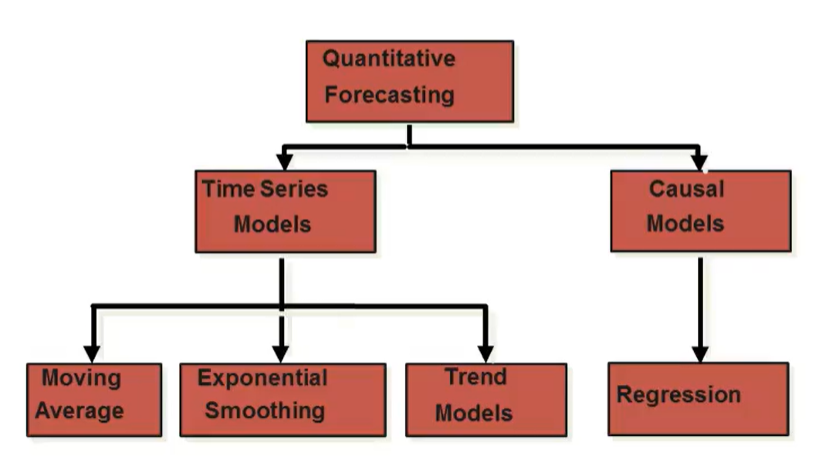
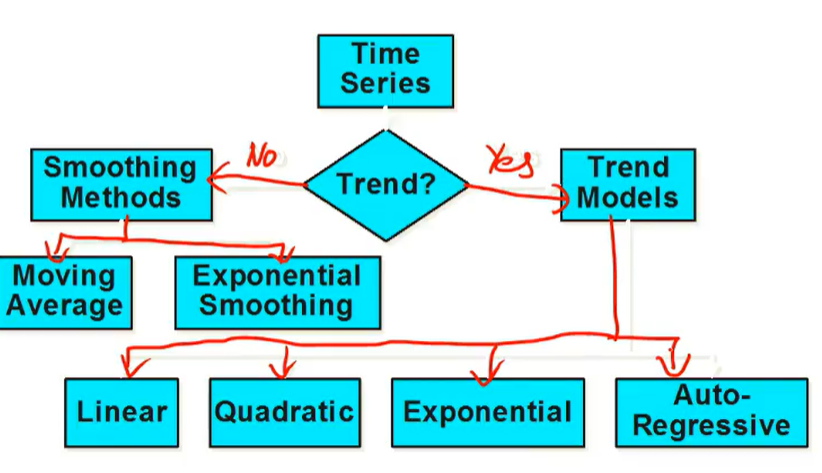
# Time series forecasting

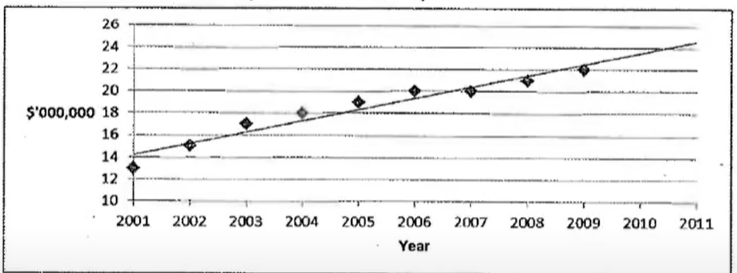
* variable – time
* x – axis -> time
* y – axis -> Magnitude
* Definition:
  + Set of observation taken at specified times usually at equal intervals.
  + Qualitative(Lots of data there by it is accurate) forecasting.
  + Evenly spaced numeric data (either monthly basis/ yearly basis)
  + Dynamic
* Uses:
  + Business forecasting
  + Analyse past behaviour
  + Plan future
  + Evaluate current accomplishment (goals met).
* Components:
  + Trends(Demand)
    - Long term underlying growth movement in a time series.
    - Up trend
    - Down trend
    - Stationary trend – No trend
  + Seasonality
    - Repeating pattern within a fixed time period.
    - Ex: Christmas(Chocolate) – Summer(IceCream)
  + De Seasonalize
    - Removing the seasonal patterns to easily analyse the underlying trend.
    - Steps
      * Find the mean for each month/quater
  + Irregularity
    - Non repeating
    - Unsystematic
    - Short duration and non – repeating
    - Ex: Flood(Medical)
  + Cyclic
    - Repeating (harder to predict)
    - Shorter duration than trend.
    - Repeated up and down movements.
    - Cycles may vary in length.
* When not to use:
  + Values are constant
  + Values in the form of function
    - Because by just substituting u’ll get the ans.
* Stationarity:
  + Stationary time series:
    - Data is gonna be independent of other year.
    - Constant mean(Average), variance(Distance form mean).
    - Autocovariance doesn’t depend on times.
    - Has the best linear predictor.
* Autocovariance:
  + Measure that quantifies linear relationship between two points in a time
    - +ve - similar
    - -ve - Not similar
* Tests to check stationarity:
  + Rolling Statistics
    - Plot the moving variance and see if it varies with time.
  + ADCF Test
    - Null Hypothesis
      * TS is not stationary
    - Test Results comprises,
      * Test Statistics and Critical values
* ARIMA Model – AR+MA
  + AR – Auto Regressive
    - Corelation between previous time period and current.
  + MA – Moving Average
  + Integration joins AR + MA
  + Parameters
    - P – Autoregressive lags(lagged values)
    - Q – Moving Average
    - d – Order of Integration
* Interpolation:
  + Predicting the point in between
* Extrapolation:
  + Predicting the point beyond the given points
* Forecast:
  + Process of predicting the future.

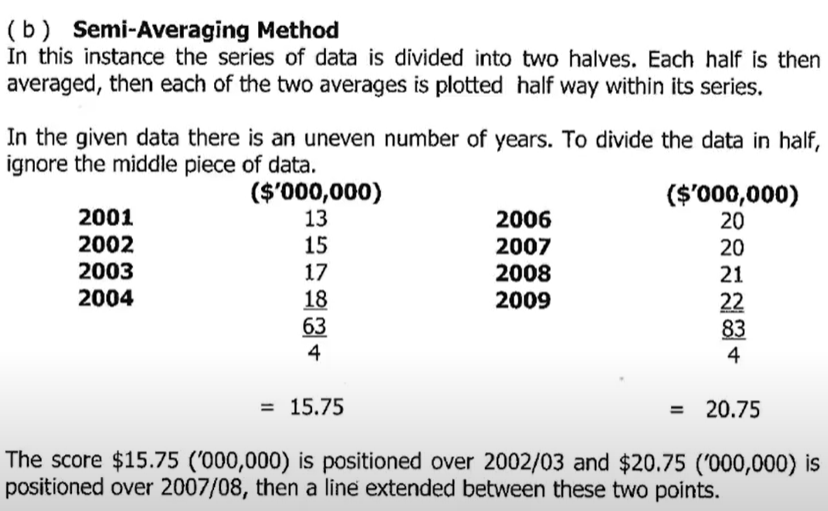
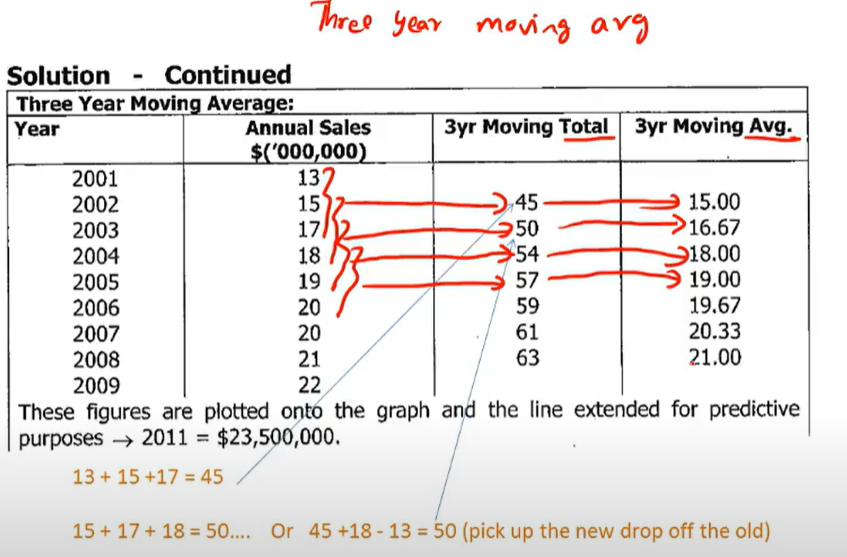
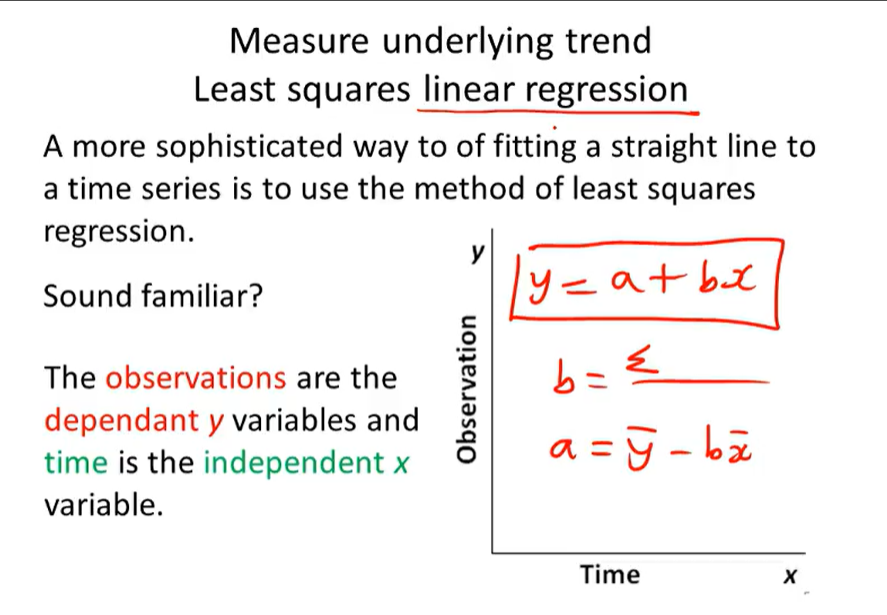




Methods for depicting trend:

* Free hand drawing



* Semi – average
  + Divide data into 2 equal time ranges
  + Calculate the average of the observations in each of 2 time ranges.
  + Draw a straight line between the 2 points.
  + Extend the line slightly past the end of the original observation to make predictions for future years.
  + 
* Moving – average
  + 
* Least – squares method(Linear Regression)
  + 
* Exponential smoothing
* Weighted exponential smoothening.

|  |  |
| --- | --- |
| Linear Regression | Time Series Analysis |
| General data points | Based on time |
|  |  |
|  |  |
|  |  |
|  |  |

Noise example

For instance, suppose a manufacturing plant has been producing a certain product with a stable daily output for several weeks. However, on a particular day, the output of the plant drops unexpectedly, despite no apparent problems with the production process or equipment. This drop in production can be attributed to noise or random fluctuations in the production process, rather than any underlying trend or pattern.

In this example, the noise in the daily output could be due to various reasons, such as fluctuations in the availability of raw materials or changes in the quality of the raw materials. Understanding the sources of noise in the production process and minimizing their impact can help improve the overall efficiency and productivity of the manufacturing plant. This could involve implementing process control measures, improving the quality of the raw materials, or optimizing the production schedule.

lag value example

For example, suppose a manufacturing plant produces a certain product, and the output is measured after a lag of one hour from the time of production. This lag value of one hour indicates that there is a delay between the time when the product is produced and the time when its quality and quantity can be assessed. If the lag value is too high, it can result in excessive inventory or production bottlenecks, which can lead to lower efficiency and increased costs.

Forecasting:

* Multiplicative - Si\*Ii\*Ti
  + When the difference between mean and the peak is not constant.
* Addictive - Si+Ii+Ti
  + When the difference between mean and the peak is constant.