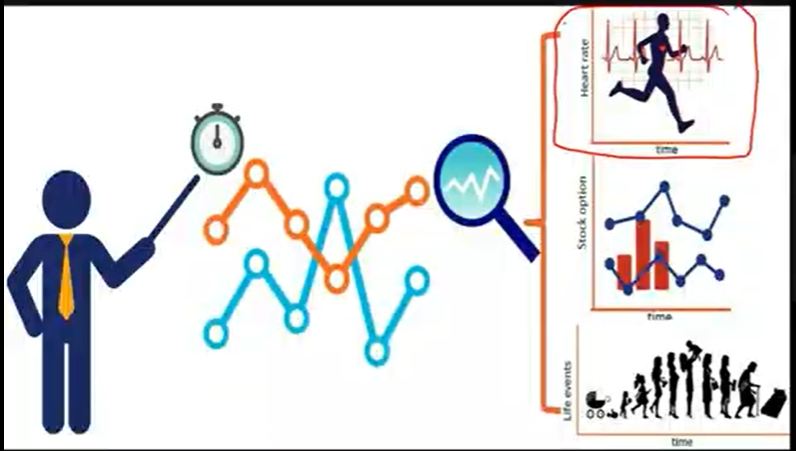


So in time series the data is timestamped may be from seconds to seconds to monthly, yearly,weekly,etc. And it is called as time-series data analysis. It’s a purely sequentially data.

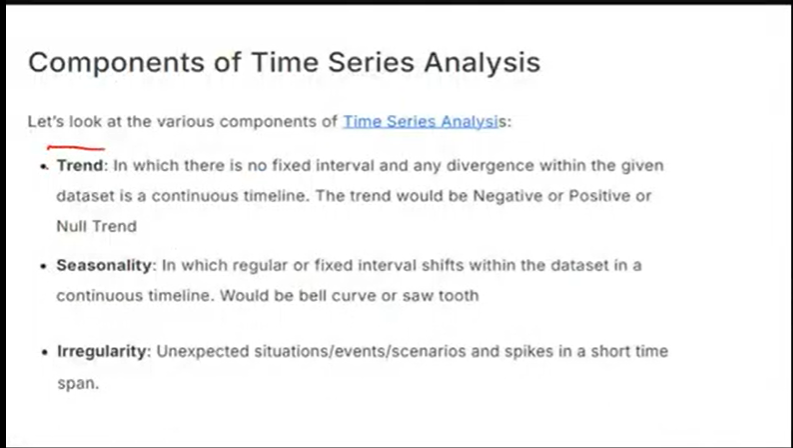
For eg. heart rate,stock options, life events,etc



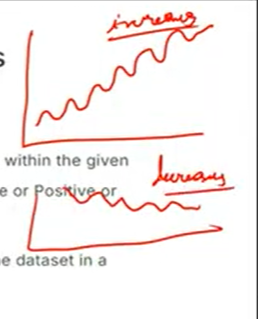
For the sequential data, rnn,lstm and gru are best.

There are some ml models too that help us in doing forecasting or doing future prediction.

Components of time-series data—



So trend is like increasing or decreasing trend. For eg—



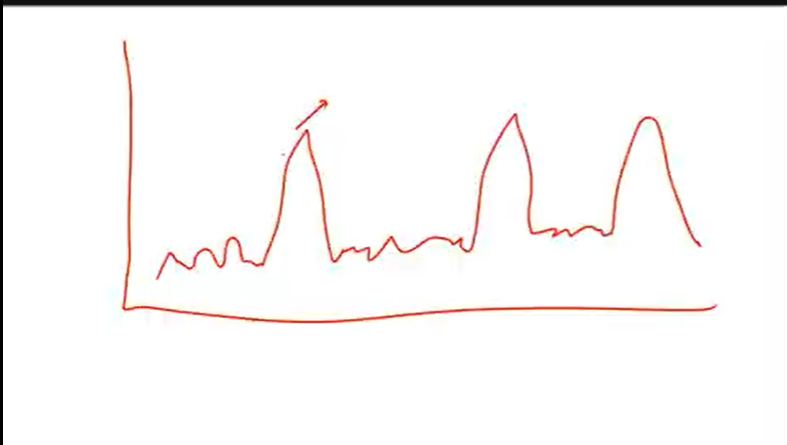
i.e. time series data is increasing wrt time or decreasing wrt time.

So time series data is composed of trend, seasonality and irregularity.

Postive trend is with increasing time stock price is also increasing,

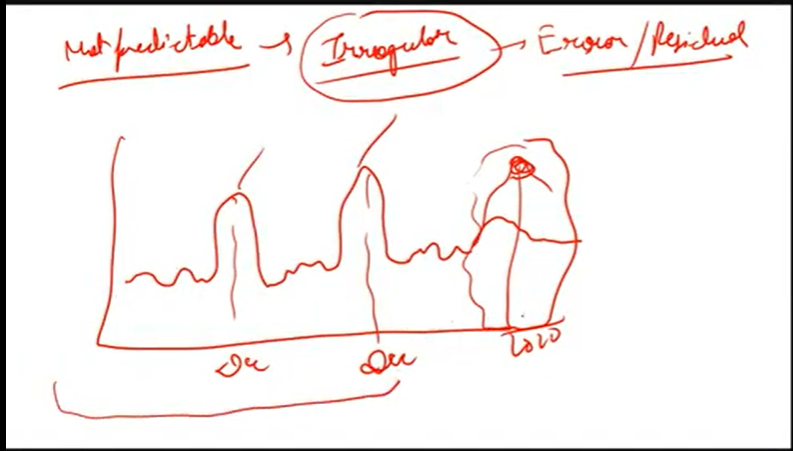
Negative trend is price decreasing with time.

Seasonality is pattern repeating itself in a fixed interval.



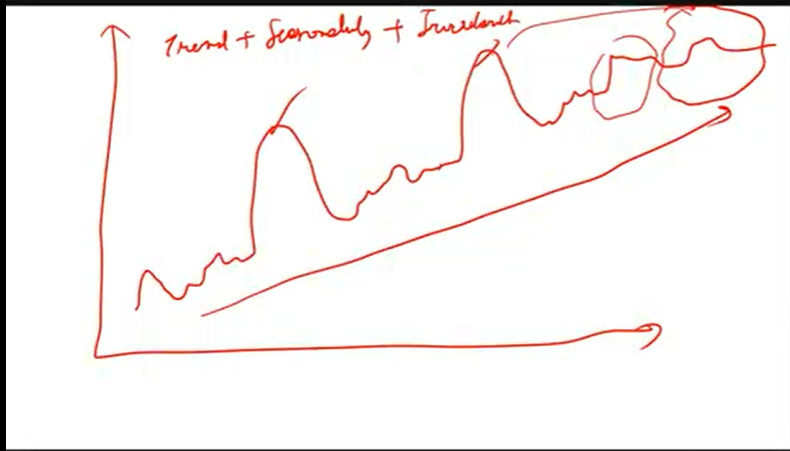
This has seasonality aspect. i.e. seasonality is repeating pattern.

Next is irregularity. For eg covid happened and stock prices fell. This can also be called as error or residual. For eg. lets suppose we have a seasonality and every December my chart is at peak. But because of covid instead of increasing the graph fell down. i.e.



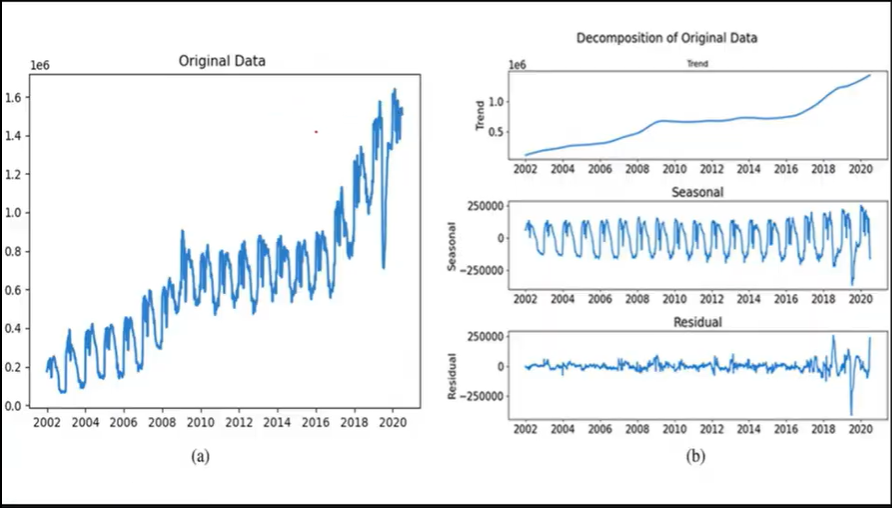
This is error/residual . Now because of this the complete analysis change. So irregularity is something which cant be predicted and because of which error comes in my data.

So this is time series data with trend, seasonality and irregularity.

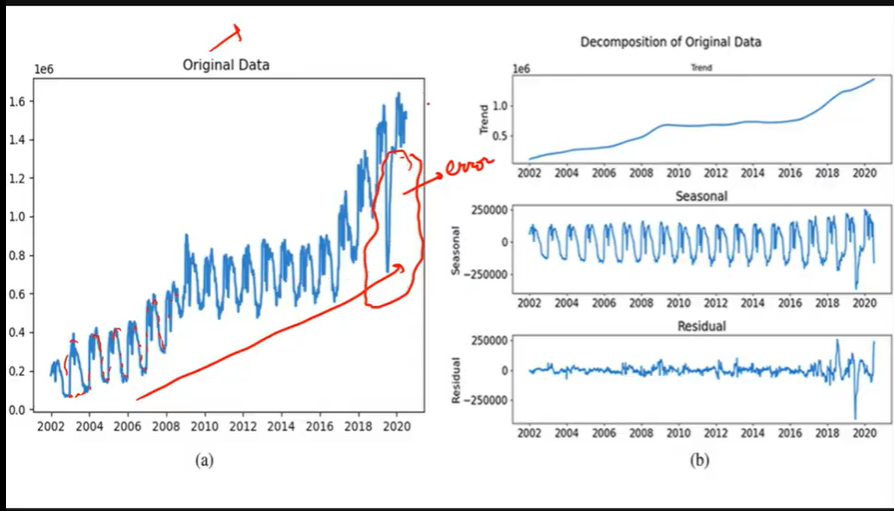


So, we can see trend, seasonality and some error at the end called as irregularity.

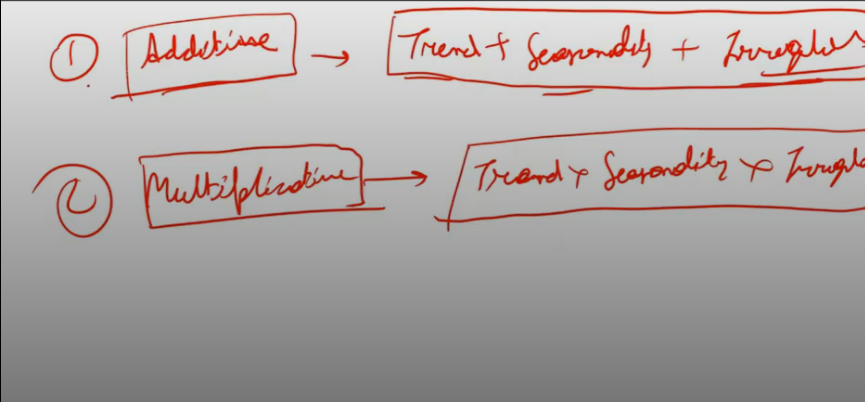
It is possible that in some dataset all of them are present while in some either one of them is present.



In the above original data if u will see, all 3 of them are present i.e. trend, seasonality and irregularity. i.e.



This is called as decomposiotion of data. Now any timeseries data has these 3 things only or a combination of these 3 things only. It is possible that timeseries data has no trend, seasonality but only irregularity is there. So we have 2 types of timeseries— Additive and multiplicative.

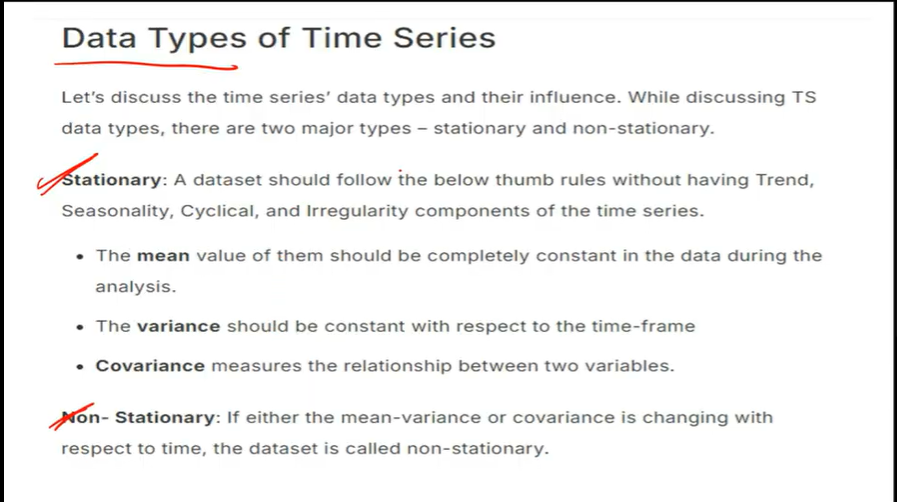


So In case of additive lets suppose--- trend is not there and seasonality and irregularity is there, then 0+1+1, the magnitude of time series data=2., if only seasonality is there then 1, and if all 3 of them are present then 3.

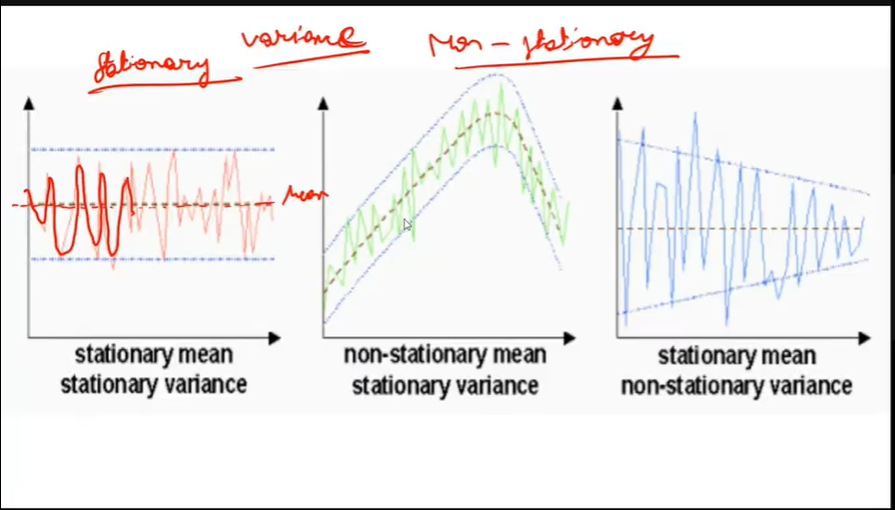
Now in case of multiplicative, if trend is not there, then itsvelue is 1 and those elements which are there, there value is >1, So additive time-series=1+1+1 whereas multyiplidate gives o/p after prod

Multipluication i.e. 1\*1\*1.

Next—



Stationary timeseries has constant mean and constant variance, for eg-

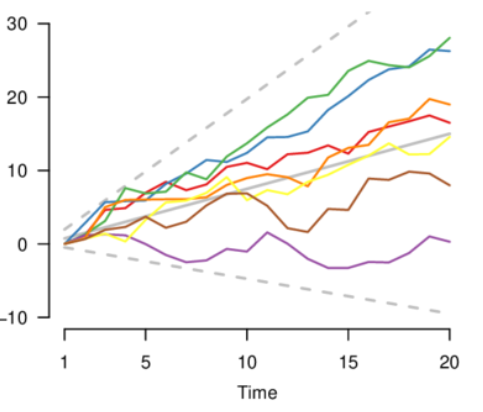


Mean and Variance i.e. dispersion between data points are important. The 2 straight lines are telling how dispersed the values are. The away lines are from the mid line that much is variance, closer to the mid line lesser variance, away from the mid line higher variance. Variance is the measure of dispersion.

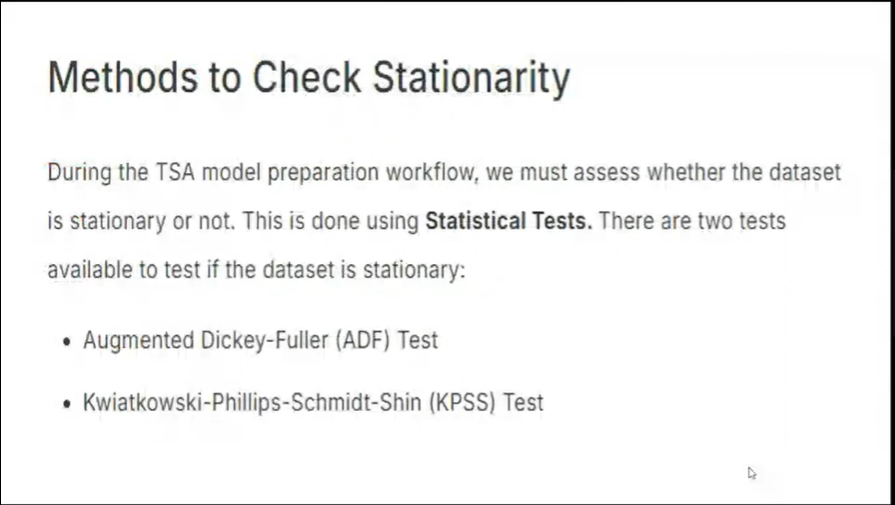
In the 2nd series , we can see mean increasing and then decreasing. So mean not constant but the variance is constant.

In 3rd series, mean is onstant but variance is not.

There can be a series where neither mean, nor variance is constant.

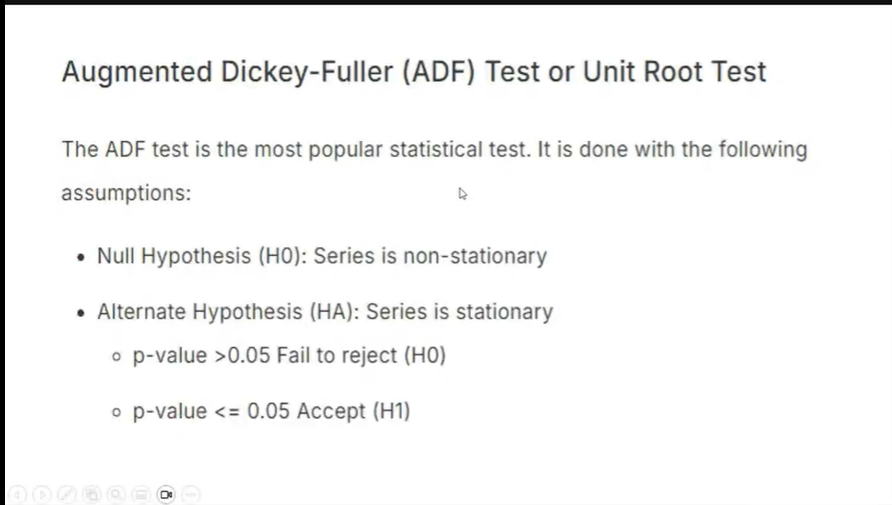


So we need some test to check if series is stationary or not.



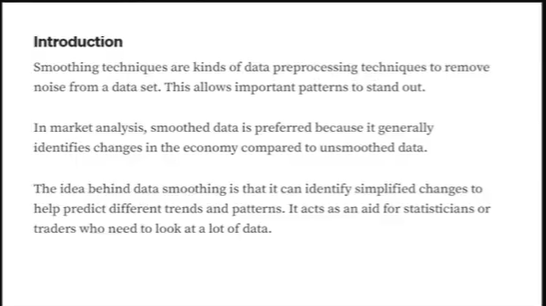
ADF and KPSS are the 2 statistical test, using which u can decide whether the series is stationary or not.

1st ADF test--



The idea is to achieve stationary data.

See time\_series.ipynb -- Only do till p-value . Next discuss smoothening techniques–



In timeseries data our goal is to predict future i.e forecast.

But in real life we generally don’t get stationary timesereis data.

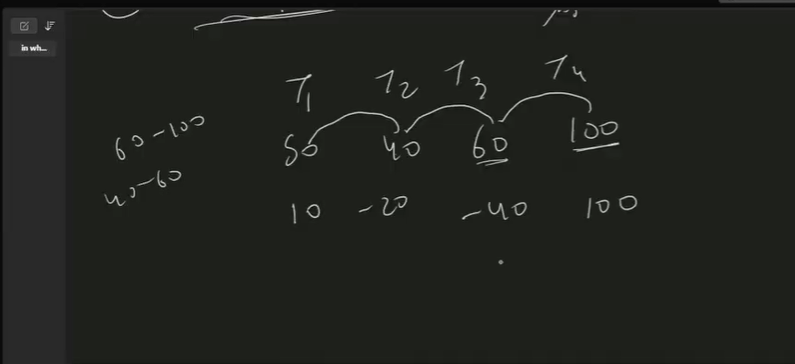
Steps to make non-stationary time series data into stationary upto some extent by using—

1. Differencing 2. Transformation.

Differencing is performed when mean is not constant.

Transformation is performed when variance is not constant.

Differencing is going back 1 time series and subtracting .

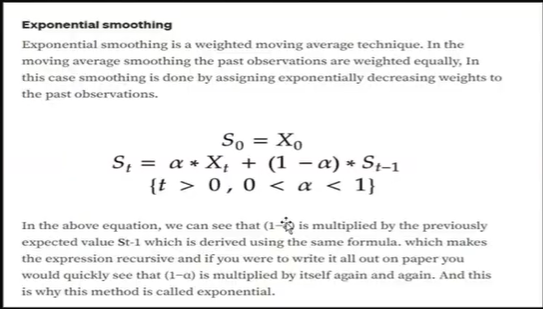


This I s done to make mean constant.

Transformation – sq root transformation or log root or cube root.Cube root ca find -ve no. cube root too. But sq root or log root I preferred. So the data on which differencing is performed over it perform log transformation or sq root transformation.

Next is forecasting or future prediction—

For that we do moving average. Now moving average alone is not applied. Because it assumes that in the that there is no trend, seasonality and residual. Moving average is also called as smoothening as it removes noise or error and u r left with seasonality and trend. Next is exponential smoothening –



Expalnation—we have time window—

T1 t2 t3 t4 t5 prediction

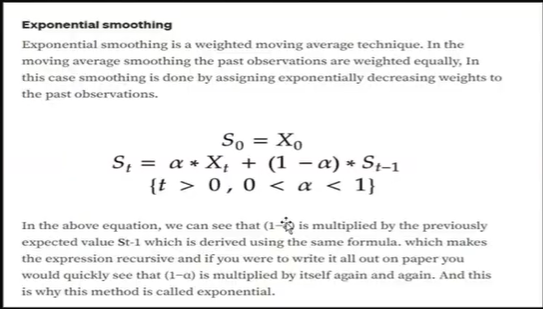
20 40 30 50 20 -- what we can do here is we can give more changes to the recent changes then to the past changes. i.e. more weightage to recent event and less to past event.

i.e. lets suppose t4 weightage is 0.2 and t5 weightage is 0.8 , then,

we can do something like (0.2\*50+0.8\*20)/ 2

i.e. giving more weightage to recent events and less weightage to past events.

So now looking at the formula—

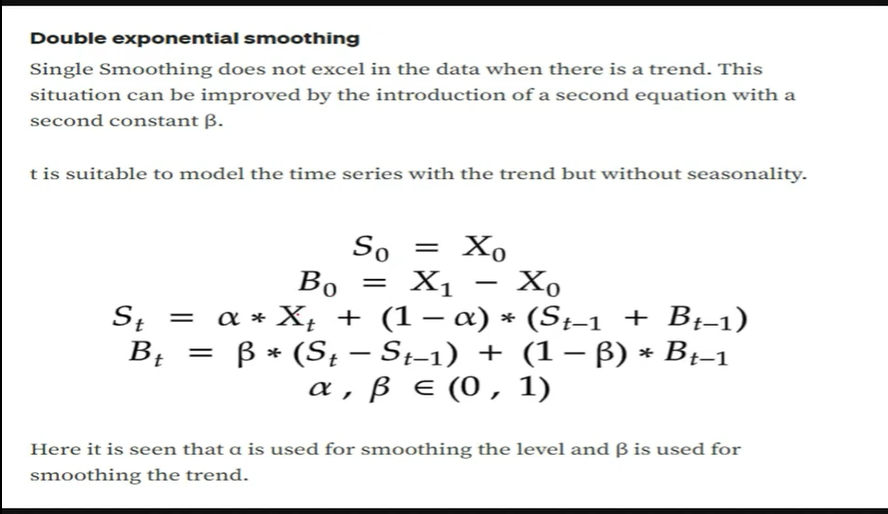


X\_t is the recent event and its weightage is alpha then the past event S\_(t-1) weightage is (1-alpha).

Now this alpha is the hyperparameter. Its value can range from 0.1 to 0.9.

Now the exponential Smoothening doesn’t perform good if there is trend in the data. So the solution is double exponential Smoothening.

In double exponential smoothening, the formula looks something like—



S-T is similar as before its just it has trend component added into it i.e. B\_(t-1).

So,B\_t captures the trend.

Now if seasonality is also there in the data, then use triple exponential smoothening.

Look for the formula in google. Noone will ask formula.

And if the data has only seasonality and no tred, then,

Open RNN\_TSA.ipynb— Here we can see trend seasonal and residual.

If seasonality wouldn’t have been there, then it would be a straight line and in seasonality we can see some small spikes,they are comparatively smaller or bigger at some palces that is irregularity.

Go to time\_series.ipynb --- airpassengers.csv – in airlines data we can see mean and variance is not constant. Also the trend and seasonality is there as pattern is repeating itself after certain amount of time. In the residuals graph we can see variance is not constant. So to get the best results triple exponentiation gives best result.

1st applying simple exponential smoothening and forecasting for next 6 steps.

Next in the graph the green line is the forecasted line. Now this is a straight line because single smoothening cannot find out trend or seasonality factor.

Next Double exponentiation is also called as holt. We can see it can detect trend but not seasonality.

That’s why in forecast, trend is slant straight line.

Next is triple exponentiation. Now it is predicting trend and seasonality.

Now these are old models.

We have new models now. So basically in time-series we have to find those past data points which can help us in predicting future data poins. In ml we try to find the correlation of different columns with the target, in the same way ,in time-series data , we try to find window size or number of past values which will affect my future prediction. Now in time seris , when I say lag=1, that means im going 1 step back and taking data.

i.e. if

t1 t2 t3 t4 t5 then lag1 will consider t2 t3 t4 t5

while the original data will be t1 t2 t3 t4 . Our moto is to use original data and predict t5. So if we can find some correlation between original and lagged data, then we can predict the t5 data and then t6 etc.

lly for lag=2 , t1,t2,t3 whereas original will be t3,t4,t5. And we will check the correlation. So we will check for last 2 data points. And I will try for lag 3 ,lag 4, lag 5, the moment I will know that now the correlation is not there, then stop. And the lag just before this will be the window size.

So to find the window size, we use acf and pacf. ACF is auto-correlation function and pacf is partial auto correlation function.

So when u r applying auto-regressor, to find window size use pacf.

And to know the window size of moving average use acf.

Now ma we don’t use separately instead we use with ar, and it becomes ARMA model i.e. auto regressor moving average model.

end\_to\_End\_time\_series\_analysis.ipynb

Next is Future prediction – forecasting. U can use lstm, gru, rnn. We will do with lstm.

36:34 – 7th August

Next is ARIMA. I is Integrated. It is better then ARMA. This factor I controls non-stationarity.

If there is even a small amount of non-stationarity left in the data, then it is handled by I. it is also called as differencing. As discussed earlier, to make a non-stationary time-series stationary, we can do differencing and transformation. Differencing is done to make means constant. So in ARIMA, u don’t have to the differencing, the model itself will do differencing. This is decided by factor d. Its value can be chosen from 0 to 3 , not more then that.