1. Exponential Smoothing?

Ans: Older data is given progressively-less relative importance whereas newer data is given progressively greater importance in TS called Exponential Smoothing.

2. Explain Exponential Smoothing?

Ans:

Exponential Smoothing

Week	Sales (A_t)	Forecast (F_t)
1	39	
2	44	
3	40	
4	45	
5	38	
6	43	
7	39	

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

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

Exponential Smoothing $\alpha = 0.2$

Week	Sales (A_t)	Forecast (F_t)
1	39	
2	44	39.00
3	40	40.00
4	45	40.00
5	38	41.00
6	43	40.40
7	$0.2 \times 39 + 0.8 \times 40.92$	
8		

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

$$F_{t+1} = 0.2(A_t) + 0.8(F_t)$$

$$F_2 = A_1$$

$$F_3 = 0.2(44) + 0.8(39.00)$$

$$F_4 = 0.2(40) + 0.8(40.00)$$

$$F_5 = 0.2(45) + 0.8(40.00)$$

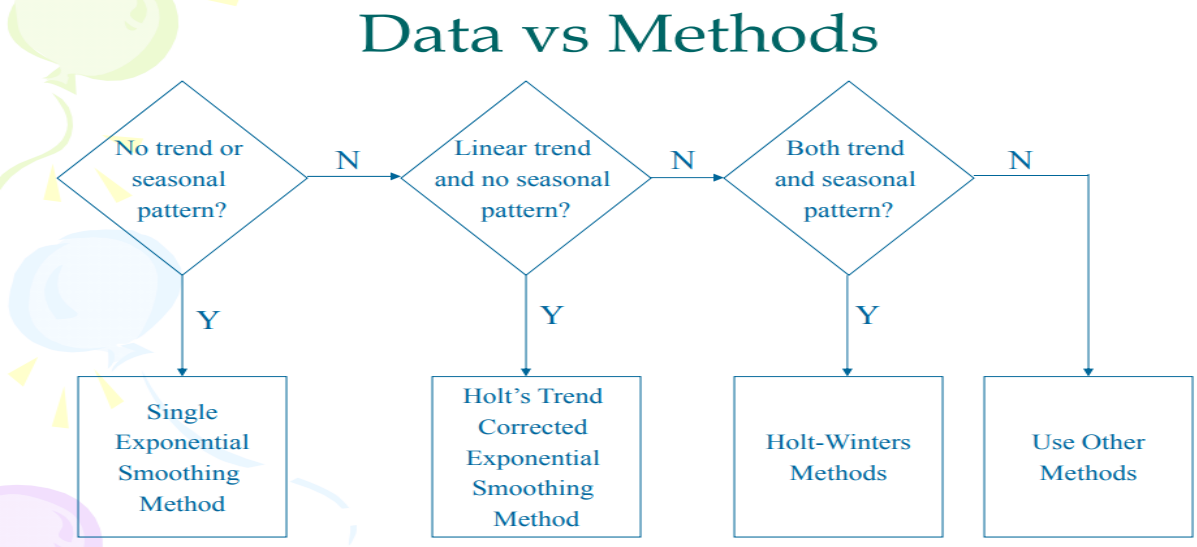
$$F_6 = 0.2(38) + 0.8(41.00)$$

$$F_7 = 0.2(43) + 0.8(40.40)$$

$$F_8 = 0.2(39) + 0.8(40.92)$$

3. When we should use Exponential Smoothing and other models?

Ans: No trend or seasonal pattern then use Exponential Smoothing.



4. What is Time Series?

Ans: A time series is a sequence of observations taken sequentially in time.
Univariate time series: Only one variable is varying over time. For example, data collected from a sensor measuring the temperature of a room every second. Therefore, each second, you will only have a one-dimensional value, which is the temperature.
Multivariate time series: Multiple variables are varying over time. For example, a tri-axial accelerometer. There are three accelerations, one for each axis (x, y, and z) and they vary simultaneously over time.

5. Components of Time Series?

Ans: Trend, Seasonality, Cyclical, Irregularity

Trend component

The trend is the long term pattern of a time series. A trend can be positive or negative depending on whether the time series exhibits an increasing long term pattern or a decreasing long term pattern.

If a time series does not show an increasing or decreasing pattern then the series is stationary in the mean.

Cyclical component

Any pattern showing an up and down movement around a given trend is identified as a cyclical pattern. The duration of a cycle depends on the type of business or industry being analyzed.

Seasonal component

Seasonality occurs when the time series exhibits regular fluctuations during the same month (or months) every year, or during the same quarter every year. For instance, retail sales peak during the month of December.

Irregular component

This component is unpredictable. Every time series has some unpredictable component that makes it a random variable. In prediction, the objective is to "model" all the components to the point that the only component that remains unexplained is the random component.

6. Types of Trend in Time Series?

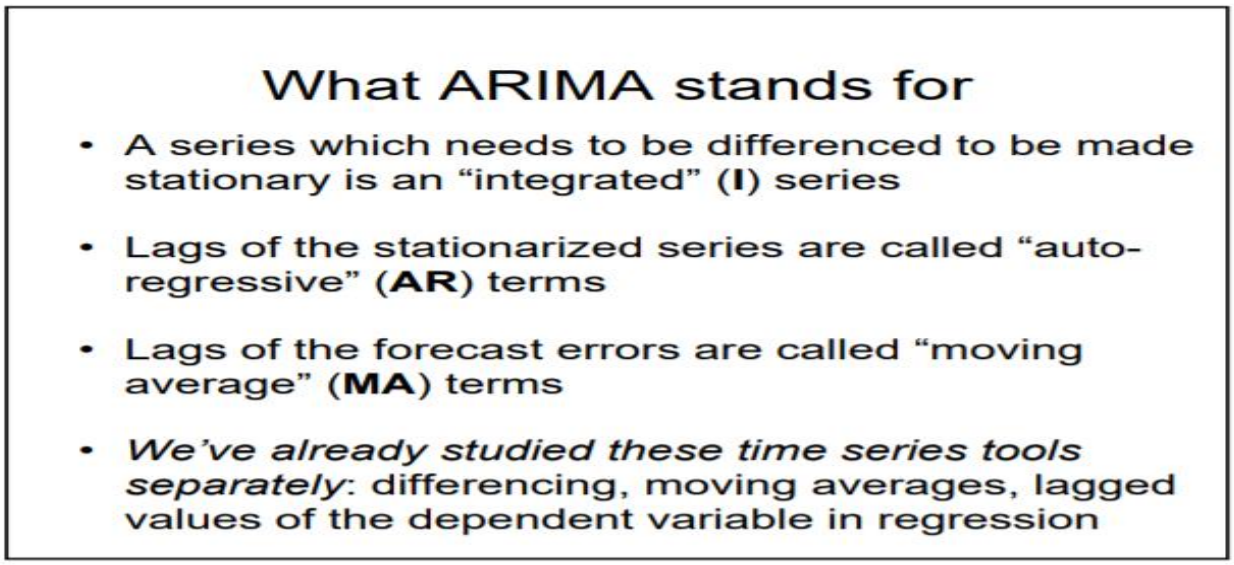
Ans: Deterministic and Stochastic

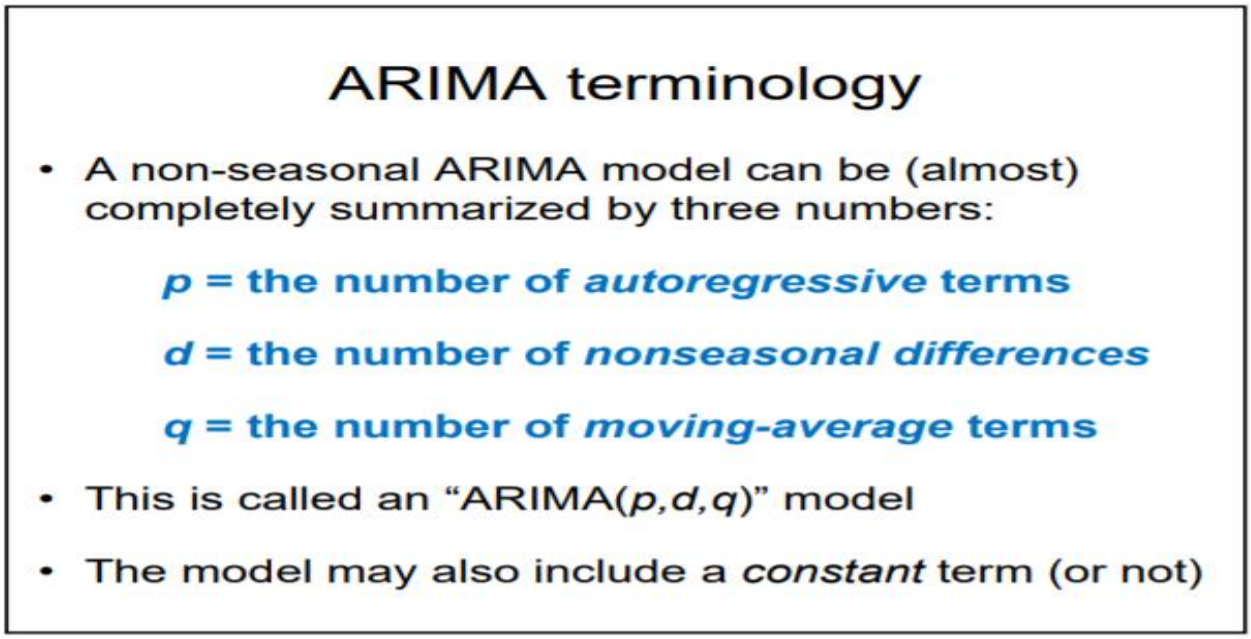
7. Time Series Forecasting Algorithms?

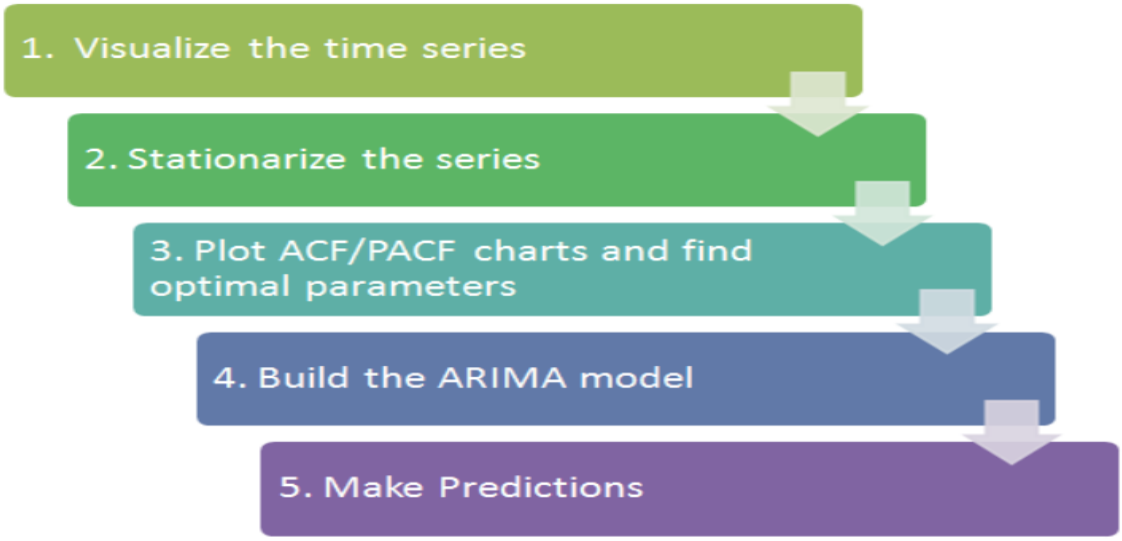
Ans: ARIMA, Holt Winters (Triple Exponential Smoothing) are two main algorithms for demand forecasting in TS.

8. Explain Arima Model?

Ans:







9. Holt Winters vs Arima Model?

Ans: Holt Winters Model is easy to use and easy to repeat for several periods.
Arima uses Auto Correlation, which explains a lot of valuable information about data.
It is complex, hard to repeat and requires many data.

Holt Winters model is good for demand forecasting because market is open to change.
Easy and cheap, to implement and cost factor is not too high.

10. Why do we need to make our data stationary only in ARIMA?

Ans: We make the data stationary only is case of Arima because the Arima model looks at the past data to predict the future values. (Naive, simple exponential etc. do not work that way).

<<<Stationery data means you mean and standard deviation of data does not change in time. In trends, your mean and standard deviation is dependent on past value so you need to do the differencing of it to remove the dependency of the data.>>>

11. What are the different ways to make our data stationary?

Ans: The most common method would be **Differencing** (one level or seasonal differencing).
You can also perform various **transformations** (square root, log, box cox).

12. What method can be used on Non Stationary Data?

Ans: exponential smoothing tends to be more robust with non-stationary time series.

13. Test for Stationarity in ARIMA model?

Ans: Augmented Dickey-Fuller (ADF) test

14. Explain Autocorrelation, ACF and PACF in Time series Analysis?

Ans:

The correlation of the time series observations is calculated with values of the same series at previous times, this is called a serial correlation, or an autocorrelation.

Since autocorrelation is the linear correlation of a signal with itself at two different points in time, ACF (autocorrelation function) is just such correlation as a function of the lag h between two points of time, like $acf(h) = corr(x_t, x_{x+h})$

PACF (partial autocorrelation function) is essentially the autocorrelation of a signal with itself at different points in time, *with linear dependency with that signal at shorter lags removed*, as a function of lag between points of time. Informally, the partial correlation between x_t and x_{t+h} is the autocorrelation between x_t and x_{t+h} without the contribution of $x_{t+1}, x_{t+2}, \ldots, x_{t+h-1}$.

15. Practice demand-forecasting challenge.

Ans: https://www.kaggle.com/c/demand-forecasting-kernels-only/

Problem Statement: You are given 5 years of store-item sales data, and asked to predict 3 months of sales for 50 different items at 10 different stores.

16. Error metrics for time series?

Ans: Root Mean Square Error (RMSE), Mean Square Error (MSE), Percentage Error (MPE), and Mean Absolute Percentage Error (MAPE)

17. Multivariate Time Series Model?

Ans: Random Forest, VAR, LSTM

18. Time Series vs Regression?

Ans: The biggest difference is that time series regression accounts for the autocorrelation between time events, which always exists, while in normal regression, independence of serial errors are presumed, or at least minimized.

19. Does Time parameter indicate the problem is time series problem?

Ans: Not Necessary.

Dataset with house prices having multiple features of the house and the year of construction. Even though the year of construction will give some correlation towards house price, but other features like the constructed area will make it even better.

Just because you have time feature, it does not mean to be a time series problem.

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)**