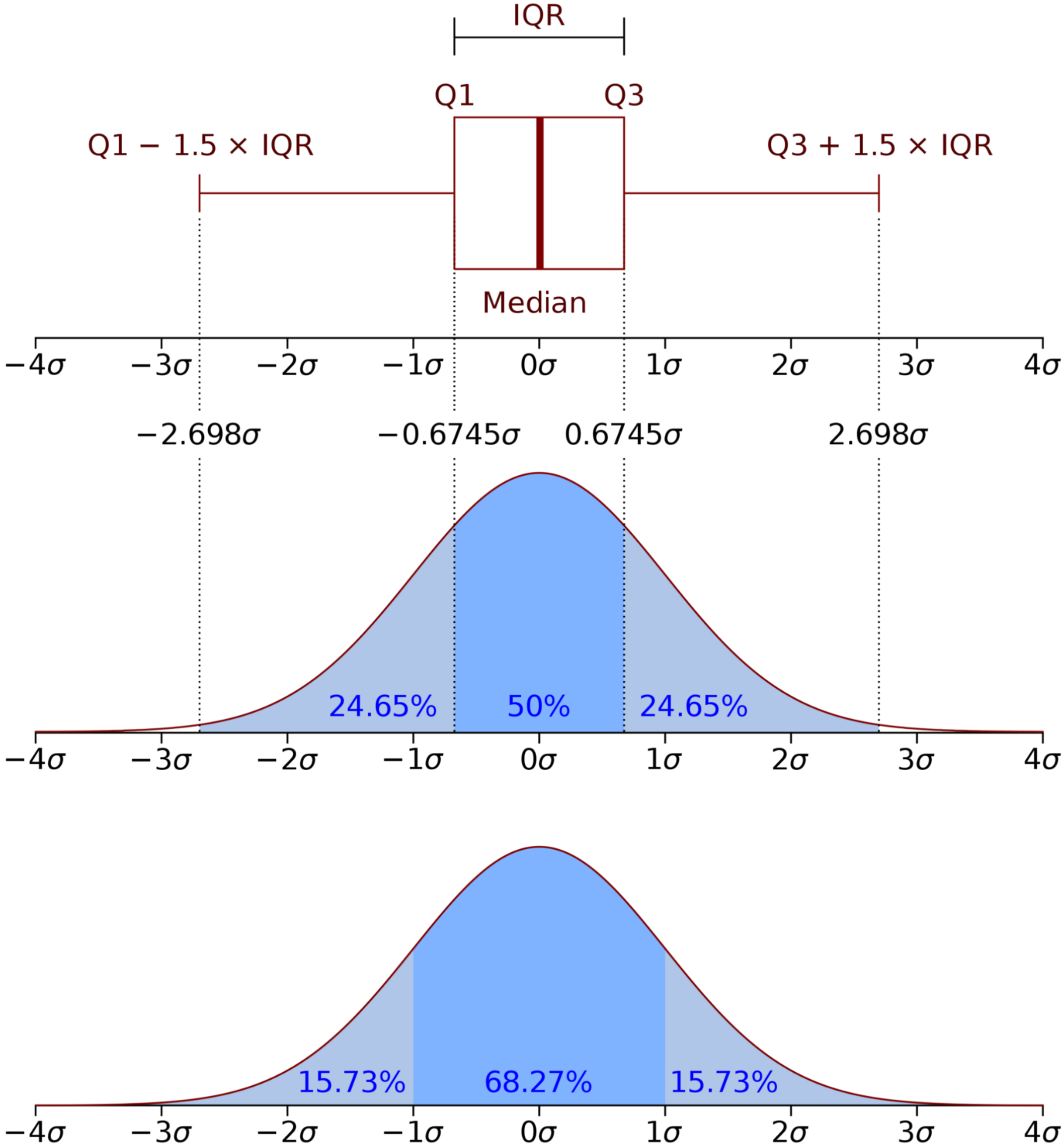
How the *IQR (Interquartile Range)* is used in Time Series Forecasting?
Basically, IQR is *the range between 1st and 3rd quartile*.



We can use IQR to detect outliers in Time Series data**. An "outlier" is a point that is which lies beyond Q1- 1.5 times the Inter Quartile Range (IQR) and Q3+1.5IQR**
What are some common *Data Preparation Operations* you would use for Time Series Data?

Junior

Time Series 38

Answer

Some *data preparation operations* which can be used are:

- *Parsing time series* information from various sources and formats.
- Generating sequences of fixed-frequency dates and time spans.
- Manipulating and converting date times with *time zone* information.
- *Resampling* or *converting* a time series to a particular frequency.
- Performing date and time arithmetic with absolute or relative time increments.

What are some examples of *Time-Series Data* which can be *Mined*?

Some examples of such data are as follows:

- **Sensor data:** Sensor data is often collected by a wide variety of hardware and other monitoring devices. Typically, this data contains continuous readings about the underlying data objects. For example, environmental data is commonly collected with different kinds of sensors that measure temperature, pressure, humidity, and so on. Sensor data is the most common form of time series data.
- **Medical devices:** Many medical devices such as an electrocardiogram (ECG) and electroencephalogram (EEG) produce continuous streams of time series data. These represent measurements of the functioning of the human body, such as the heartbeat, pulse rate, blood pressure, etc. Real-time data is also collected from patients in intensive care units (ICU) to monitor their condition.
- **Financial market data:** Financial data, such as stock prices, is often temporal. Other forms of temporal data include commodity prices, industrial trends, and economic indicators.

What are some real-world applications of Time-Series Forecasting?

- **Time-Series in Financial and Business Domain:** Time series analysis and forecasting essential processes for explaining the dynamic and influential behavior of financial markets. Via examining financial data, an expert can predict required forecasts for important financial applications in several areas such **as risk evolution, option pricing & trading, portfolio construction, etc.**
- **Time-Series in Medical Domain:** Medical instruments like *electrocardiograms*, and *electroencephalograms* are used to diagnose cardiac conditions, and measure electrical activity in the brain, respectively.

These inventions made more opportunities for medical practitioners to deploy time series for medical diagnosis.

- **Time-Series in Astronomy:** Being specific in its domain, astronomy hugely relies on plotting objects, trajectories, and accurate measurements, and due to the same, astronomical experts are proficient in time series in calibrating instruments and studying objects of their interest. In the past century, time series analysis was used to discover variable stars that are used to surmise stellar distances, and observe transitory events such as supernovae to understand the mechanism of the changing of the universe with time.

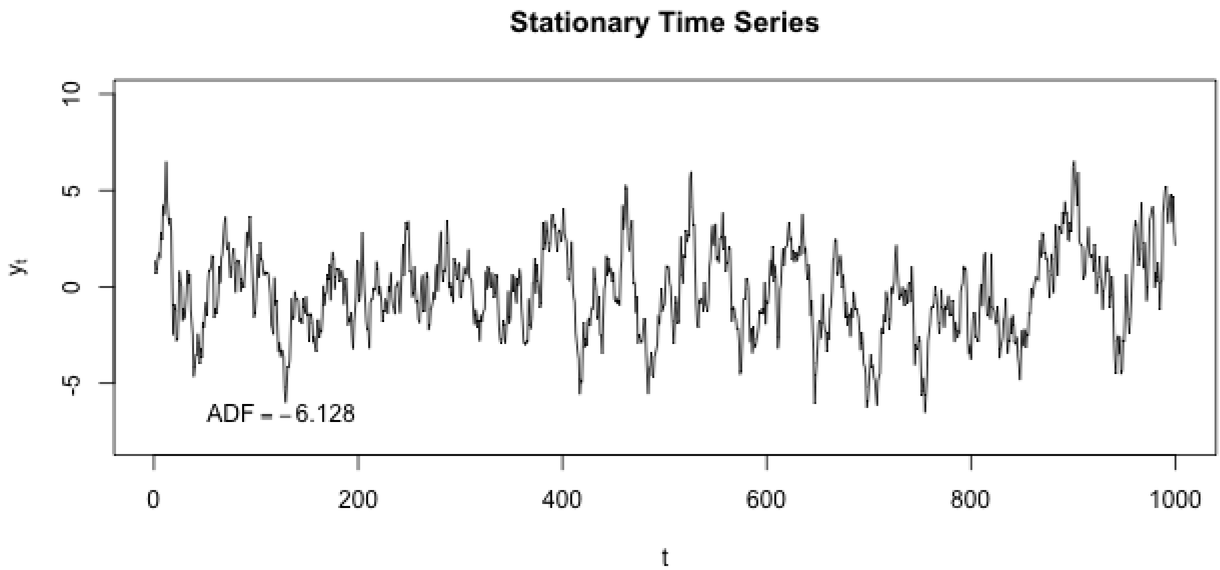
What is the *Sliding Window* method for Time Series Forecasting?

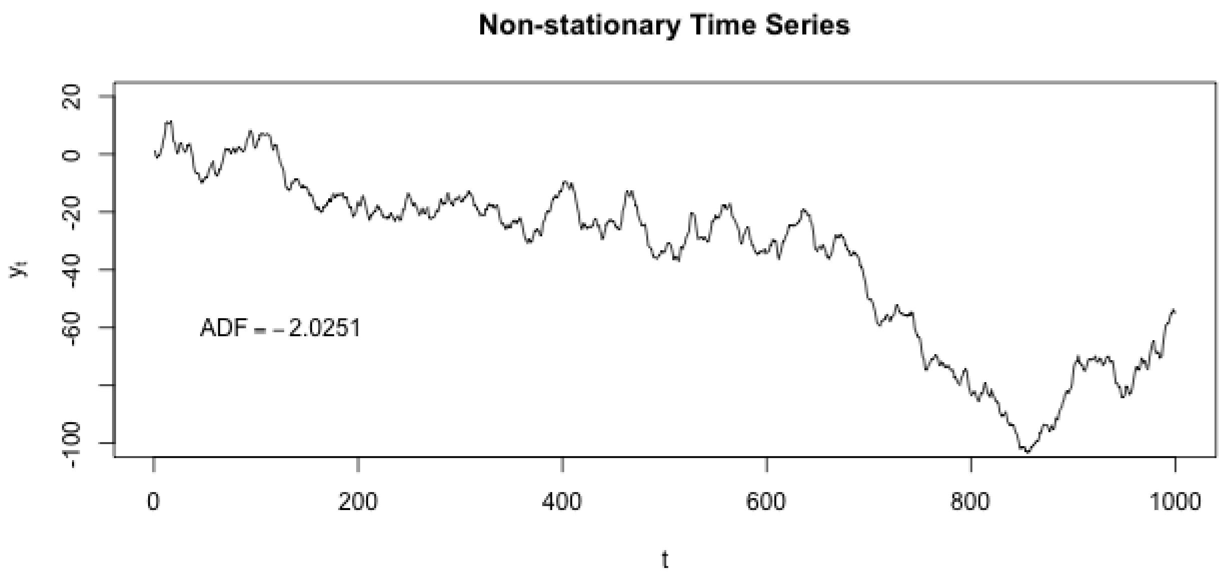
- *Time series* can be phrased as supervised learning. Given a sequence of numbers for a time series dataset, we can restructure the data to look like a supervised learning problem.
- In the **sliding window** method, the previous time steps can be used as input variables, and the next time steps can be used as the output variable.
- In statistics and time series analysis, this is called a **lag** or **lag method**. The number of previous time steps is called the *window width* or *size of the lag*. This sliding window is the basis for how we can turn any time series dataset into a supervised learning problem.
- A pictorial representation of the *sliding window* is shown in the figure below:

Why does a Time Series have to be *Stationary*?

Stationarity is important because, in its absence, *a model describing the data will vary in accuracy at different time points*. As such, **stationarity is required for sample statistics such as means, variances, and correlations to accurately describe the data at all time points of interest**.

Looking at the time series plots below, you can (hopefully) see how the mean and variance of any given segment of time would do a good job representing the whole stationary time series but a relatively poor job representing the whole non-stationary time series. For instance, the mean of the non-stationary time series is much lower from 600<t<800 ,600<t<800 and its variance is much higher in this range than in the range from 200<t<400 200<t<400.





What quantities are we typically interested in when we perform statistical analysis on a time series? We want to know

- Its expected value,
- Its variance, and
- The correlation between values $s$ss periods apart for a set of $s$ss values.

How do we calculate these things? Using a mean across many time periods.

The mean across many time periods is only informative if the expected value is the same across those time periods. If these population parameters can vary, what are we really estimating by taking an average across time?

(Weak) stationarity requires that these population quantities must be the same across time, making the sample average a reasonable way to estimate them.

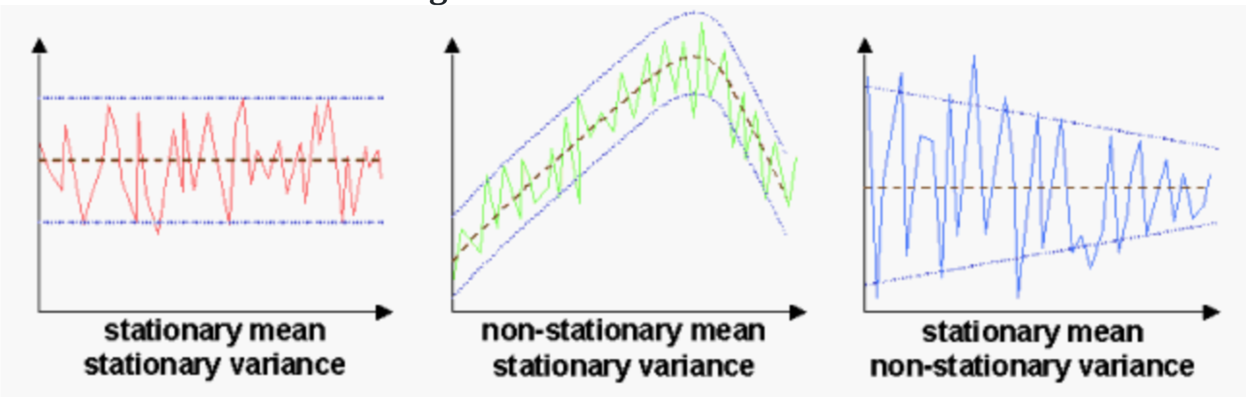
Can *Non-Sequential Deep Learning Models* outperform *Sequential Models* in Time-Series Forecasting?

- Yes, *non-sequential* deep learning models like *CNN* can outperform *sequential models* like *RNN*.
- Even though there are no big differences in terms of results there are some nice properties that *CNN* based models offer such as, ***parallelism*, *stable gradients*, and *low training memory footprint*.**
- So depending on what type of properties you want your model to have, architectures can be chosen accordingly.

What is weak stationary,non stationary data

- **Weak stationarity**
- *Time-series data* which exhibit **weak stationarity** properties, will have the *mean* and *covariance* between approximately adjacent time series values be non-zero and constant over time. This is referred to as *covariance stationarity*.
- This kind of weak stationarity can be assessed relatively easily and is also useful for forecasting models that are dependent on specific parameters such as the *mean* and *covariance*.

  Non stationarity:
- In other **nonstationary series**, the average value of the series can be described by a *trend-line* that is not necessarily horizontal, as required by a stationary series.
- Periodically, the series will deviate from the trend line, possibly because of some changes in the generative process, and then return to the trend line. **This is referred to as a *trend stationary series*.**
- Nonstationary data can be converted to stationary series for forecasting. One method is called **differencing**.



How do you *Normalize* Time-Series Data?

Two normalization methods that are commonly used are:

- **Range-based Normalization:** In range-based normalization, the minimum and maximum values of the time series are determined. Let these values be denoted by
`min` and `max`, respectively. Then, the time series value y_i is mapped to the new value y_i' in the range `(0,1)` as follows:

$$y_i' = \frac{y_i - min}{max - min}$$

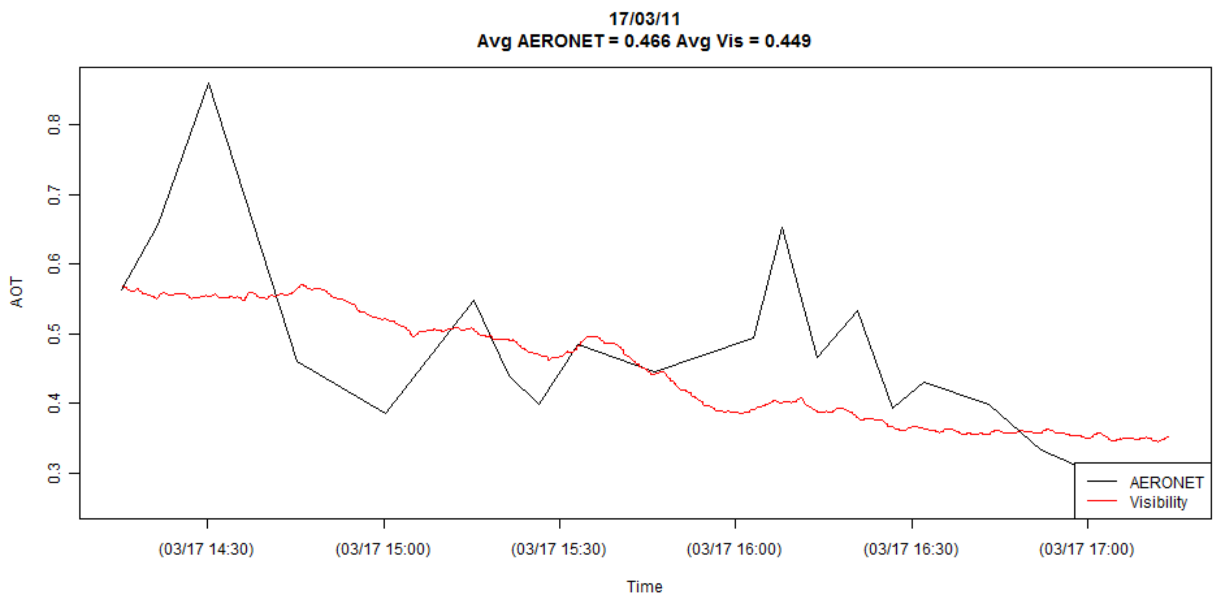
- **Standardization:** In standardization, the mean and standard deviation of the series are used for normalization. This is essentially the Z-value of the time series. Let \mu$\mu$ and \sigma$\sigma$ represent the mean and standard deviation of the values in the time series. Then, the time series value y_iyi is mapped to a new value z_izi as follows:

$$z_i = \frac{y_i - \mu}{\sigma}$$

How is *Pearson Correlation* used with Time Series?

- **Pearson correlation** is used to look at the correlation between series, but being time series, the correlation is looked at across different lags - the *cross-correlation function*.

How would you compare the two Time Series shown below?



- The plot appears to obscure what may be a crucial difference between these series: they might be sampled at different frequencies. The black line (Aeronet) seems to be sampled only about 20 times and the red line (Visibility) hundreds of times or more.

If your Time-Series Dataset is very long, what *architecture* would you use?

- f the dataset for *time-series* is very long, **LSTMs** are ideal for it because it can not only process *single data points*, but also entire *sequences of data*. A time-series being a sequence of data makes LSTM ideal for it.
- For an even stronger representational capacity, making the LSTM's multi-layered is better.
- Another method for long time-series dataset is to use **CNNs** to extract information.
- **1) 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.
- **2) 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, **Exponential smoothing** of time series data assigns exponentially decreasing weights for newest to oldest observations. In other words, the older the data, the less priority ("weight") the data is given; newer data is seen as more relevant and is assigned more weight. Smoothing parameters (smoothing constants)— usually denoted by α— determine the weights for observations. Exponential smoothing is usually used to make short term forecasts, as longer term forecasts using this technique can be quite unreliable.

- What is moving averages in time series?

- In time series analysis, Moving averages have the property to reduce the amount of variation present in the data. In the case of time series, this property is used **to eliminate fluctuations**, and the process is called smoothing of time series.Moving averages alone aren't that useful for forecasting. Instead, they are mainly used for analysis. For example, moving averages help stock investors in technical analysis by smoothing out the volatility of constantly changing prices. This way, short-term fluctuations are reduced, and investors can get a more generalized picture.

Which of the following are the component for a time series plot?

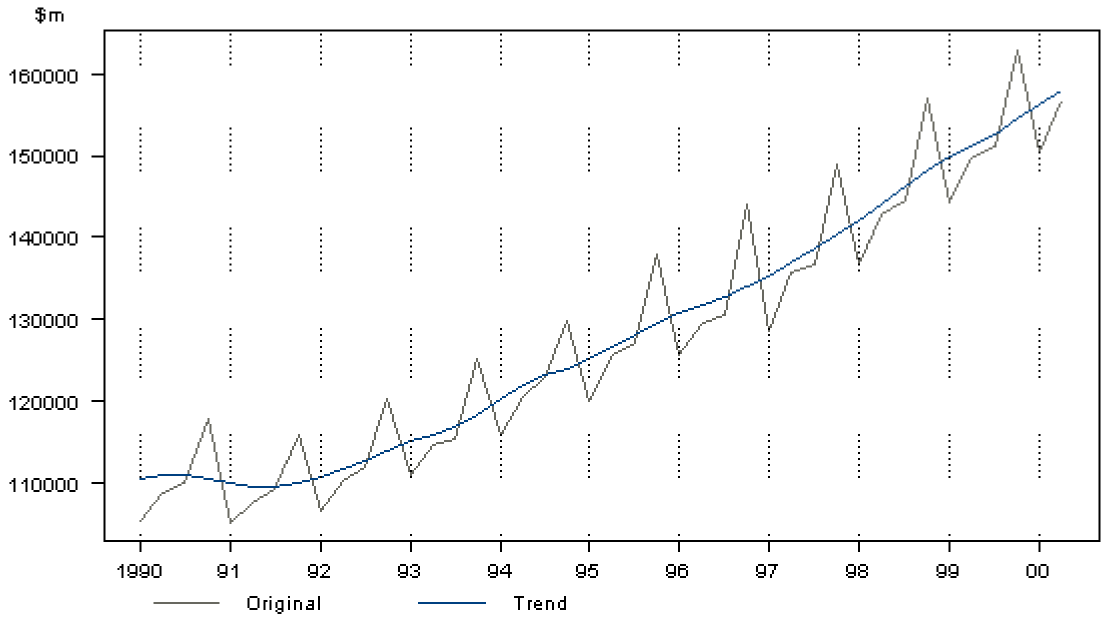
A) Seasonality
B) Trend
C) Cyclical
D) Noise
E) All of the above

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.

4) Which of the following is relatively easier to estimate in time series modeling?

A) Seasonality

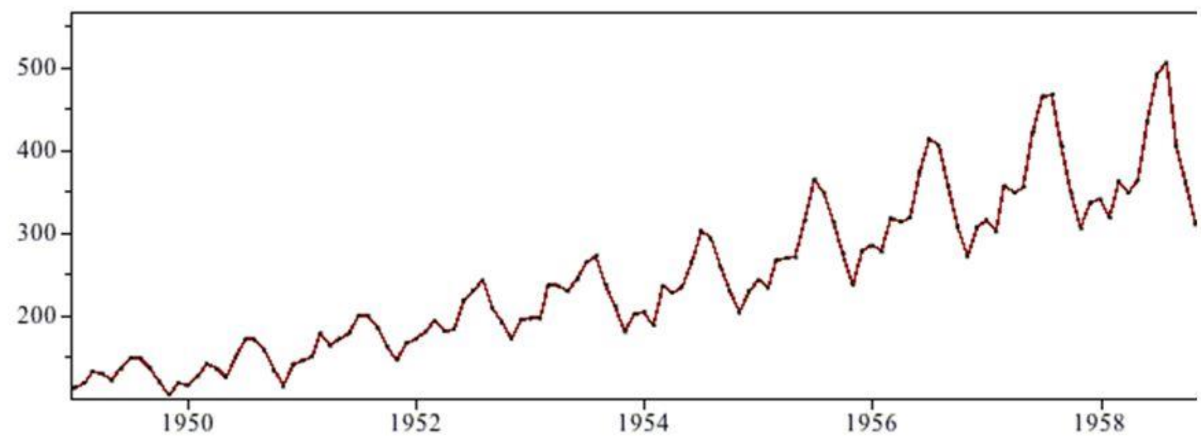
B) Cyclical

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.

5) 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.

6) 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.

7) 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:
$\hat{y}_{T+1|T}=\alpha y_T+\alpha(1-\alpha)y_T-1+\alpha(1-\alpha)2y_{T-2}+\cdots,(7.1)$

where 0≤α≤10≤α≤1 is the smoothing parameter. The one-step-ahead forecast for time T+1T+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 αα.

8) Sum of weights in exponential smoothing is _____.
A)<1
B) 1
C) >1
D) None of the above
Solution: **(B)**
Table 7.1 shows the weights attached to observations for four different values of αα when forecasting using simple exponential smoothing. Note that the sum of the weights even for a small αα will be approximately one for any reasonable sample size.

Observation	α=0.2	α=0.4	α=0.6	α=0.8
yT	0.2	0.4	0.6	0.8
yT−1	0.16	0.24	0.24	0.16
yT−2	0.128	0.144	0.096	0.032
yT−3	0.102	0.0864	0.0384	0.0064
yT−4	(0.2)(0.8)	(0.4)(0.6)	(0.6)(0.4)	(0.8)(0.2)
yT−5	(0.2)(0.8)	(0.4)(0.6)	(0.6)(0.4)	(0.8)(0.2)

9) 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)**
Y_{t-1}= 70
S_{t-1}= 60
Alpha = 0.4
Substituting the values we get
0.4*60 + 0.6*70= 24 + 42= 66

10) What does autocovariance 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 autocovariance.

11) 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.

What is smoothing?

Smoothing helps us to identify trend, seasonality and level.It smooth out the irregular roughness to see a clearer signal.

13) 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?

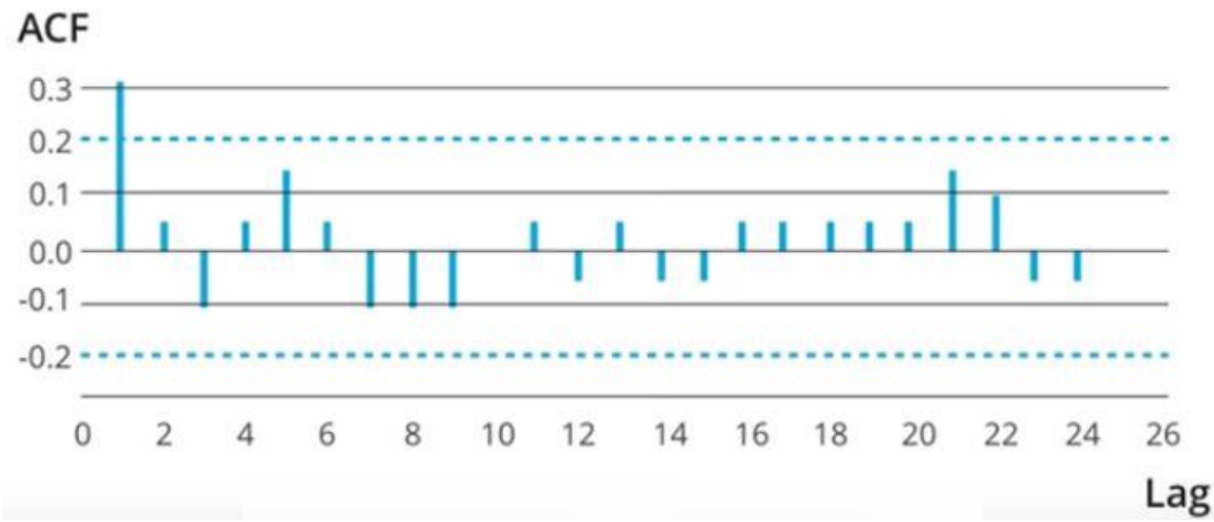
A) 300
B)350
C)400
D) Need more information

Solution: **(A)**

X\`= (x_{t-3} + x_{t-2} + x_{t-1}) /3

(200+300+400)/ 3 = 900/3 =300

14) 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 "over differenced"–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 preffered.

15) Suppose, you are a data scientist at Analytics Vidhya. And you observed the views on the articles 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.

16) Which of the following graph can be used to detect seasonality in time series data?
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.

17) 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.

1. Mean is constant and does not depend on time
2. Autocovariance function depends on s and t only through their difference |s-t| (where t and s are moments in time)
3. 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.

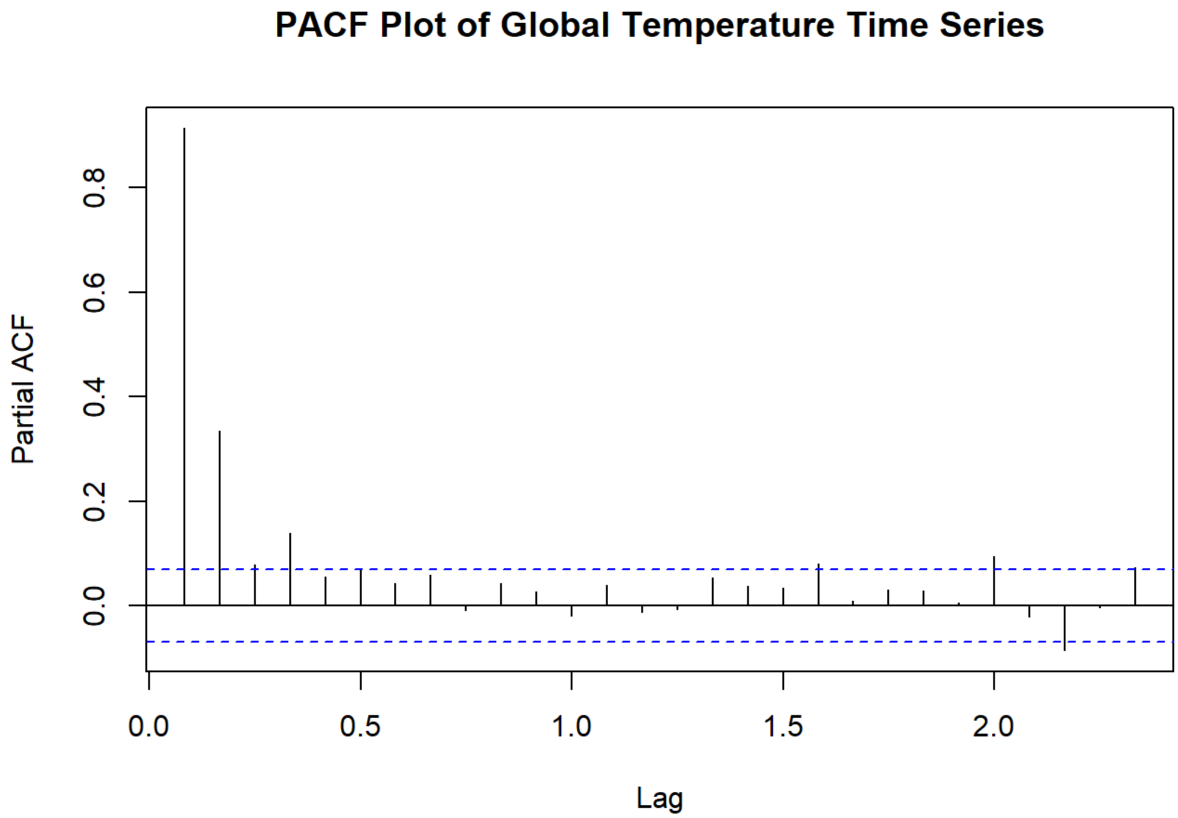
19) Imagine, you are working on a time series dataset. Your manager has asked you to build a highly accurate model. You started to build two types of models which are given below.
Model 1: Decision Tree model
Model 2: Time series regression model
At the end of evaluation of these two models, you found 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.

20) 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)TimeSeriesAnalysis
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.

What are Autoregressive Model: AR(p)
The autoregressive model uses observations from previous time steps as input to a regression equations to predict the value at the next step. The AR model takes in one argument, p, which determines how many previous time steps will be inputted.
The order, p, of the autoregressive model can be deterimined by looking at the **partial autocorrelation function (PACF)**. The PACF gives the partial correlation of a stationary time series with its own lagged values, regressed of the time series at all shorter lags.



What should we look for in this plot? The primary goal is to look for the number of significant spikes outside of the blue confidence intervals. In this plot, I would determine there to be 2 spikes, one at 0.1 and the other at 0.3. The spike at the 0 does not count and any spikes outside of the blue later in the plot are likely due to random error. Therefore, this looks like an AR(2) model.

Let's look at an AR(2) model for the global temp time series. You can use the *ar()* function in R; however, I recommend using the *Arima()* function from the **forecast** package

Moving Average Model: MA(q)

The moving average model is a time series model that accounts for very short-run autocorrelation. **It basically states that the next observation is the mean of every past observation.** The order of the moving average model, q, can usually be estimated by looking at the ACF plot of the time series. Let's take a look at the ACF plot again.

Autoregressive Moving Average Model: ARMA(p,q)

Autoregressive moving average models are simply a combination of an AR model and an MA model. Let's take a look at what our ARMA model would be.

Autoregressive Integrated Moving Average Model: ARIMA(p,d,q)

This model is the same as the previous, except now it has this weird d argument. What does this d stand for? **d represents the number of nonseasonal differences needed for stationarity**. Simply, **purpose of d is to make nonstationary data to stationary by removing trends!**

How do you pick your differencing term?

Usually, small terms are picked for the differencing term. If you pick too high, you will likely cause your model to incorrectly represent your data. Some general rules for picking your differencing term are that differencing should not increase your variance and the autocorrelation of the model should be less than -0.5.

(The value of d must be small and range between 0 to 2)

Seasonal Autoregressive Integrated Moving Average Model: SARIMA(p,d,q)(P,D,Q)s

We know that ARIMA makes time series stationary by using d which removes trends from time series to make it stationary.SARIMA is going to remove both trend and seasonality.

The SARIMA model is an extension of the ARIMA model. The only difference now is that this model added on a seasonal component. As we saw, ARIMA is good for making a non-stationary time series stationary by adjusting the trend. However, the SARIMA model can adjust a non-stationary time series by removing trend and seasonality.

As we know:

- p - the order of the autoregressive trend
- d - the order of the trend differencing
- q - the order of the moving average trend

What do (P,D,Q)s mean?

- P - the order of the autoregressive seasonality
- D - the order of the seasonal differncing
- Q - the order of the moving average seasonality
- s - the number of periods in your season

How do you pick these new terms?

There are several ways to pick these orders; however, when trying to use the SARIMA model in practice, it is likely best to let R or other software estimate the parameters for you. This article attached **here** mentions this in more detail.

In our example, we may not have a SARIMA model because our time series did not have seasonality. Therefore, it may follow a SARIMA(2,1,5)(0,0,0)12.

The s term would be 12 because there would be 12 periods (months) in the season if we had seasonality. We will still follow through with an example. We can use the sarima function from the **astsa** package in R.