Stationary Time Series Analysis

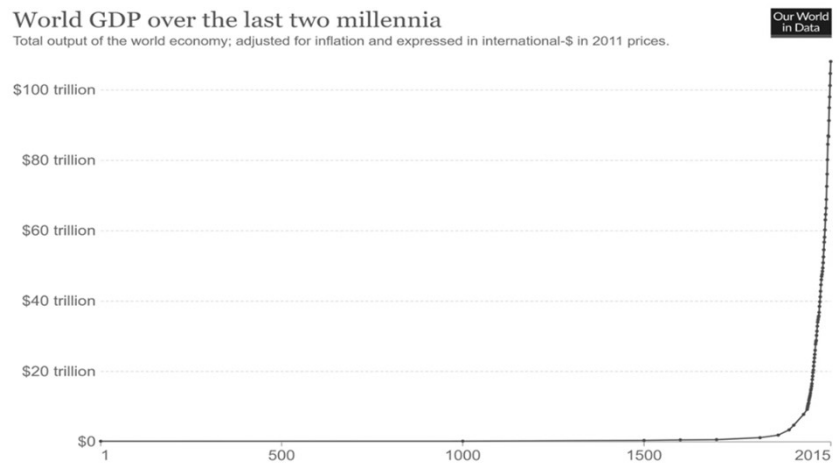
Muhammad Rafi

November 21, 2023

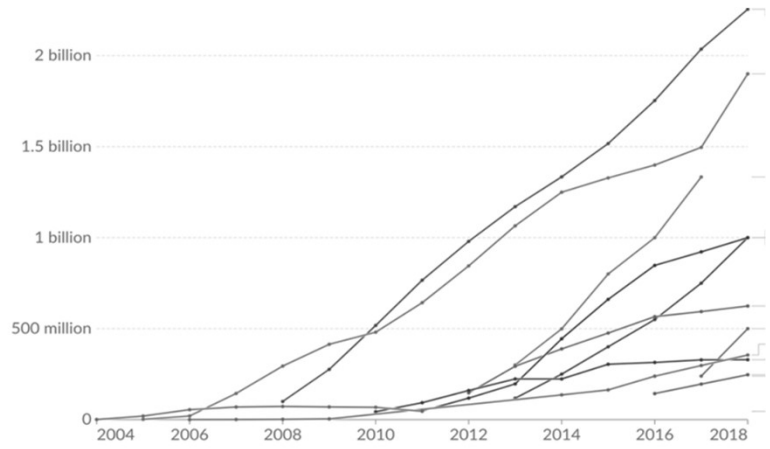
Outline

- Forecasting/ Time Series Terminologies
- Components of a Time Series
- Stationary Time Series
- AR Model
- MA Model
- ARMA and ARIMA Model
- Evaluation of TS
- Conclusion

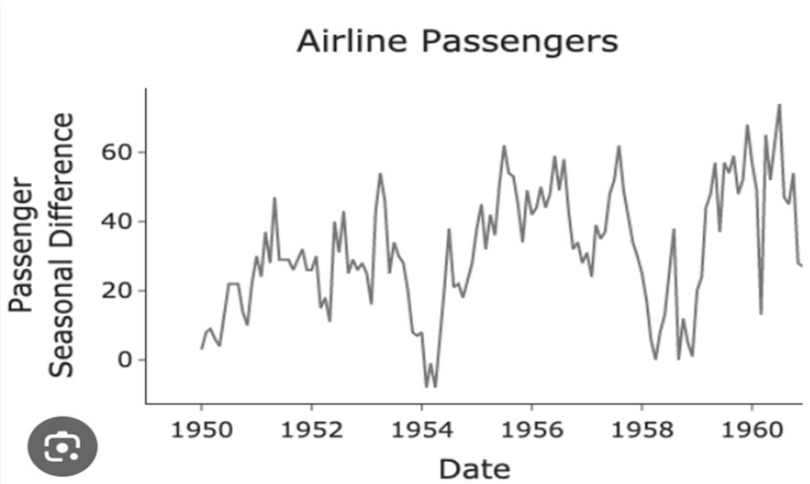
Some Time Series Graphs



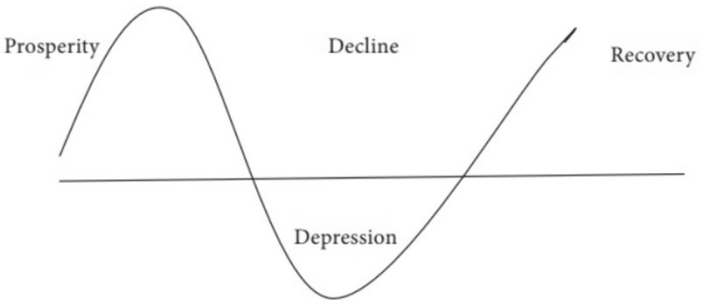
Some Time Series Graphs



Some Time Series Graphs



Financial Time Series Graphs



What is Forecasting?

- A forecast is a prediction of some future event or events.
- Forecasting is an important problem that spans many fields.
- Forecasting problems are often classified as short-term, medium-term, and long-term.
- Forecasting vs. Prediction

Forecasting Vs. Prediction

- Forecasting is more scientific
- Forecasting generally consider temporal dimension and try to estimate a target variable for future.
- Generally not bias as based data assessment
- Forecasting is more accurate

- Prediction is more intuitive
- Prediction encompasses any dimension of data.
- Subject to bias and can be subject to more error
- Forecasting is a kind of prediction but not all predictions are forecasts.

What is Forecasting?

- Short-term
  - Short-term forecasting problems involve predicting events only a few time periods (days, weeks, and months) into the future.
- Medium-term
  - Medium-term forecasts extend from 1 to 2 years into the future.
- Long-term.
  - Long-term forecasting problems can extend beyond that by many years.

Some basic Terminologies

- Point Forecast
- Interval Forecast
- Forecasting Horizon

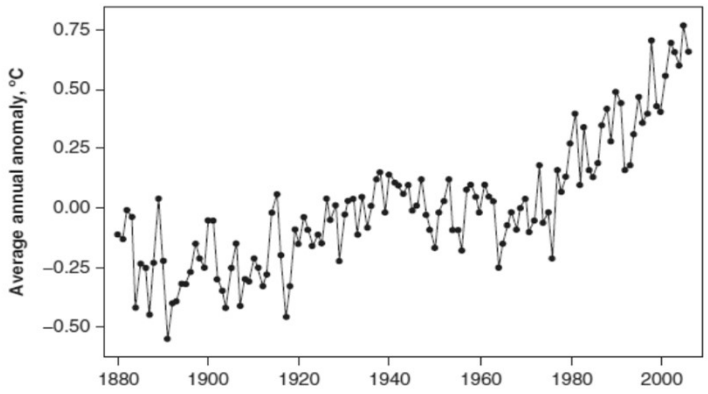
Time Series

- A time series is mathematically defined by the values $y_1, y_2, ..y_n$ of variable Y at time $t_1, t_2, ..t_n$.
- Symbolically, Y=F(t) for t →infinity
- Some examples of time series data
  - Annual rainfall in Pakistan
  - Enrollment of students in a university over a number of years.
  - Monthly sales of a given item.
  - Quarterly earning of a company.
  - Daily stock price of an stock item.

Time Series Plots

- A time series plot is a graph where some measure of time is the unit on the x-axis.
- In fact, we label the x-axis the time-axis.
- The y-axis is for the variable that is being measured.
- Data points are plotted and generally connected with straight lines, which allows for the analysis of the graph generated.

Time Series Plots



Global mean surface air temperature annual anomaly.
Source: NASA-GISS.

Components of a Time Series

- A typical Time Series(TS) has four types of movements, usually called components of a TS. (Systematic vs. Random)
- Systematic
  - Secular
  - Cyclic
  - Seasonal
- Random
  - Irregular movement (random force)

Analysis of Time Series

- Analysis of time series covers the description, measurement and isolation (generally mathematically) of the various components of movements presents in the series.
- It is generally believed by statisticians that these components Secular Trend (T), Cyclical Trend (C), Seasonal Trend (S) and Irregular Trend (I) follows:
  - Additive Law { Y = T + C + S + I}
  - Product Law { Y = T x C x S x I}

Model for Stationary Time Series

- Autoregressive Model AR(p)
- Moving Average Model MA(q)
- Autoregressive Moving Average Model ARMA(p,q)
- ARIMA

Model for Stationary Time Series

- Autoregressive Model AR(p)

$$Y_t = \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \cdots + \alpha_p Y_{t-p} + e_t$$

The model is based on weighting functions of past p observations.

Model parameters ($\alpha_1, \alpha_2, \cdots, \alpha_p$)

Observations:

- Tends to carry high past observation longer period
- No effective of errors in each observation
- Use of PACF for determining p lag values.

Model for Stationary Time Series

- Moving Average Model MA(q)

$$Y_t = \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + \cdots + \beta_q \varepsilon_{t-q} + e_t$$

The model is based on weighting functions of past q errors.

Model parameters ($\beta_1, \beta_2, \cdots, \beta_q$)

Observations:

- Tends to carry high shocks / low shocks longer period
- No effective of values from past in each observation
- Use of ACF for determining q lag values.

Model for Stationary Time Series

- Autoregressive Moving Average Model ARMA(p,q)

$Y_t = \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \cdots + \alpha_p Y_{t-p} + \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + \cdots + \beta_q \varepsilon_{t-q} + e_t$

The model is based on weighting functions of past p observations.

Model parameters $(\alpha_1, \alpha_2, \cdots, \alpha_p, \beta_1, \beta_2, \cdots, \beta_q)$

Observations:

- A complex model with large parameters only good for stationarity processes.
- A balance effect of past values and errors
- Use of ACF and PACF for determining q, p lag values.

Model for Non-Stationary Time Series

- ARIMA
  - This model is the basic interface for ARIMA-type models, including those with exogenous regressors and those with seasonal components. It also allows all specialized cases, including
  - Autoregressive models: AR(p)
  - Moving average models: MA(q)
  - Mixed autoregressive moving average models: ARMA(p, q)
  - Integration models: ARIMA(p, d, q)

Forecasting Model Evaluation

- This goodness-of-fit approach often uses the residuals and does not really reflect the capability of the forecasting technique to successfully predict future observations.
- The user of the forecasts is very concerned about the accuracy of future forecasts, not model goodness of fit, so it is important to evaluate this aspect of any recommended technique.
- Sometimes forecast accuracy is called “out-of-sample” forecast error, to distinguish it from the residuals that arise from a model-fitting process.

Forecasting Model Evaluation

- Mean Error (or Average Error)
  - This is the average of all the error

$$\mathrm{ME} = \frac{1}{n}\sum_{t=1}^{n} e_t(1),$$

  - The mean forecast error is an estimate of the expected value of forecast error, which we would hope to be zero (unbiased forecasts)
  - If the mean forecast error differs appreciably from zero, bias in the forecast is indicated.
  - If the mean forecast error drifts away from zero, it an indicator of changing the time series and model is not capturing the changes.

Forecasting Model Evaluation

- Mean Absolute Deviation (MAD)
  - This is an absolute deviation between estimated and actual forecast.

$$\mathrm{MAD} = \frac{1}{n}\sum_{t=1}^{n} |e_t(1)|,$$

  - It is a measure the variability in forecast errors. Obviously, we want the variability in forecast errors to be small.

Forecasting Model Evaluation

- Root Mean Squared Error (RMSE) - is the square root of the mean squared error between the predicted and actual values.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (Predicted_i - Actual_i)^2}{N}}$$

- Mean Absolute Percentage Error (MAPE) is the mean of all absolute percentage errors between the predicted and actual values.

$$MAPE = \frac{1}{n}\sum_{i=1}^{n} \frac{|y_i - \hat{y}_i|}{y_i}.100\%$$

Conclusion

- Time Series data is quite abundant and better models are required for many applications