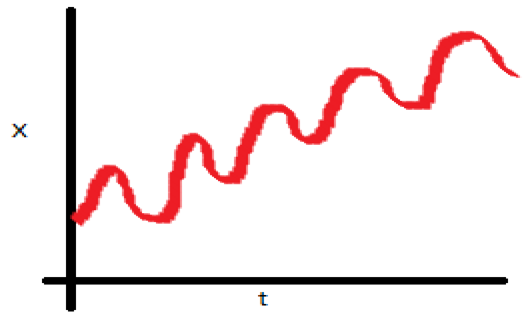
Step 5: Make Predictions
Once we have the final ARIMA model, we are now ready to make predictions on the future time points. We can also visualize the trends to cross validate if the model works fine.

Q. Can you elaborate on Stationarity of time series?

Stationary Series
There are three basic criteria for a series to be classified as stationary series:
1. The mean of the series should not be a function of time rather should be a constant. The image below has the left hand graph satisfying the condition whereas the graph in red has a time dependent mean.

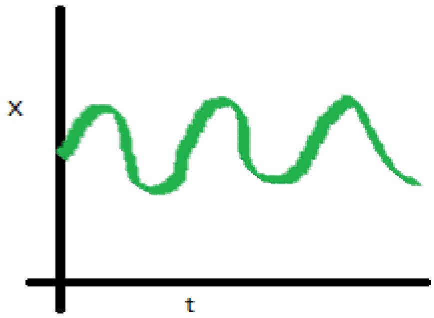


Stationary series

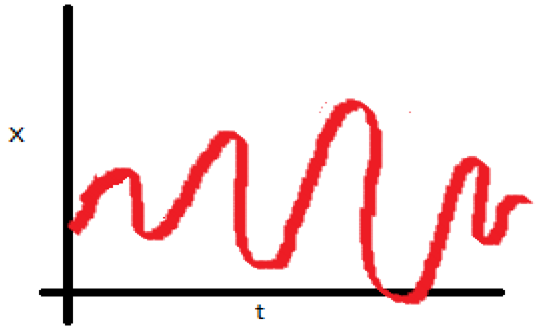


Non-Stationary series

2. The variance of the series should not be a function of time. This property is known as homoscedasticity. Following graph depicts what is and what is not a stationary series. (Notice the varying spread of distribution in the right hand graph)

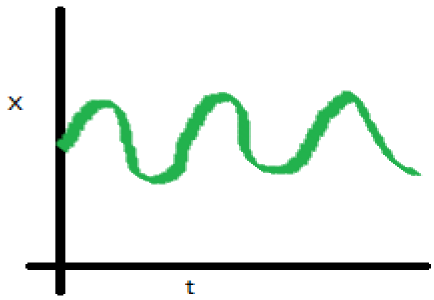


Stationary series

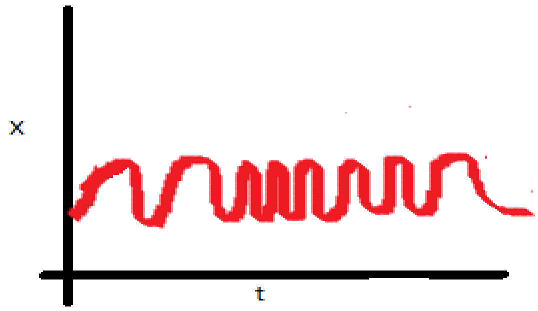


Non-Stationary series

3. The covariance of the i th term and the (i + m) th term should not be a function of time. In the following graph, you will notice the spread becomes closer as the time increases. Hence, the covariance is not constant with time for the 'red series'.



Stationary series



Non-Stationary series

Q. What is Random Walk?
Ans: This is the most basic concept of the time series. Let's take an example.

Example: Imagine a girl moving randomly on a giant chess board. In this case, next position of the girl is only dependent on the last position.

Now imagine, you are sitting in another room and are not able to see the girl. You want to predict the position of the girl with time. How accurate will you be? Of course you will become more and more inaccurate as the position of the girl changes. At t=0 you exactly know where the girl is. Next time, she can only move to 8 squares and hence your probability dips to 1/8 instead of 1 and it keeps on going down. Now let's try to formulate this series :X(t) = X(t-1) + Er(t)

where Er(t) is the error at time point t. This is the randomness the girl brings at every point in time. Now, if we recursively fit in all the Xs, we will finally end up to the following equation:
X(t) = X(0) + Sum(Er(1),Er(2),Er(3).....Er(t))

Q. Is random work stationary series?

Ans: Is the Mean constant?

E[X(t)] = E[X(0)] + Sum(E[Er(1)],E[Er(2)],E[Er(3)].....E[Er(t)])
We know that Expectation of any Error will be zero as it is random.
Hence we get E[X(t)] = E[X(0)] = Constant.

Is the Variance constant?

Var[X(t)] = Var[X(0)] + Sum(Var[Er(1)],Var[Er(2)],Var[Er(3)].....Var[Er(t)])
Var[X(t)] = t * Var(Error) = Time dependent.
Hence, we infer that the random walk is not a stationary process as it has a time variant variance. Also, if we check the covariance, we see that too is dependent on time.

Objective Type Questions

Q. Which of the following is an example of time series problem?

1. Estimating number of hotel rooms booking in next 6 months.
2. Estimating the total sales in next 3 years of an insurance company.
3. Estimating the number of calls for the next one week.

A) Only 3
B) 1 and 2
C) 2 and 3
D) 1 and 3
E) 1,2 and 3

Solution: (E)

All the above options have a time component associated.

Q. Which of the following is not an example of a time series model?

A) Naive approach
B) Exponential smoothing
C) Moving Average
D) None of the above

Solution: (D)

Naïve approach: Estimating technique in which the last period's actuals are used as this period's forecast, without adjusting them or attempting to establish causal factors. It is used only for comparison with the forecasts generated by the better (sophisticated) techniques.

In exponential smoothing, older data is given progressively-less relative importance whereas newer data is given progressively-greater importance.

In time series analysis, the moving-average (MA) model is a common approach for modelling univariate time series. The moving-average model specifies that the output variable depends linearly on the current and various past values of a stochastic (imperfectly predictable) term.

Q. Which of the following can't be a component for a time series plot?

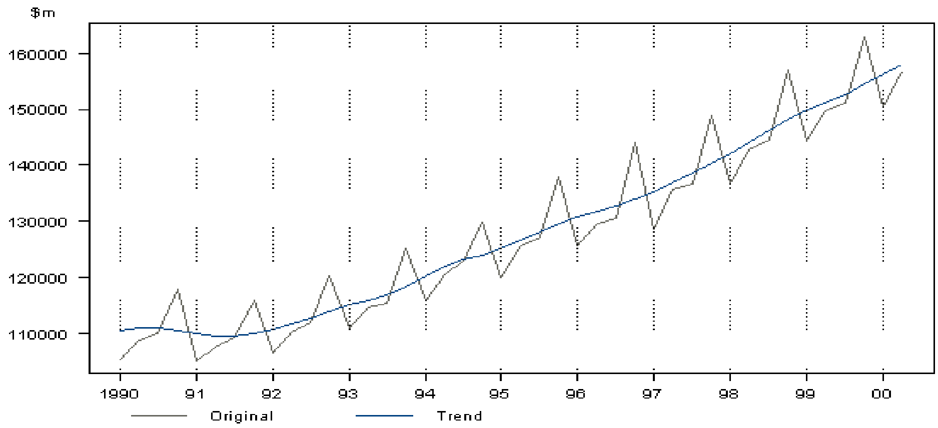
A) Seasonality
B) Trend
C) Cyclical
D) Noise
E) None of the above

Solution: (E)

A seasonal pattern exists when a series is influenced by seasonal factors (e.g., the quarter of the year, the month, or day of the week). Seasonality is always of a fixed and known period. Hence, seasonal time series are sometimes called periodic time series

Seasonality is always of a fixed and known period. A cyclic pattern exists when data exhibit rises and falls that are not of fixed period.

Trend is defined as the 'long term' movement in a time series without calendar related and irregular effects, and is a reflection of the underlying level. It is the result of influences such as population growth, price inflation and general economic changes. The following graph depicts a series in which there is an obvious upward trend over time.

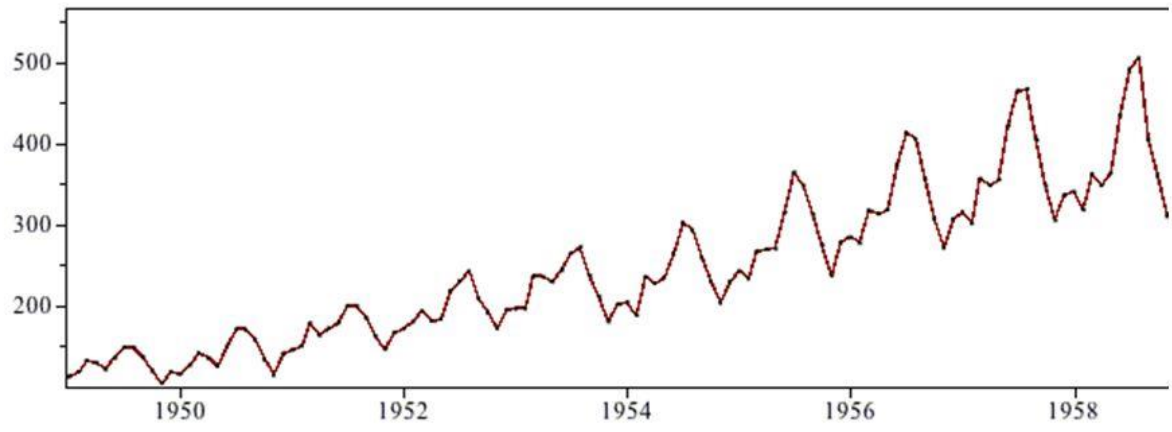


Quarterly Gross Domestic Product

Noise: In discrete time, white noise is a discrete signal whose samples are regarded as a sequence of serially uncorrelated random variables with zero mean and finite variance.
Thus all of the above mentioned are components of a time series.

Q. Which of the following is relatively easier to estimate in time series modeling?
A) Seasonality
B) Cyclical
C) No difference between Seasonality and Cyclical

Solution: (A)
As we seen in previous solution, as seasonality exhibits fixed structure; it is easier to estimate.
Q. The below time series plot contains both Cyclical and Seasonality component.



A) TRUE
B) FALSE

Solution: (B)
There is a repeated trend in the plot above at regular intervals of time and is thus only seasonal in nature.

Q. Adjacent observations in time series data (excluding white noise) are independent and identically distributed (IID).

A) TRUE

B) FALSE

Solution: (B)

Clusters of observations are frequently correlated with increasing strength as the time intervals between them become shorter. This needs to be true because in time series forecasting is done based on previous observations and not the currently observed data unlike classification or regression.

Q. Smoothing parameter close to one gives more weight or influence to recent observations over the forecast.

A) TRUE

B) FALSE

Solution: (A)

It may be sensible to attach larger weights to more recent observations than to observations from the distant past. This is exactly the concept behind simple exponential smoothing. Forecasts are calculated using weighted averages where the weights decrease exponentially as observations come from further in the past — the smallest weights are associated with the oldest observations:

$$Y_{t+1}=\alpha Y_t+\alpha(1-\alpha)Y_{t-1}+\alpha(1-\alpha)(1-\alpha)Y_{t-2}+\cdots$$

where $0\leq\alpha\leq1$ is the smoothing parameter. The one-step-ahead forecast for time t+1 is a weighted average of all the observations in the series $Y_1,\ldots,Y_t$. The rate at which the weights decrease is controlled by the parameter α.

Q. Sum of weights in exponential smoothing is _____.

A) <1

B) 1

C) >1

D) None of the above

Solution: (B)

Below table shows the weights attached to observations for four different values of α when forecasting using simple exponential smoothing. Note that the sum of the weights even for a small α will be approximately one for any reasonable sample size.

Observation	α=0.2	α=0.4	α=0.6	α=0.8
y_T	0.2	0.4	0.6	0.8
y_{T-1}	0.16	0.24	0.24	0.16
y_{T-2}	0.128	0.144	0.096	0.032
y_{T-3}	0.102	0.0864	0.0384	0.0064
y_{T-4}	(0.2)(0.8)	(0.4)(0.6)	(0.6)(0.4)	(0.8)(0.2)
y_{T-5}	(0.2)(0.8)	(0.4)(0.6)	(0.6)(0.4)	(0.8)(0.2)

Q. The last period's forecast was 70 and demand was 60. What is the simple exponential smoothing forecast with alpha of 0.4 for the next period?

A) 63.8
B) 65
C) 62
D) 66
Solution: (D)
Yt-1= 70, St-1= 60, Alpha = 0.4

Substituting the values we get
0.4*60 + 0.6*70= 24 + 42= 66

Q. What does auto covariance measure?
A) Linear dependence between multiple points on the different series observed at different times
B) Quadratic dependence between two points on the same series observed at different times
C) Linear dependence between two points on different series observed at same time
D) Linear dependence between two points on the same series observed at different times

Solution: (D)
Option D is the definition of auto covariance.
Q. Which of the following is not a necessary condition for weakly stationary time series?
A) Mean is constant and does not depend on time
B) Autocovariance function depends on s and t only through their difference |s-t| (where t and s are moments in time)
C) The time series under considerations is a finite variance process
D) Time series is Gaussian

Solution: (D)
A Gaussian time series implies stationarity is strict stationarity.

Q. Which of the following is not a technique used in smoothing time series?
A) Nearest Neighbor Regression
B) Locally weighted scatter plot smoothing
C) Tree based models like (CART)
D) Smoothing Splines

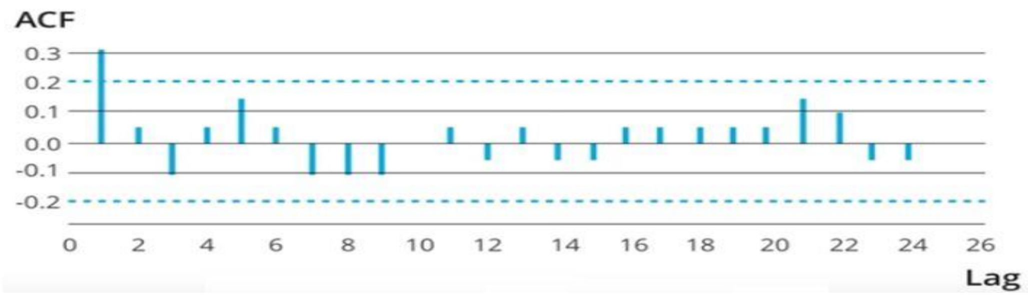
Solution: (C)
Time series smoothing and filtering can be expressed in terms of local regression models. Polynomials and regression splines also provide important techniques for smoothing. CART based models do not provide an equation to superimpose on time series and thus cannot be used for smoothing. All the other techniques are well documented smoothing techniques.

Q. If the demand is 100 during October 2016, 200 in November 2016, 300 in December 2016, 400 in January 2017. What is the 3-month simple moving average for February 2017?
A) 300
B) 350
C) 400

D) Need more information

Solution: (A)
X`= (xt-3 + xt-2 + xt-1) /3
(200+300+400)/ 3 = 900/3 =300

Q. Looking at the below ACF plot, would you suggest to apply AR or MA in ARIMA modeling technique?



A) AR
B) MA
C) Can't Say

Solution: (A)
MA model is considered in the following situation, If the autocorrelation function (ACF) of the differenced series displays a sharp cutoff and/or the lag-1 autocorrelation is negative–i.e., if the series appears slightly "overdifferenced"–then consider adding an MA term to the model. The lag beyond which the ACF cuts off is the indicated number of MA terms.

But as there are no observable sharp cutoffs the AR model must be preferred.

Q. Suppose, you are a data scientist and you observed the views on the website increases during the month of Jan-Mar. whereas the views during Nov-Dec decreases. Does the above statement represent seasonality?
A) TRUE
B) FALSE
C) Can't Say

Solution: (A)
Yes this is a definite seasonal trend as there is a change in the views at particular times.
Remember, Seasonality is a presence of variations at specific periodic intervals.

Q. Which of the following graph can be used to detect seasonality in time series data?
Ans:
1. Multiple box
2. Autocorrelation

A) Only 1
B) Only 2

C) 1 and 2
D) None of these

Solution: (C)
Seasonality is a presence of variations at specific periodic intervals.

The variation of distribution can be observed in multiple box plots. And thus seasonality can be easily spotted. Autocorrelation plot should show spikes at lags equal to the period.

Q. Stationarity is a desirable property for a time series process.
A) TRUE
B) FALSE
Solution: (A)
When the following conditions are satisfied then a time series is stationary.

Mean is constant and does not depend on time
Autocovariance function depends on s and t only through their difference |s-t| (where t and s are moments in time)
The time series under considerations is a finite variance process
These conditions are essential prerequisites for mathematically representing a time series to be used for analysis and forecasting. Thus stationarity is a desirable property.

Q. Suppose you are given a time series dataset which has only 4 columns (id, Time, X, Target).
Ans:

id	Time	X	Target
1	1	100	10
2	2	200	20
3	3	300	30
1	4	400	40
2	5	500	50
3	6	600	60
1	7	500	50
2	8	400	40
3	9	500	30
4	10	700	20

What would be the rolling mean of feature X if you are given the window size 2?
Note: X column represents rolling mean.

Solution:

Quater	Time	X'	Target
1	1	NaN	10
2	2	NaN	20
3	3	150	30
1	4	250	40
2	5	350	50
3	6	450	60
1	7	550	50
2	8	550	40
3	9	450	30
4	10	450	20

X`= xt-2 + xt-1 /2
Based on the above formula: (100 +200) /2 =150; (200+300)/2 = 250 and so on.

Q. Imagine, you are working on a time series dataset. Your manager has asked you to build a highly accurate model. You started to build two types of models which are given below.

Model 1: Decision Tree model
Model 2: Time series regression model

At the end of evaluation of these two models, you found that model 2 is better than model 1. What could be the possible reason for your inference?
A) Model 1 couldn't map the linear relationship as good as Model 2
B) Model 1 will always be better than Model 2
C) You can't compare decision tree with time series regression
D) None of these

Solution: (A)
A time series model is similar to a regression model. So it is good at finding simple linear relationships. While a tree based model though efficient will not be as good at finding and exploiting linear relationships.

Q. What type of analysis could be most effective for predicting temperature on the following type of data?

Date	Temperature	precipitation	temperature/precipitation
12/12/12	7	0.2	35
13/12/12	9	0.123	73.1707317073
14/12/12	9.2	0.34	27.0588235294
15/12/12	10	0.453	22.0750551876
16/12/12	12	0.33	36.3636363636
17/12/12	11	0.8	13.75

A) Time Series Analysis
B) Classification
C) Clustering
D) None of the above

Solution: (A)
The data is obtained on consecutive days and thus the most effective type of analysis will be time series analysis.

Q. What is the first difference of temperature / precipitation variable?

Date	Temperature	precipitation	temperature/precipitation
12/12/12	7	0.2	35
13/12/12	9	0.123	73.1707317073
14/12/12	9.2	0.34	27.0588235294
15/12/12	10	0.453	22.0750551876
16/12/12	12	0.33	36.3636363636
17/12/12	11	0.8	13.75

A) 15,12.2,-43.2,-23.2,14.3,-7
B) 38.17,-46.11,-4.98,14.29,-22.61
C) 35,38.17,-46.11,-4.98,14.29,-22.61
D) 36.21,-43.23,-5.43,17.44,-22.61

Solution: (B)
73.17-35 = 38.17
27.05-73.17 = – 46.11 and so on..

Q. Consider the following set of data:
{23.32 32.33 32.88 28.98 33.16 26.33 29.88 32.69 18.98 21.23 26.66 29.89}
What is the lag-one sample autocorrelation of the time series?

A) 0.26
B) 0.52
C) 0.13
D) 0.07

Solution: (C)
$\hat{\rho}_1 = (23.32-\bar{x})(32.33-\bar{x})+(32.33-\bar{x})(32.88-\bar{x})+\cdots = 0.130394786$

Where $\bar{x}$ is the mean of the series which is 28.0275

Q. Any stationary time series can be approximately the random superposition of sines and cosines oscillating at various frequencies.
A) TRUE
B) FALSE

Solution: (A)

A weakly stationary time series, xt, is a finite variance process such that

The mean value function, µt, is constant and does not depend on time t, and (ii) the autocovariance function, γ(s,t), defined in depends on s and t only through their difference |s−t|.

random superposition of sines and cosines oscillating at various frequencies is white noise. white noise is weakly stationary or stationary. If the white noise variates are also normally distributed or Gaussian, the series is also strictly stationary.

Q. Auto covariance function for weakly stationary time series does not depend on _______ ?

A) Separation of xs and xt

B) h = | s – t |

C) Location of point at a particular time

Solution: (C)

By definition of weak stationary time series described in previous question.

Q. Two time series are jointly stationary if _____?

A) They are each stationary

B) Cross variance function is a function only of lag h

A) Only A

B) Both A and B

Solution: (D)

Joint stationarity is defined based on the above two mentioned conditions.

Q. In autoregressive models _______?

A) Current value of dependent variable is influenced by current values of independent variables

B) Current value of dependent variable is influenced by current and past values of independent variables

C) Current value of dependent variable is influenced by past values of both dependent and independent variables

D) None of the above

Solution: (C)

Autoregressive models are based on the idea that the current value of the series, xt, can be explained as a function of p past values, xt−1,xt−2,…,xt−p, where p determines the number of steps into the past needed to forecast the current value. Ex. xt = xt−1 −.90xt−2 + wt,

Where xt-1 and xt-2 are past values of dependent variable and wt the white noise can represent values of independent values.

The example can be extended to include multiple series analogous to multivariate linear regression.

Q. For MA (Moving Average) models the pair σ = 1 and θ = 5 yields the same auto covariance function as the pair σ = 25 and θ = 1/5.

$$x_t = w_t + \frac{1}{5} w_{t-1}, \quad w_t \sim \text{iid } N(0, 25)$$

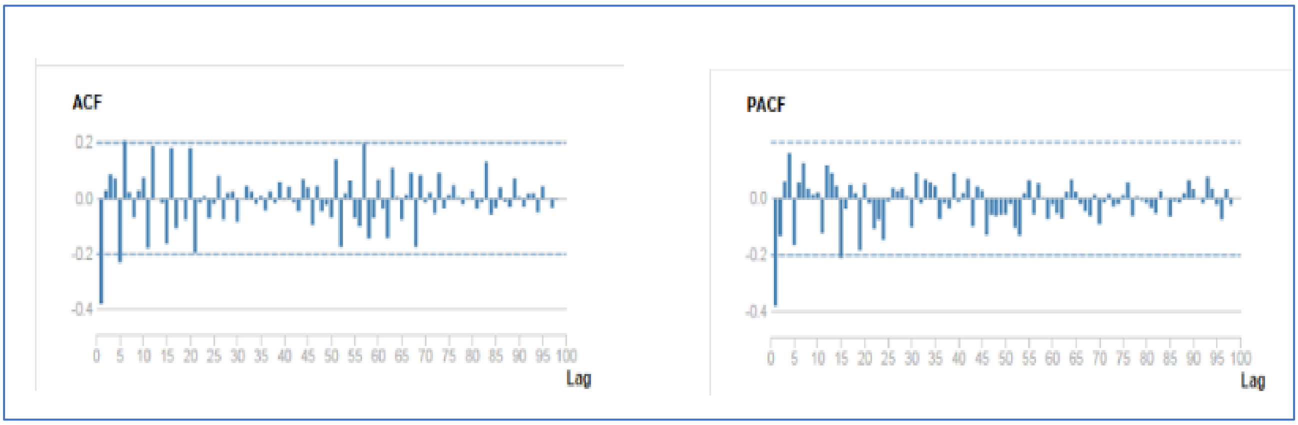
$$y_t = v_t + 5 v_{t-1}, \quad v_t \sim \text{iid } N(0, 1)$$

A) TRUE
B) FALSE

Solution: (A)
True, because autocovariance is invertible for MA models

Note: that for an MA(1) model, ρ(h) is the same for θ and 1 /θ
The pair σ2 w = 1 and θ = 5 yield the same autocovariance function as the pair σ2 w = 25 and θ = 1/5.

Q. How many AR and MA terms should be included for the time series by looking at the above ACF and PACF plots?



A) AR (1) MA(0)
B) AR(0)MA(1)
C) AR(2)MA(1)
D) AR(1)MA(2)
E) Can't Say

Solution: (B)
Strong negative correlation at lag 1 suggest MA and there is only 1 significant lag

Q. Which of the following is true for white noise?
A) Mean =0
B) Zero autocovariances
C) Zero autocovariances except at lag zero
D) Quadratic Variance

Solution: (C)

A white noise process must have a constant mean, a constant variance and no autocovariance structure (except at lag zero, which is the variance).

Q. For the following MA (3) process $Y_t = \mu + E_t + \theta_1 E_{t-1} + \theta_2 E_{t-2} + \theta_3 E_{t-3}$, where σ_t is a zero mean white noise process with variance σ^2

A) ACF = 0 at lag 3
B) ACF =0 at lag 5
C) ACF =1 at lag 1
D) ACF =0 at lag 2
E) ACF = 0 at lag 3 and at lag 5

Solution: (B)

Recall that an MA(q) process only has memory of length q. This means that all of the autocorrelation coefficients will have a value of zero beyond lag q. This can be seen by examining the MA equation, and seeing that only the past q disturbance terms enter into the equation, so that if we iterate this equation forward through time by more than q periods, the current value of the disturbance term will no longer affect y. Finally, since the autocorrelation function at lag zero is the correlation of y at time t with y at time t (i.e. the correlation of y_t with itself), it must be one by definition.

Q. Consider the following AR(1) model with the disturbances having zero mean and unit variance.
$y_t = 0.4 + 0.2y_{t-1} + u_t$ **The (unconditional) variance of y will be given by?**

A) 1.5
B) 1.04
C) 0.5
D) 2

Solution: (B)
Variance of the disturbances divided by (1 minus the square of the autoregressive coefficient
Which in this case is : 1/(1-(0.2^2))= 1/0.96= 1.041

Q. The pacf (partial autocorrelation function) is necessary for distinguishing between ______?
A) An AR and MA model is solution: False
B) An AR and an ARMA is solution: True
C) An MA and an ARMA is solution: False
D) Different models from within the ARMA family

Solution: (B)

Table 3.1. Behavior of the ACF and PACF for ARMA Models

	AR(p)	MA(q)	ARMA(p,q)
ACF	Tails off	Cuts off after lag q	Tails off
PACF	Cuts off after lag p	Tails off	Tails off

Q. Second differencing in time series can help to eliminate which trend?

A) Quadratic Trend

B) Linear Trend

C) Both A & B

D) None of the above

Solution: (A)

The first difference eliminates a linear trend. A second difference, can eliminate a quadratic trend, and so on.

Q. Which of the following cross validation techniques is better suited for time series data?

A) k-Fold Cross Validation

B) Leave-one-out Cross Validation

C) Stratified Shuffle Split Cross Validation

D) Forward Chaining Cross Validation

Solution: (D)

Time series is ordered data. So the validation data must be ordered to. Forward chaining ensures this. It works as follows:

fold 1 : training [1], test [2]

fold 2 : training [1 2], test [3]

fold 3 : training [1 2 3], test [4]

fold 4 : training [1 2 3 4], test [5]

fold 5 : training [1 2 3 4 5], test [6]

Q. BIC penalizes complex models more strongly than the AIC?

A) TRUE

B) FALSE

Solution: (A)

AIC = -2*ln(likelihood) + 2*k,

BIC = -2*ln(likelihood) + ln(N)*k,

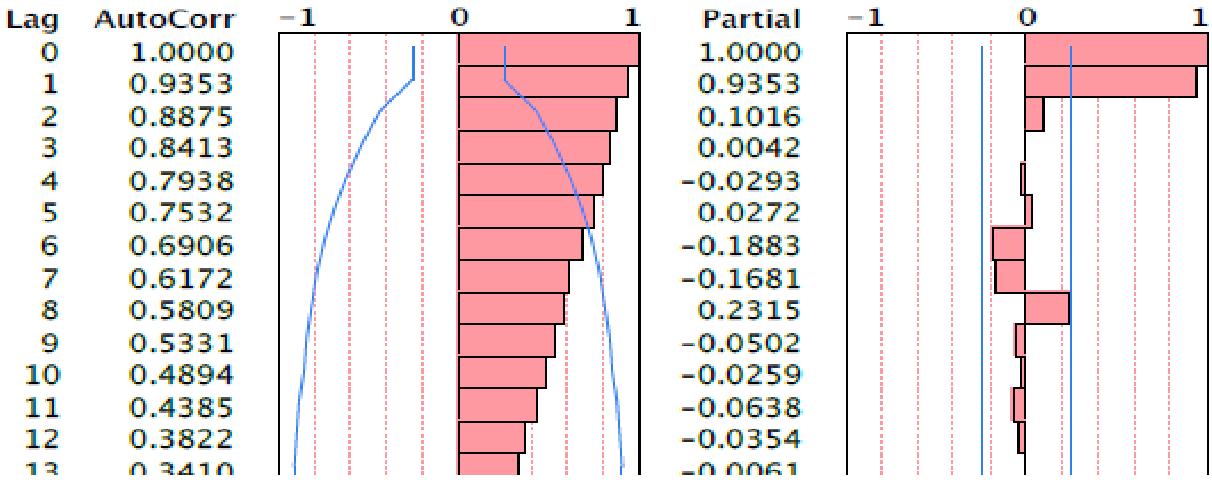
where:

k = model degrees of freedom

N = number of observations

At relatively low N (7 and less) BIC is more tolerant of free parameters than AIC, but less tolerant at higher N (as the natural log of N overcomes 2).

Q. The figure below shows the estimated autocorrelation and partial autocorrelations of a time series of n = 60 observations. Based on these plots, we should.



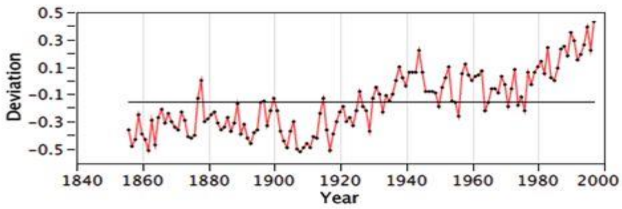
Lag	AutoCorr	Partial
0	1.0000	1.0000
1	0.9353	0.9353
2	0.8875	0.1016
3	0.8413	0.0042
4	0.7938	-0.0293
5	0.7532	0.0272
6	0.6906	-0.1883
7	0.6172	-0.1681
8	0.5809	0.2315
9	0.5331	-0.0502
10	0.4894	-0.0259
11	0.4385	-0.0638
12	0.3822	-0.0354
13	0.3410	-0.0061

A) Transform the data by taking logs
B) Difference the series to obtain stationary data
C) Fit an MA(1) model to the time series

Solution: (B)
The autocorr shows a definite trend and partial autocorrelation shows a choppy trend, in such a scenario taking a log would be of no use. Differencing the series to obtain a stationary series is the only option.

Q. Question Context for next two questions:

The remaining questions consider a time series model for annual global temperature. The data for the time series in this analysis begin in 1856 and run through 1997 (n = 142). The measurements give the deviation from typical temperature in degrees Celsius. (Zero would be considered consistent with the long-run average.)



Model Summary	
DF	140.0000
Sum of Squared Errors	1.7726
Variance Estimate	0.0127
Standard Deviation	0.1125
Akaike's 'A' Information Criterion	-214.4648
Schwarz's Bayesian Criterion	-211.5160
RSquare	0.7328

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Level Smoothing Weight	0.39680005	0.0900926	4.40	<.0001*

	Actual Deviation	Year	Predicted Deviation
133	0.25	1988	0.13245007
134	0.18	1989	0.17909389
135	0.35	1990	0.17945343
136	0.29	1991	0.24712632
137	0.15	1992	0.2641386
138	0.19	1993	0.2188484
139	0.26	1994	0.20740135
140	0.39	1995	0.2282725
141	0.22	1996	0.29244598
142	0.43	1997	0.26369941

Q1. Use the estimated exponential smoothening given above and predict temperature for the next 3 years (1998-2000)
These results summarize the fit of a simple exponential smooth to the time series.

A) 0.2,0.32,0.6
B) 0.33, 0.33,0.33
C) 0.27,0.27,0.27
D) 0.4,0.3,0.37

Solution: (B)
The predicted value from the exponential smooth is the same for all 3 years, so all we need is the value for next year. The expression for the smooth is

Smootht = α yt + (1 – α) smooth t-1Hence, for the next point, the next value of the smooth (the prediction for the next observation) is
Smoothn = α yn + (1 – α) smooth n-1
= 0.3968*0.43 + (1 – 0.3968)* 0.3968
= 0.3297

Q2. Find 95% prediction intervals for the predictions of temperature in 1999.
These results summarize the fit of a simple exponential smooth to the time series.
A) 0.3297 2 * 0.1125
B) 0.3297 2 * 0.121
C) 0.3297 2 * 0.129
D) 0.3297 2 * 0.22

Solution: (B)
The sd of the prediction errors is
1 period out 0.1125
2 periods out 0.1125 sqrt(1+α2) = 0.1125 * sqrt(1+ 0.39682) ≈ 0.121

Q. Which of the following statement is correct?
1. If autoregressive parameter (p) in an ARIMA model is 1, it means that there is no auto-correlation in the series.
2. If moving average component (q) in an ARIMA model is 1, it means that there is auto-correlation in the series with lag 1.
3. If integrated component (d) in an ARIMA model is 0, it means that the series is not stationary.

A) Only 1B) Both 1 and 2C) Only 2D) All of the statements

Solution: (C)
Autoregressive component: AR stands for autoregressive. Autoregressive parameter is denoted by p. When p =0, it means that there is no auto-correlation in the series. When p=1, it means that the series auto-correlation is till one lag.

Integrated: In ARIMA time series analysis, integrated is denoted by d. Integration is the inverse of differencing. When d=0, it means the series is stationary and we do not need to take the difference of it. When d=1, it means that the series is not stationary and to make it stationary, we need to take the first difference. When d=2, it means that the series has been differenced twice. Usually, more than two time difference is not reliable.

Moving average component: MA stands for moving the average, which is denoted by q. In ARIMA, moving average q=1 means that it is an error term and there is auto-correlation with one lag.

Q. In a time-series forecasting problem, if the seasonal indices for quarters 1, 2, and 3 are 0.80, 0.90, and 0.95 respectively, what can you say about the seasonal index of quarter 4?

A) It will be less than 1

B) It will be greater than 1

C) It will be equal to 1

D) Seasonality does not exist

E) Data is insufficient

Solution: (B)