# Week 4: Introduction to Time series modeling

Predictive modeling!

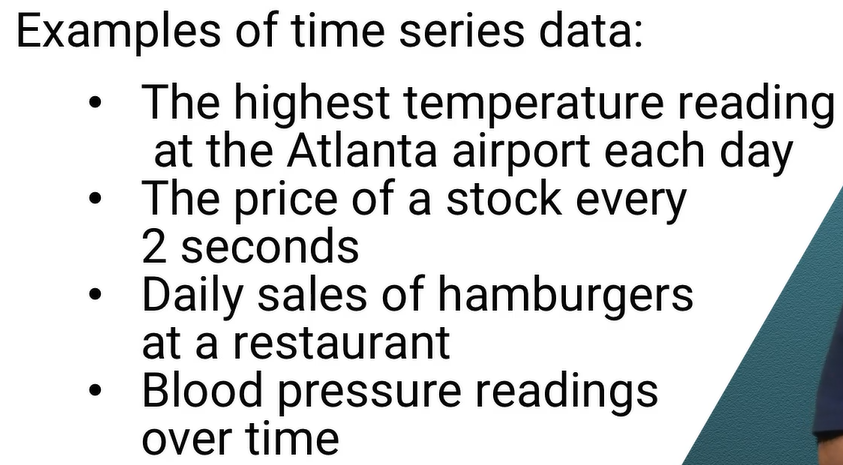
Time series models for smoothing,

short-term forecasting, and variance prediction.

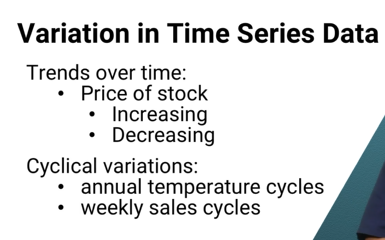
# 4.1 Introduction to exponential Smoothing

- for time series data

- What is time series data : same response for many time periods. Usually same intervals(may be not as well)



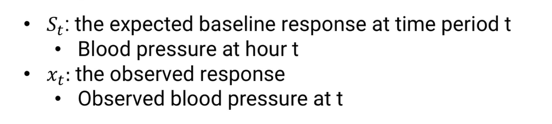
## Variations



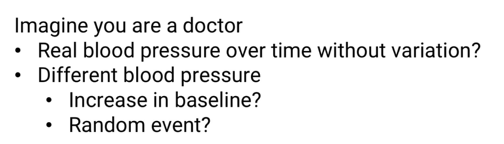
**Other variations: Random**

* **Challenge:** is an action required or not?
* Here is where an exponential smoothing comes to play (to get a baseline of the trend without the diversion of the peaks and valleys)..

Example:



**Scenario:** we are thedoctor and want to see BP over time without variation



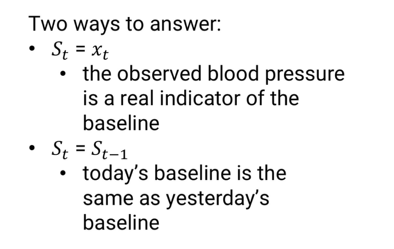
**Two ways to answer this questions**

1. Assume observed BP is a real indicator of baseline

S(t) =X(t)

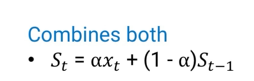
1. Assume no change to baseline and observed reading is just random

S(t) =s(t-1)

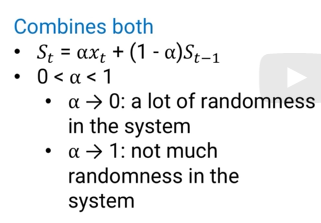


**Exponential smoothing:**

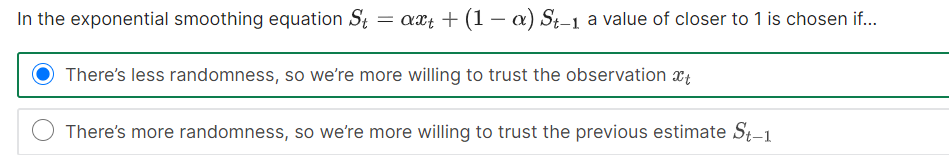
Combines the above two ideas



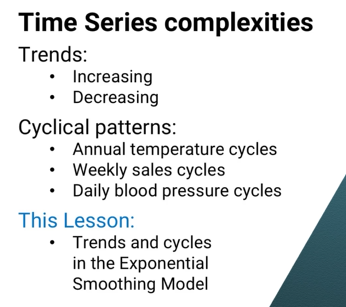
“**[α” = value between 0 and 1](https://www.greeksymbols.net/alpha-symbol" \l ":~:text=Alpha%20(uppercase%20%CE%91%3B%20lowercase%20%CE%B1,derived%20from%20the%20Phoenician%20alphabet.)**



NOTE : This doesn’t include trends and cyclical(Seasonality).usually, we start with S(1) =x(1)

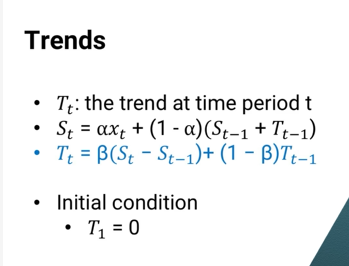


# 4.2 Trends & Cyclic Effects



Trend = expected baseline for time “t” – expected baseline for (t-1)

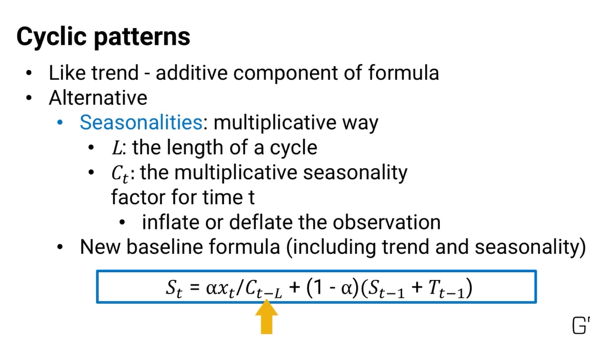
= s(t) – s(t-1)



**Cyclic pattern:**

Van be added to exponential smoothing as additive or multiplicative way

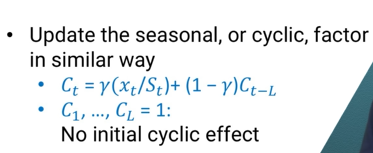
Here we will see multiplicative way

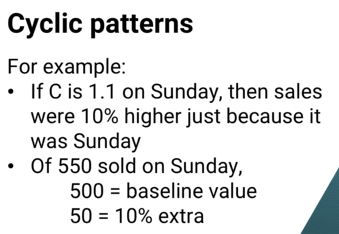


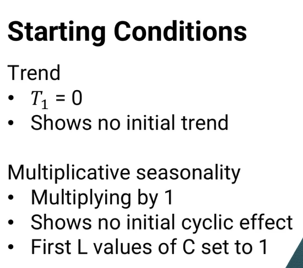
**Why we use c(t-L)?**

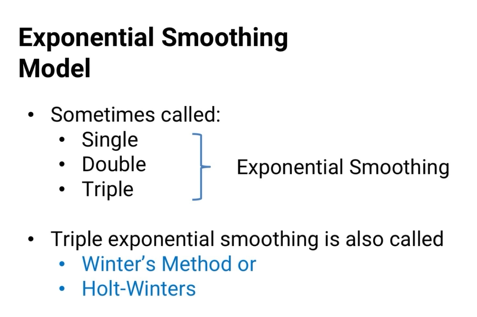
Because when we consider cyclic effect, we want to consider previous time period as t-L , reason L =7 when we consider weeks and L=24 when we consider every hour

So we compare data of a Monday to that of a previous Monday



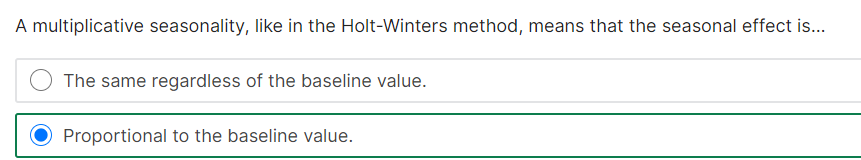






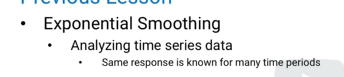
Triple exponential smoothing : exponential smoothing+trends+smoothing

A multiplicative seasonality is larger when the baseline value is larger, because its effect is a multiple of the baseline.

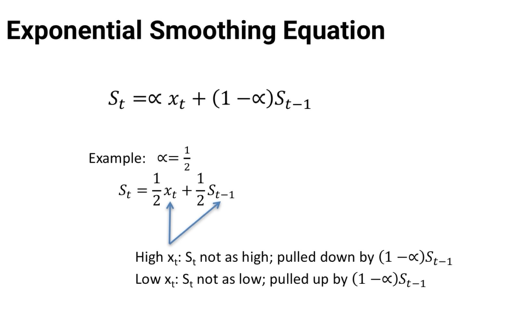


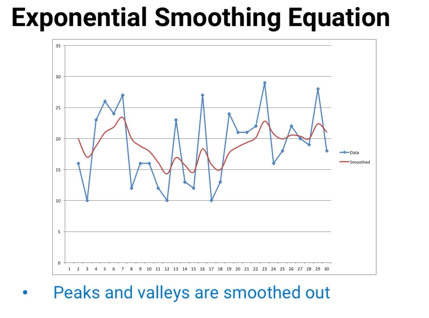
# 4.3 Exponential Smoothing: What the name means

Exponential smoothing can NOT only used for descriptive model, but also for forecasting/predictive model

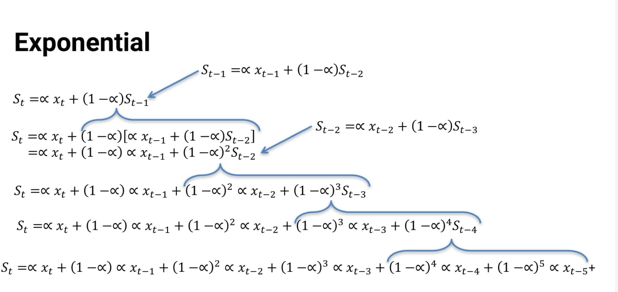


**SMOOTHING?**

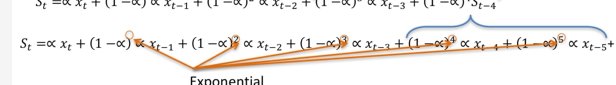




**Exponential:**

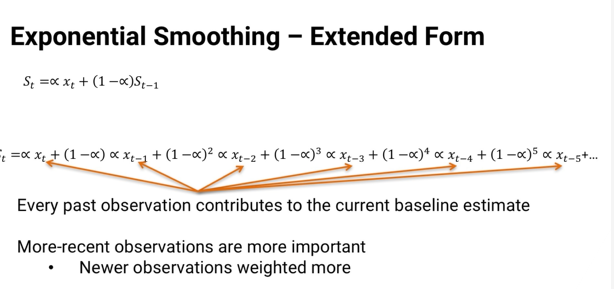
****

**If we observe, we can see each expansion will increase the (1-alpha) in exponential order.**

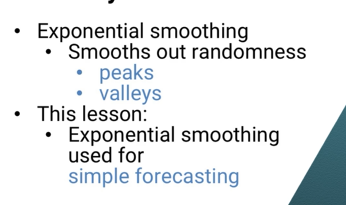
****

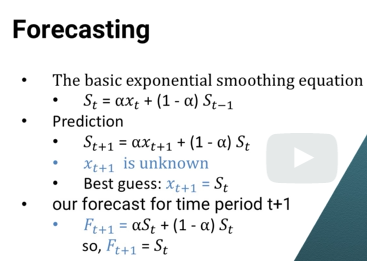
Which means, exponential will look for previous value to determine the current value (ALL OLDER ONES are considered with exponential smoothing)

however, the newer observation add more weightage than the old ones



# 4.5 Forecasting



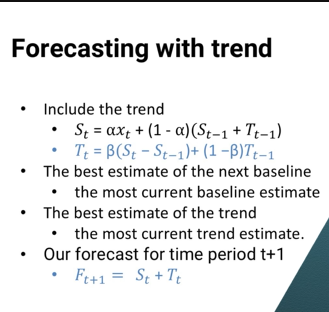


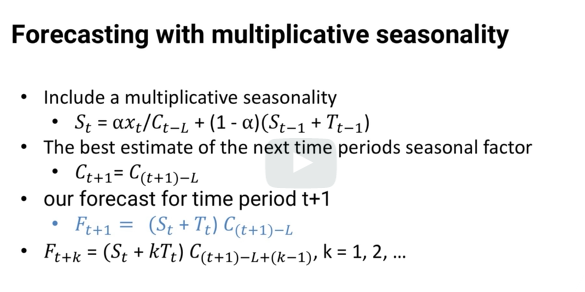
The future estimate of the time period is same of the current time period’ baseline value.

This is same for (t+1), (t+2), (t+3)….

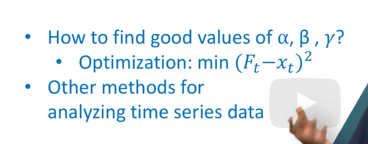
This is applicable for shorter time period forecasting

* Of course because there's more uncertainty the farther into the future we go,the anticipated forecast error gets larger, too.





How to find good values of alphz, gamma and beta?

* By optimization
* -for every forecast, we can take different of forecast and actual and take the mean square error
* **Objective** : minimize mean square error for whole training set.
* Whatever values of alpha, beta, and gamma minimize that sum of squared error is the model's best fit
* 

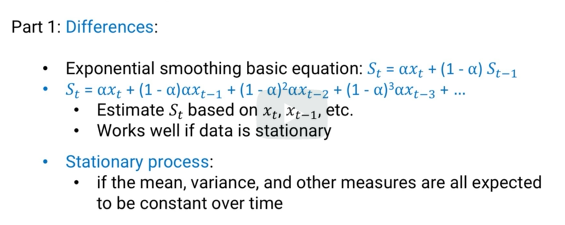
# 4.5 ARIMA

AutoRegressive Integrated Moving Average

-no theory

3 key parts

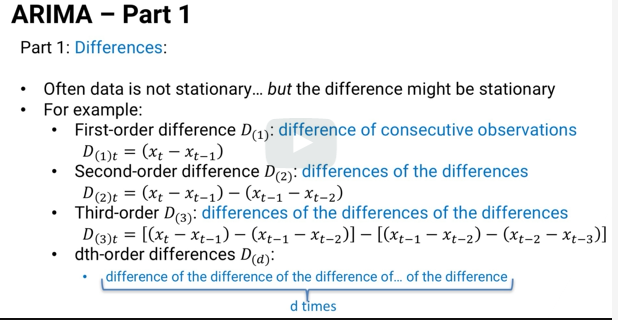
## 4.5.1 Differences



-exponential smoothing works better when data is stationary.

But data is not stationary. Trend or seasonality.

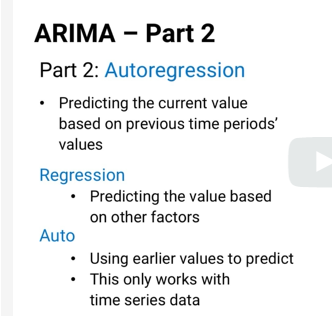
-sometimes even if data is not stationary, differences in data might be stationary



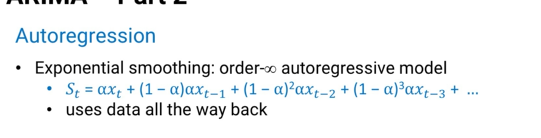
## 4.5.2 Auto regression

Regression -predict some value based on other factors

Auto Regression -predict some value based on SAME metrics (works only for time series)

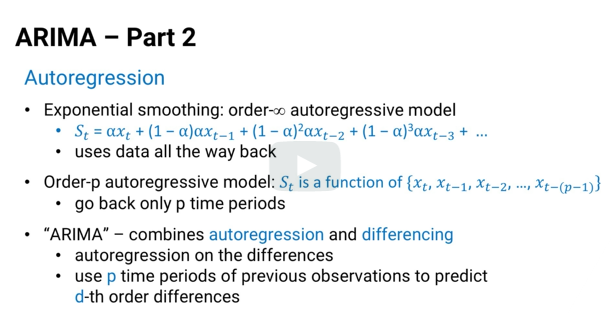


In forecasting, exponential smoothing is form of ARIMA Auto regression



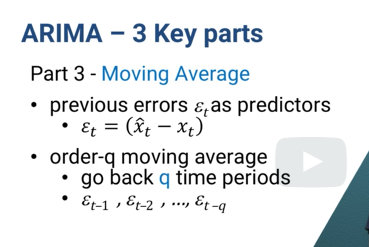
This is infinity order auto regression model.

If we use only certain data points, say “p”, it is called “order-p auto regression model”



### Instead of using autoregression with the actual observed data, , we use autoregresion on the differences.

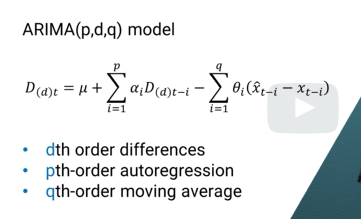
## 4.5.3 Moving Average



P => pth order auto regression (of difference)

D=> d-th order difference

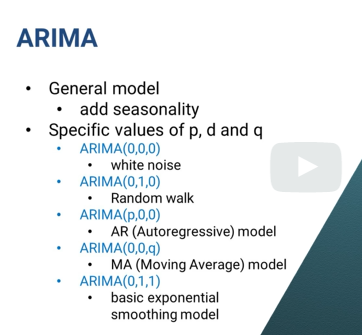
Q=> order-q moving average of errors



Statistical software can find best value for p,d,q

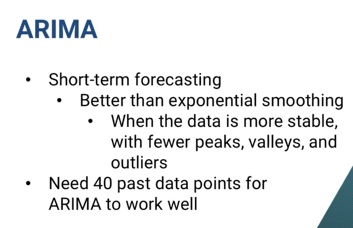
There are 3 main ways to predict the best fit of the data model

-we can get different values for p,d, and q and use validation techniques to find best value of p,d,q



-This generalizes a lot of simpler models

-  can be used for short-term forecasting,



They use to estimate or forecast a value

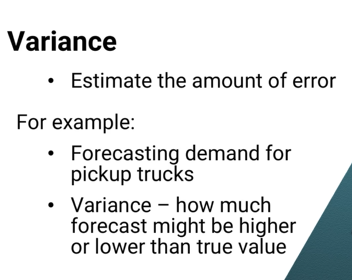
What if we want to estimate a variance?

# 4.6 G(ARCH)

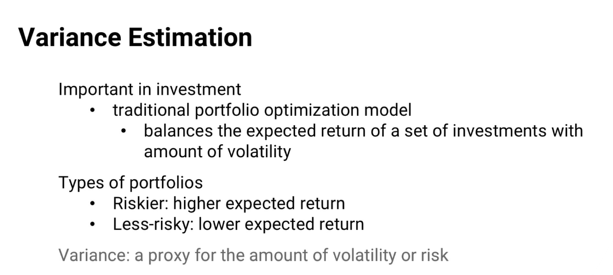
Generalized Auto regressive Conditional heteroscedasticity

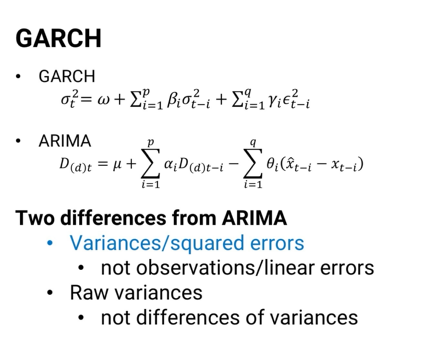
* Estimate or forecast the variance
* Time series

**Why to estimate a variance?**



If we know the variance of true vs. estimated value, we can do forecast based on these variance for future time series data





Garch is PQ model .

There is no “D”, because it doesn’t use diffeence