Business Forecasting

ADIA Course

Day 2 – Session 1

Dr. Tanujit Chakraborty

Sorbonne University, Abu Dhabi, UAE

Sorbonne Centre for Artificial Intelligence, Paris, France

Research Areas: Time Series Forecasting, Machine Learning, Econometrics, Health Data Science

July 1, 2025

Business Forecasting Process

1. Problem definition: Involves understanding how the forecast will be used, expectations of the customer or user of the forecast, how often forecast need to be revised, what level of forecast accuracy is required to make good business decision.

2. Data collection: Consists of obtaining the relevant history for the variable(s) to be forecasted, including historical information on potential predictor variables.

3. Preliminary data analysis: Needed for selection of the appropriate forecasting model, to identify the patterns such as trends, seasonal and other cyclic components. Numerical summaries such as sample mean, standard deviation, percentiles, stationarity, nonlinearity, and autocorrelations need to be computed and evaluated. Unusual observations or potential outliers need to be identified and flagged for possible further study. If predictor variables are involved, scatter plots of each pair of variables should be examined.

Business Forecasting Process

4. Model selection and fitting: Choosing one or more forecasting models and fitting the model to the data (estimating the unknown parameters of the model).

4. Model validation: Evaluation of the forecasting model to determine how it is likely to perform in the intended application. The magnitude of forecast error need to be examined not only on historical data but also on fresh or new data.

5. Forecast model deployment: Consists of using the model to forecast the future values of the variable of interest for the customer.

6. Monitoring forecasting model performance: An ongoing activity after the model has been deployed to ensure that the model still performing satisfactorily. Monitoring of forecast errors is an essential part of good forecasting system design.

Business Forecasting Case Study Problem 1

Client: Car fleet companies around the world.

Business Forecasting Case Study Problem 1

Problem: how to forecast resale value of vehicles? How should this affect leasing and sales policies?

Additional information:

- They can provide a large amount of data on previous vehicles and their eventual resale values.
- The resale values are currently estimated by a group of specialists. Statistical/ML-based forecasters can help the Car fleet companies.

Business Forecasting Case Study Problem 2

SORBONNE UNIVERSITÉ

Client: Airline Company.

Business Forecasting Case Study Problem 2

SORBONNE UNIVERSITÉ

Problem: How to forecast passenger traffic on major routes (say, Economy class passengers)?

Additional information:

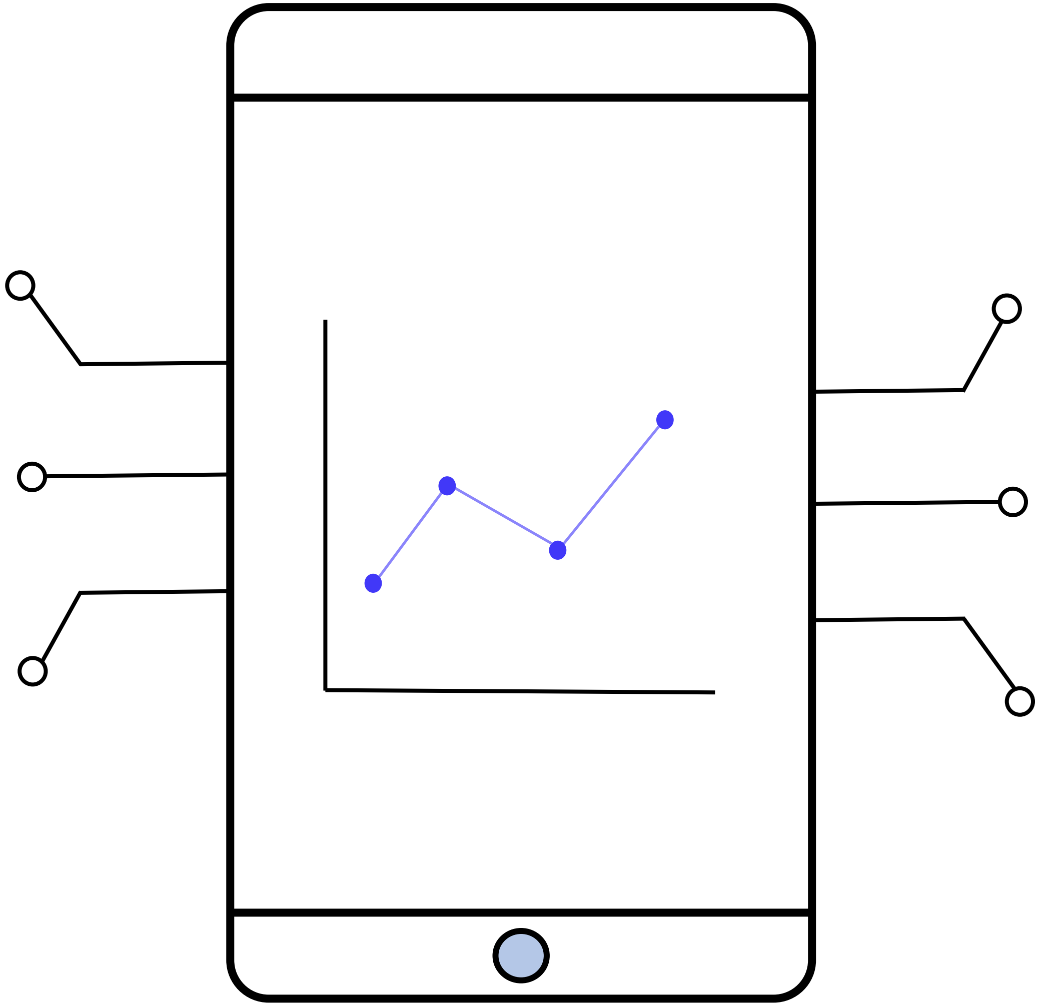
- They can provide a large amount of data on previous routes.
- Traffic is affected by school holidays, special events such as the FIFA World Cup, advertising campaigns, competition behaviour, etc.
- They have a highly capable team of people who are able to do most of the computing.

What makes for a good forecasting model?

Forecast accuracy

Resource requirements

Speed



Data leakage

Understandability & explainability

Maintainability & debuggability

Baseline Forecasting Methods

SORBONNE UNIVERSITÉ

Pretrained Models

Single model used across all tasks

- LLMTime
- ForecastPFN
- Lag-Llama
- Monai
- Chronos

Task-specific Models

Separate model fine-tuned for each task

- PatchTST
- DeepAR
- WaveNet
- TFT
- DLinear
- NBEATS
- NHITS
- CatBoost

Local Models

Separate model for each time series

- Naïve
- ETS
- ARIMA

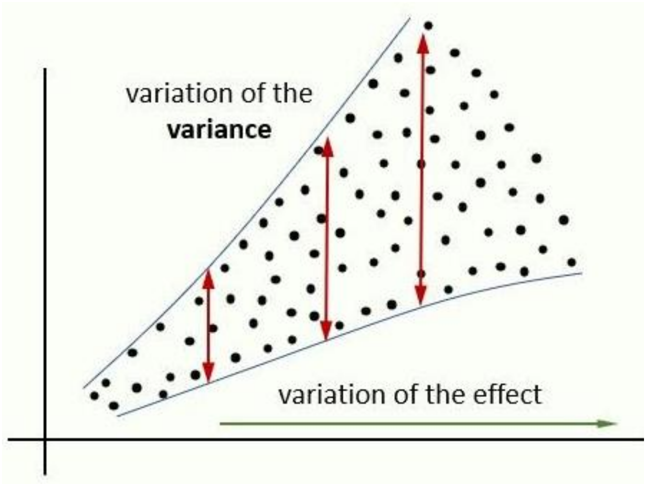
Auto-correlation Analysis

Auto Regression Analysis

- Regression analysis for time-ordered data is known as Auto-Regression Analysis

- Time series data are data collected on the same observational unit at multiple time period.

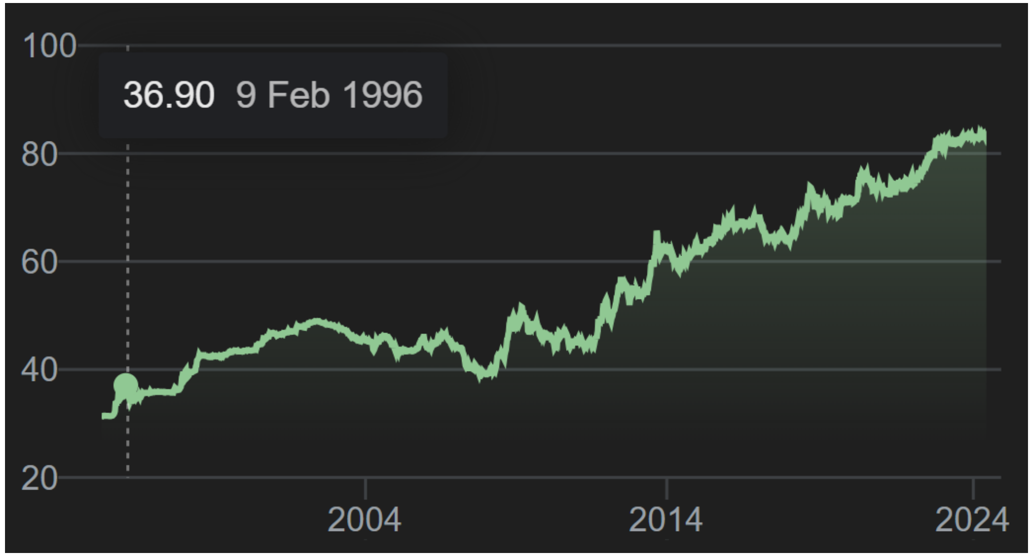
Example: Indian rate of price inflation



Heteroscedasticity

Modeling with Time Series Data

- Correlation over time
  - Serial correlation, also called autocorrelation
  - Calculating standard error
- To estimate dynamic causal effects
  - Under which dynamic effects can be estimated?
  - How to estimate?
- Forecasting model
  - Forecasting model build on regression model



Can we predict the trend (USD vs. INR) at a time, say 2025?

Some Notations and Concepts

- Y_t = Value of Y in a period t
- Data set [Y_1, Y_2, … Y_{T-1}, Y_T]: T observations on the time series random variable Y

There are four ways to have the time series data for AutoRegression analysis

- Lag: The first lag of Y_t is Y_{t-1}, its j-th lag is Y_{t-j}
- Difference: The fist difference of a series, Y_t is its change between period t and $t-1$, that is, $y_t = Y_t - Y_{t-1}$
- Log difference: $y_t = \log(Y_t) - \log(Y_{t-1})$
- Percentage: $y_t = \frac{Y_{t-1}}{Y_t} \times 100$

Related Concepts and Notations

Assumptions

1. **Uniform:** We consider only consecutive, evenly spaced observations

   For example, say monthly data in 2010-2021 for each year, and without any missing month(s); no other data, for example, on daily basis for a year is admissible.

2. **Stationarity:** A time series Y_t is stationary if its probability distribution does not change over time, that is, if the joint distribution of $(Y_{i+1}, Y_{i+2}, Y_{i+3}, \ldots . Y_{i+T})$ does not depend on i.

   Stationary property implies that history is relevant. In other words, stationary requires the future to be like the past (in a probabilistic sense).

   Auto-regression analysis assumes that Y_t is both uniform and stationary.

Auto-correlation coefficient

Just as correlation measures the extent of a linear relationship between two variables, autocorrelation measures the linear relationship between lagged values of a time series.

- There are several autocorrelation coefficients, corresponding to each panel in the lag plot.
- For example, r_1 measures the relationship between y_t and y_{t-1}, r_2 measures the relationship between y_t and y_{t-2}, and so on.

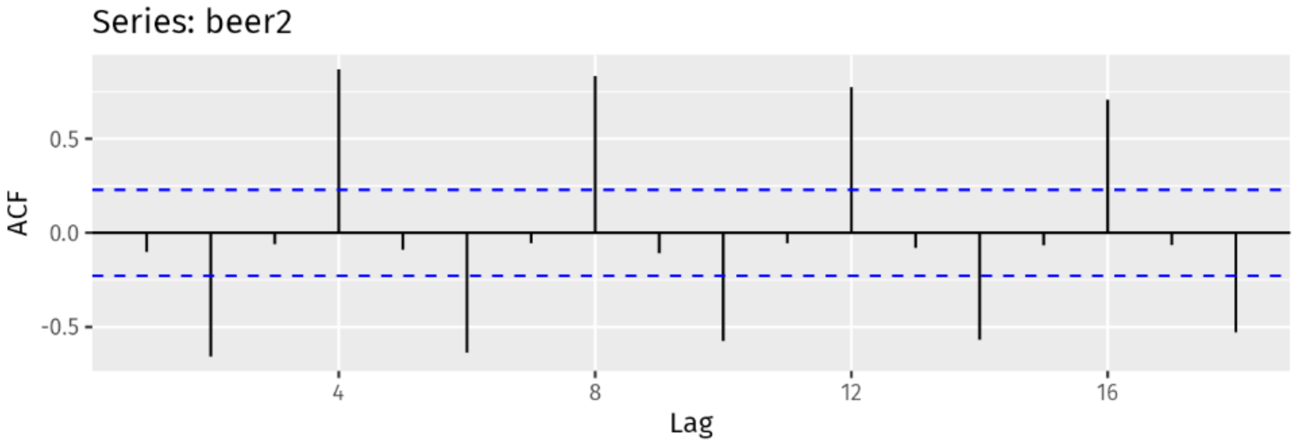
The value of r_k can be written as

$$r_k = \frac{\sum_{t=k+1}^{T} (y_t - \bar{y})(y_{t-k} - \bar{y})}{\sum_{t=1}^{T} (y_t - \bar{y})^2}$$

where T is the length of the time series. The autocorrelation coefficients make up the autocorrelation function or ACF.

ACF

The autocorrelation coefficients are plotted to show the autocorrelation function or ACF of quarterly beer production in Australia. The plot is also known as a correlogram.



In this graph:

- r_4 is higher than for the other lags. This is due to the seasonal pattern in the data: the peaks tend to be four quarters apart and the troughs tend to be four quarters apart.
- r_2 is more negative than for the other lags because troughs tend to be two quarters behind peaks.
- The dashed blue lines indicate whether the correlations are significantly different from zero.

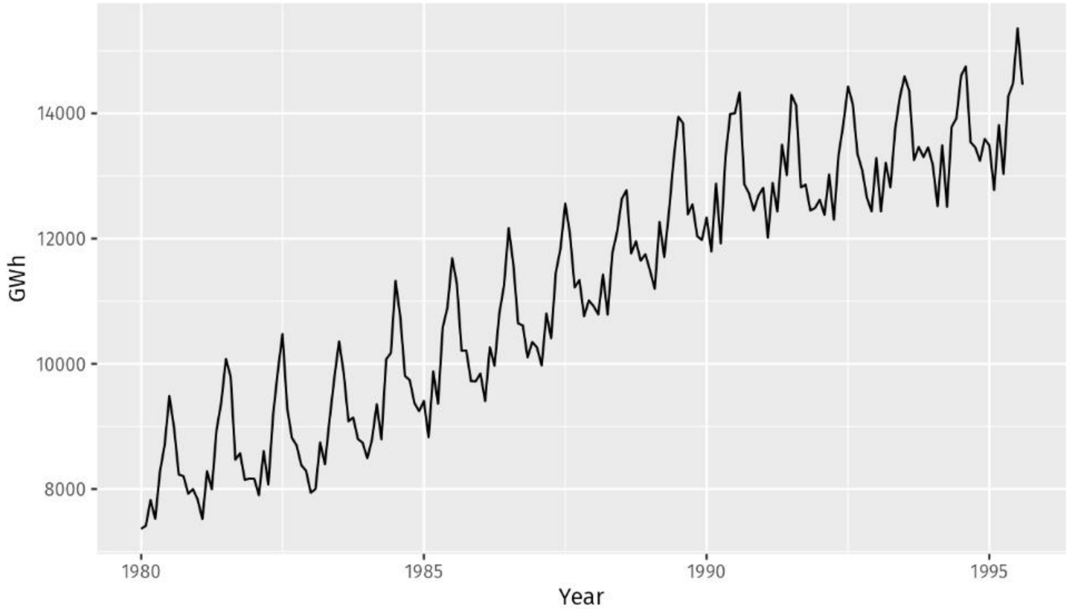
Example: Autocorrelation

- When data have a trend, the autocorrelations for small lags tend to be large and positive because observations nearby in time are also nearby in size. So the ACF of trended time series tend to have positive values that slowly decrease as the lags increase.

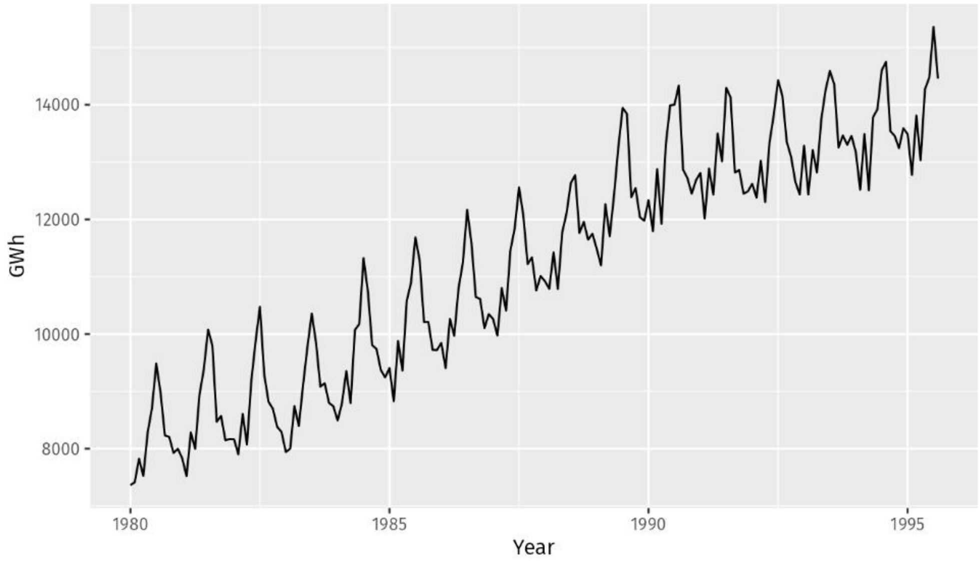
- When data are seasonal, the autocorrelations will be larger for the seasonal lags (at multiples of the seasonal frequency) than for other lags.

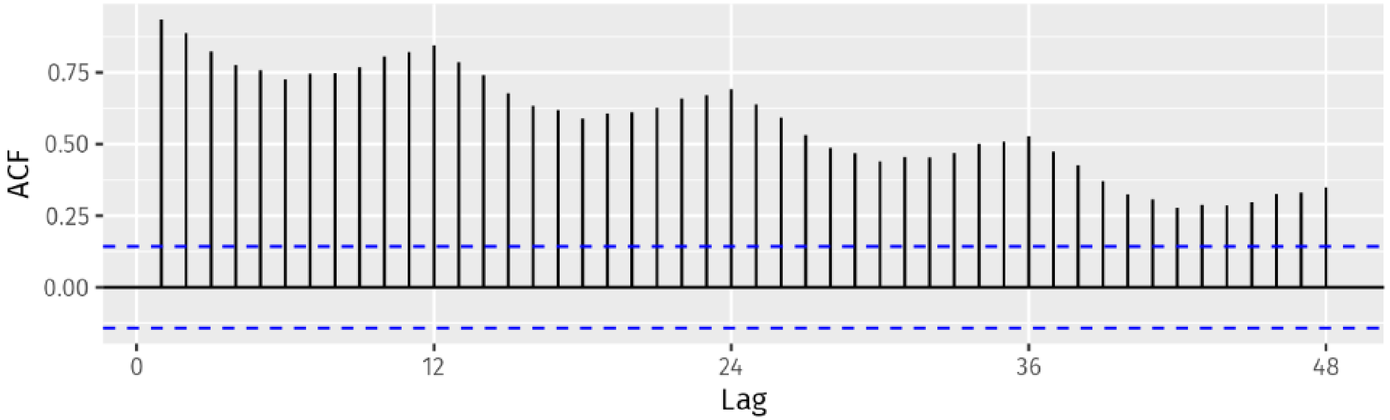
- When data are both trended and seasonal, you see a combination of these effects.

- The monthly Australian electricity demand series plotted in Figure shows both trend and seasonality.



Example: Autocorrelation





The slow decrease in the ACF as the lags increase is due to the trend, while the “scalloped” shape is due to the seasonality.

Baseline Model: Exponential Smoothing

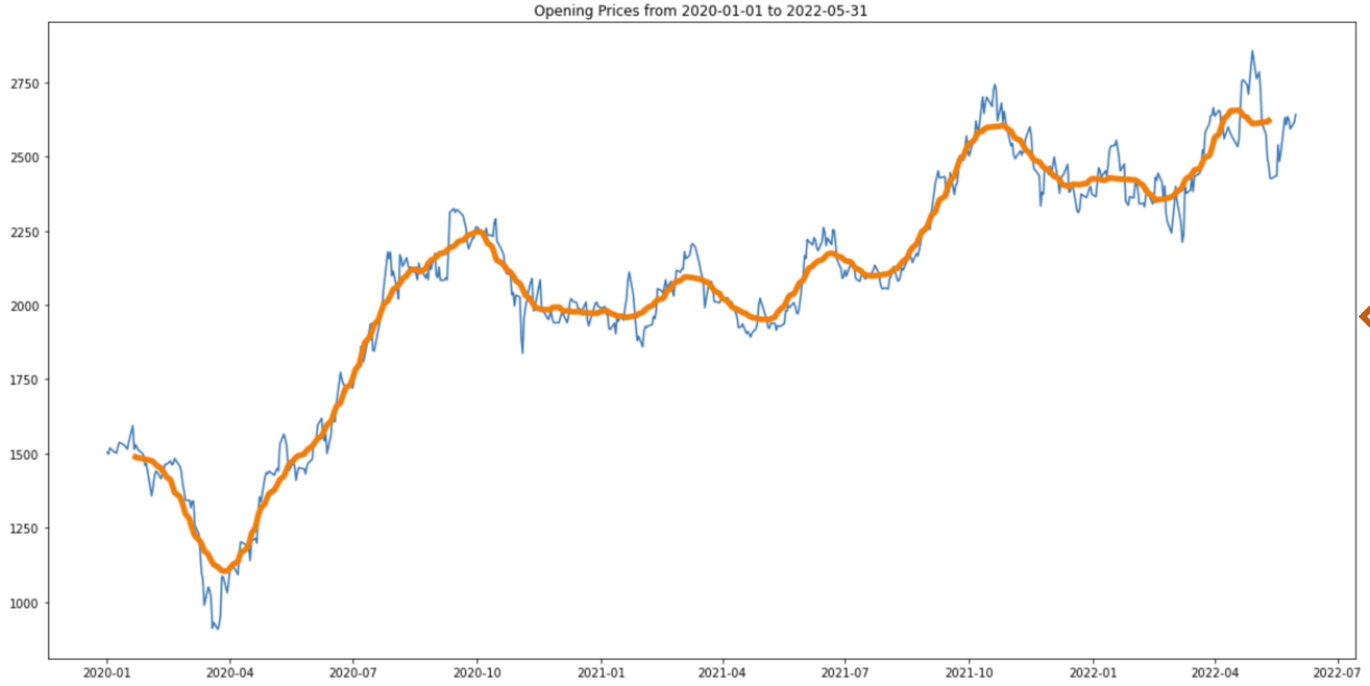
Finally at Non-Naive Forecasting!

Which models to be learnt?

- **Simple Exponential Smoothing** (SES)
  - Non Trending, Non Seasonal (Simplest)
- **Holt Model**
  - Trending, Non Seasonal (Average)
- **Holt Winters**
  - Trending, Seasonal (Amazing)

The Simple Moving Averages

- Simple is simple. No arguement.
- It is used to show the *zaggy* data in a smooth manner
- Helps to give the overall trend of the time-series
- Used to show the intuition of the data without showing much information



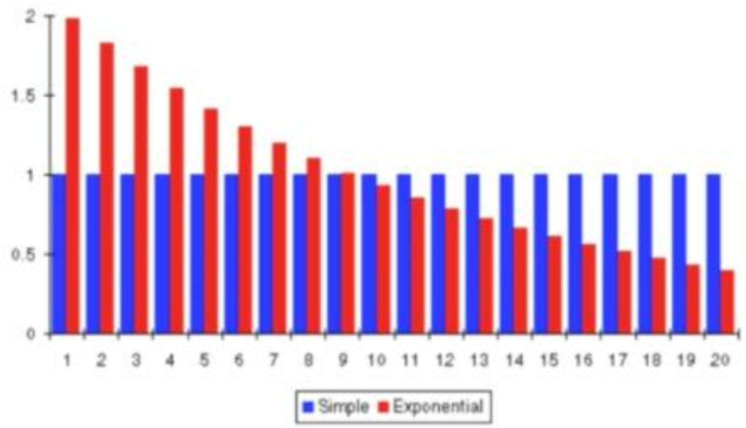
Too Simple right?

Can you think of something better!

The Exponentially Weighted Moving Average

- This is the Exponentially Weighted Moving Average (EWMA) where the weight is given exponentially lower as the data point gets older.
- The moving average is designed as such that older observations are given lower weights. The weights fall exponentially as the data point gets older – hence the name exponentially weighted.

$$\text{EWMA}(\hat{x}_{t+1}) = \alpha x_t + (1-\alpha)\hat{x}_t$$



Why Exponential?

Current Situation:

$$\text{EWMA} = \alpha x_t + (1-\alpha)\hat{x}_t$$

Which includes:

$$\text{EWMA} = \alpha x_t + (1-\alpha)[\alpha x_{t-1} + (1-\alpha)\hat{x}_{t-1}]$$

Which Becomes:

$$\text{EWMA} = \alpha x_t + (1-\alpha)\alpha x_{t-1} + (1-\alpha)^2\hat{x}_{t-1}$$

Again:

$$\text{EWMA} = \alpha x_t + (1-\alpha)\alpha x_{t-1} + (1-\alpha)^2[\alpha x_{t-2} + (1-\alpha)\hat{x}_{t-2}]$$

Which becomes:

$$\text{EWMA} = \alpha x_t + (1-\alpha)\alpha x_{t-1} + (1-\alpha)^2\alpha x_{t-2} + (1-\alpha)^3\hat{x}_{t-2}$$

We can see that:

$$(1-\alpha) \to (1-\alpha)^2 \to (1-\alpha)^3 \to \cdots \to (1-\alpha)^{t-1}$$

Exercise

Consider the time series with nine periods of data:

34, 38, 46, 41, 43, 48, 51, 50, 56

Use EWMA to forecast the value for time period 10.

Assume $\alpha = 0.2, x_1 = \hat{x}_1 = 34$.

Solution

$$\text{EWMA}(\hat{x}_{t+1}) = \alpha x_t + (1 - \alpha)\hat{x}_t$$

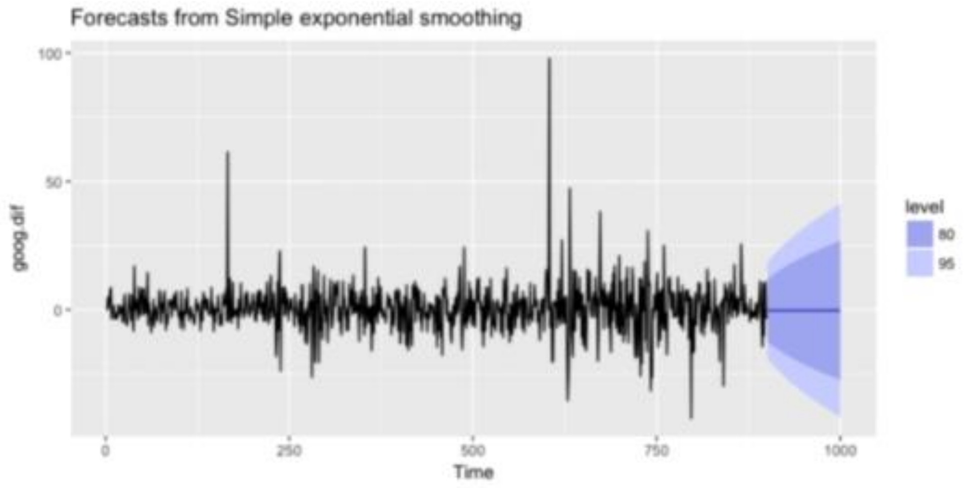
Time Period (t)	Actual (x_t)	$\hat{x}_{t+1} = \alpha x_t + (1 - \alpha)\hat{x}_t$
1	34	34.00
2	38	0.2 * 34 + 0.8 * 34.00 = 34.00
3	46	0.2 * 38 + 0.8 * 34.00 = 34.80
4	41	0.2 * 46 + 0.8 * 34.80 = 37.40
5	43	0.2 * 41 + 0.8 * 37.40 = 38.12
6	48	0.2 * 43 + 0.8 * 38.12 = 39.09
7	51	0.2 * 48 + 0.8 * 39.09 = 40.87
8	50	0.2 * 51 + 0.8 * 40.87 = 42.89
9	56	0.2 * 50 + 0.8 * 42.89 = 44.31
10	-	0.2 * 56 + 0.8 * 44.31 = 46.64

Hence, the forecast for time period 10 using EWMA is 46.64.

SES Model

- No Trend, No Seasonality model
- Assumes that there are some **fluctuations** in the data & those fluctuations are around some constant value.
- Thus, the model tries to learn what the average value is - by using the EWMA method.
- To forecast (I repeat), to forecast the value **it assumes that the same EWMA value will be propagated to the future**, because that was what the value was back in time. (around which the values were fluctuating)
- That constant value is called: The Level in this ETS terminology

$$\text{level } (t+h) = \text{ EWMA(time series)}$$



Holt Linear Trend Model

- There is Trend, No Seasonality
- See, the trend has to be linear - either positive or negative (assuming not like the cosine ~ curve! (:))
- It uses **2** EWMAs!! ***For the level*** ***For the trend***
- In the forecast, it is just the linear combination Level and Trend - and that is our linear equation!

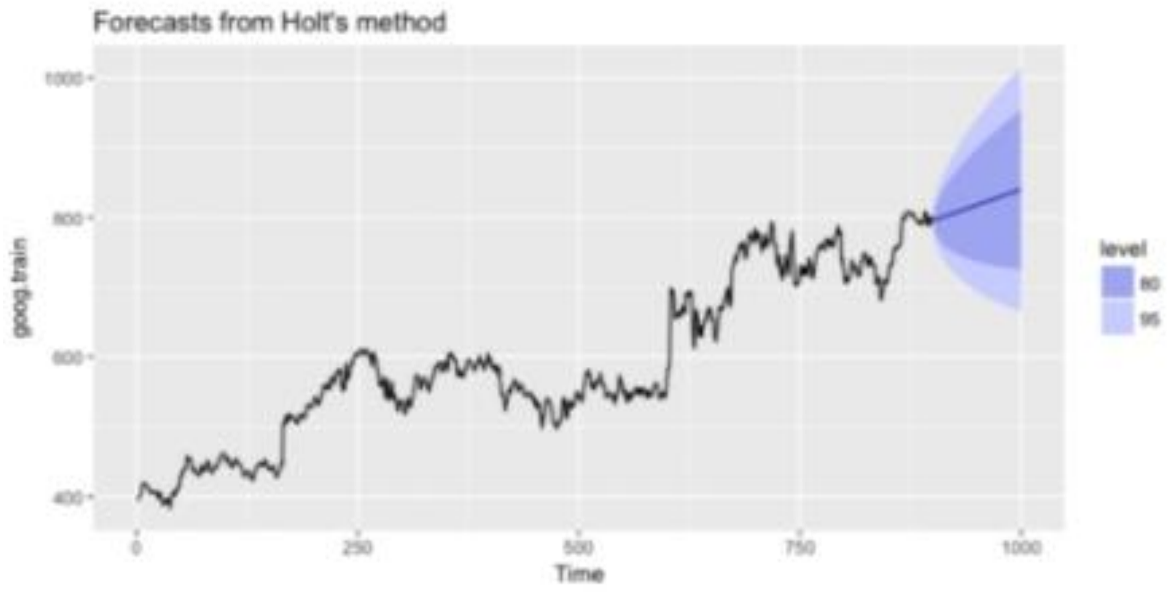
$$y = \beta_0 + \beta_1 x$$

$$\text{forecast} = \text{level} + \text{trend} \times \text{h}$$

- Again, h: is the number of steps in the future
- So the 2 EWMAs are:

$$\text{level } (t + h) = \text{ EWMA(level of time−series)}$$
$$\text{trend } (t + h) = \text{EWMA(trend of time−series)}$$

Holt Linear Trend Model

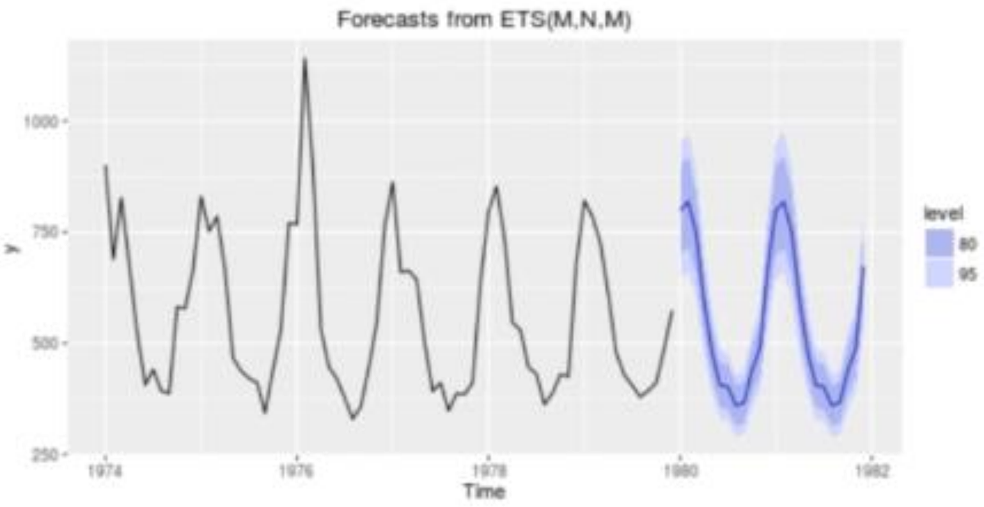


- As can be seen in the image, the line goes in the upward direction.
- Which is the result of the intercept = level and 5 lope = trend. (Better than the SES model, just the horizontal line)

Holt Winter Model

- There is seasonality, There is trend
- It means, the constant value will stay constant for each year's each season.
- If the selling of book in April 2021 is +5 then it will be +5 in the April 2022 (might increase or decrease based on the trend and level) but as per to say... it will be +5 .

$$
\begin{aligned}
level\ (t+h) &= EWMA(level\ of\ time{-}series) \\
\mathrm{tren\ d}(t+h) &= EWMA(rend\ of\ time{-}series) \\
seasonal\ (t+h) &= EWMA(seasonal\ of\ time{-}series)
\end{aligned}
$$



And, there can be different ways to forecast
1. Either: Forecast = Trend + Level + Season
2. Either: Forecast = Trend x Level + Season
3. Either: Forecast = Trend + Level x Season
4. Either: Forecast = Trend x Level x Season

Time Series Approaches

SORBONNE UNIVERSITÉ