Time Series Analysis

- Time series analysis is extensively used to forecast company sales, product demand, stock market trends, agricultural production etc.

- The fundamental idea for time series analysis is to decompose the original time series (sales, stock market trends, etc.) into several independent components.

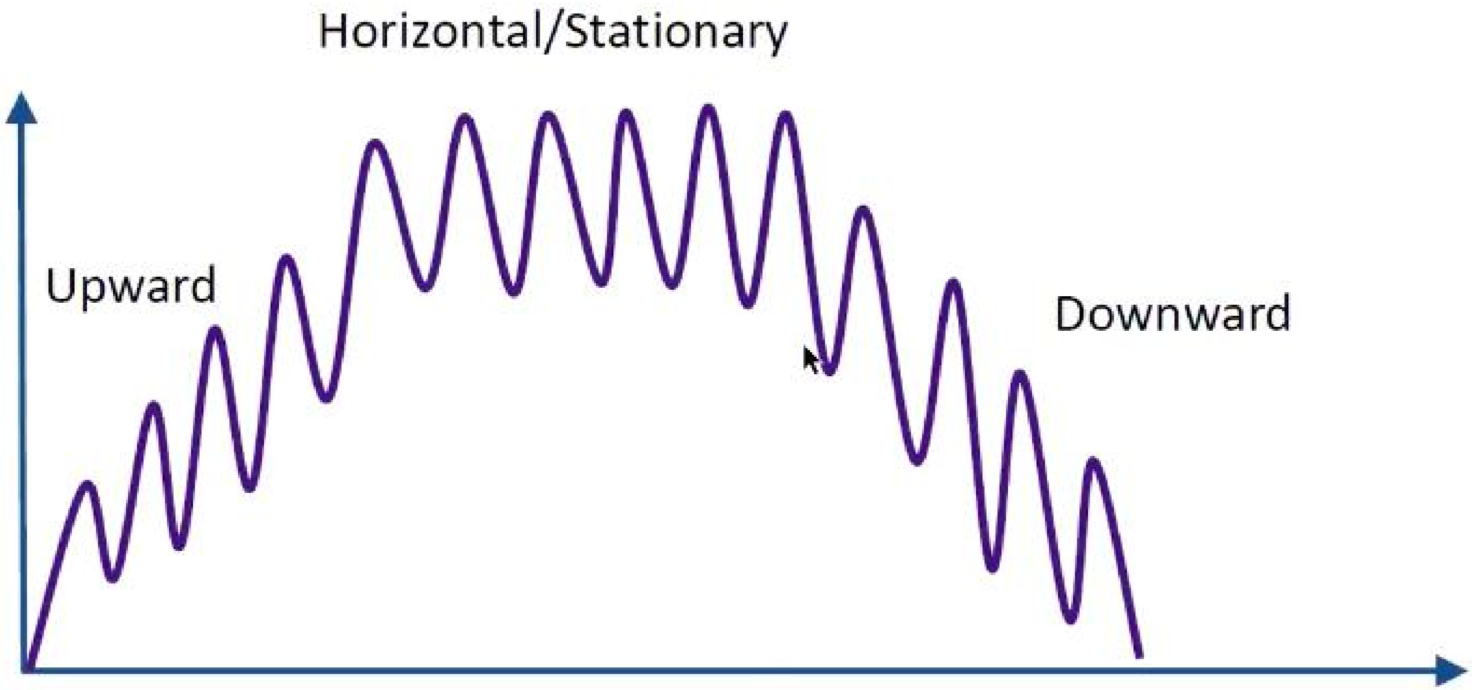
Time Series Analysis

- Typically, business time series are divided into the following four components:
- **Trend** – overall direction of the series i.e. upwards, downwards etc.
- **Seasonality** – monthly or quarterly patterns
- **Cycle** – long-term business cycles, they usually come after 5 or 7 years
- **Irregular remainder** – random noise left after extraction of all the components

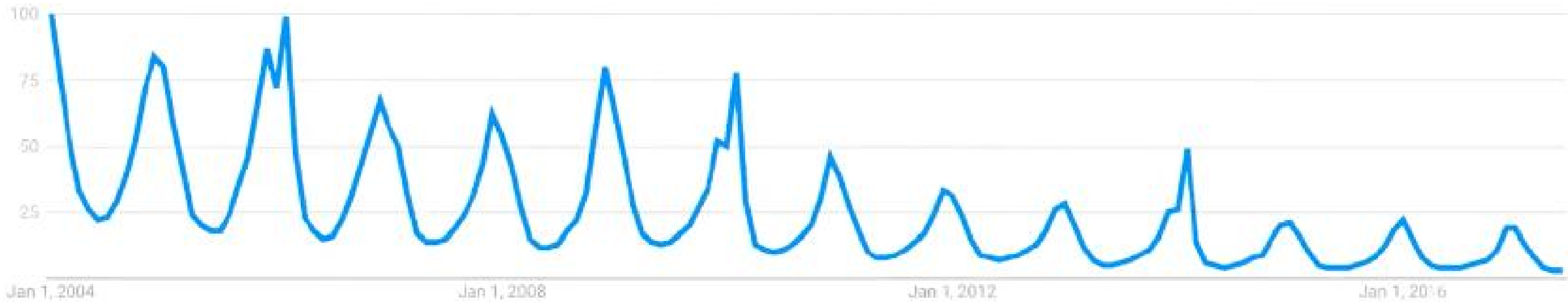
Time Series Analysis

- Interference of these components produces the final series.
- Why decomposing the original / actual time series into components?
- It is much easier to forecast the individual regular patterns produced through decomposition of time series than the actual series.

- Trends



- Seasonality - Repeating trends



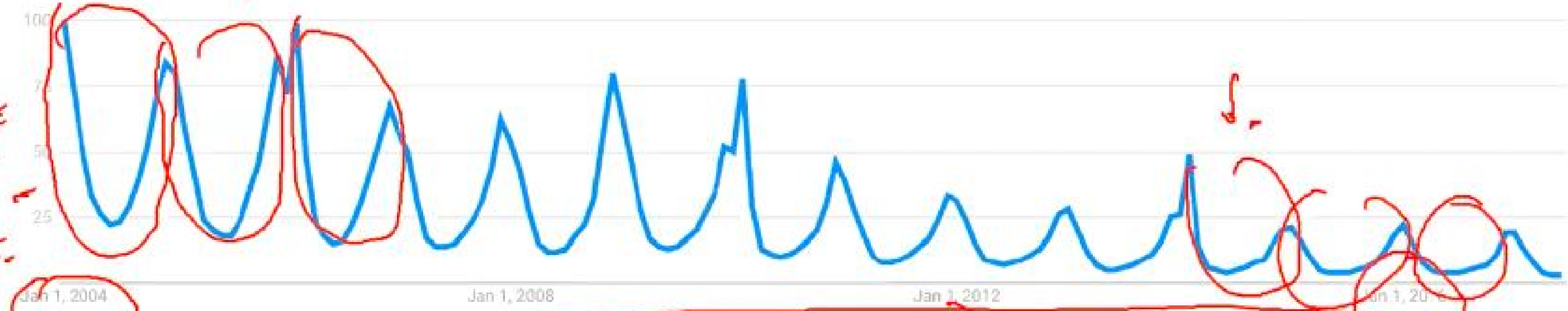
Google Trends - “Snowboarding”

winter summer rainy → year

sale

- Seasonality - Repeating trends

repeating repeating season



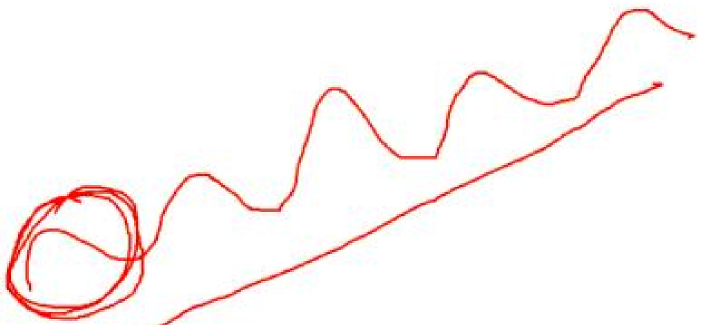
Google Trends - “Snowboarding”

Time Series Analysis

- Typically, business time series are divided into the following four components:
- **Trend** – overall direction of the series i.e. upwards, downwards etc.
- **Seasonality** – monthly or quarterly patterns olympic
- **Cycle** – long-term business cycles, they usually come after 5 or 7 years goverment 5 year worldcup → 4 year
- **Irregular remainder** – random noise left after extraction of all the components census → 10 year calculation

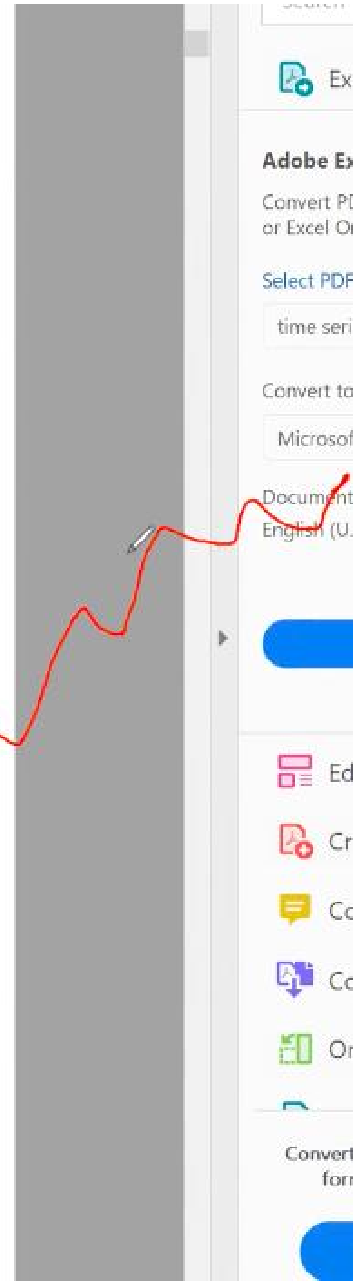
- Cyclical - Trends with no set repetition.

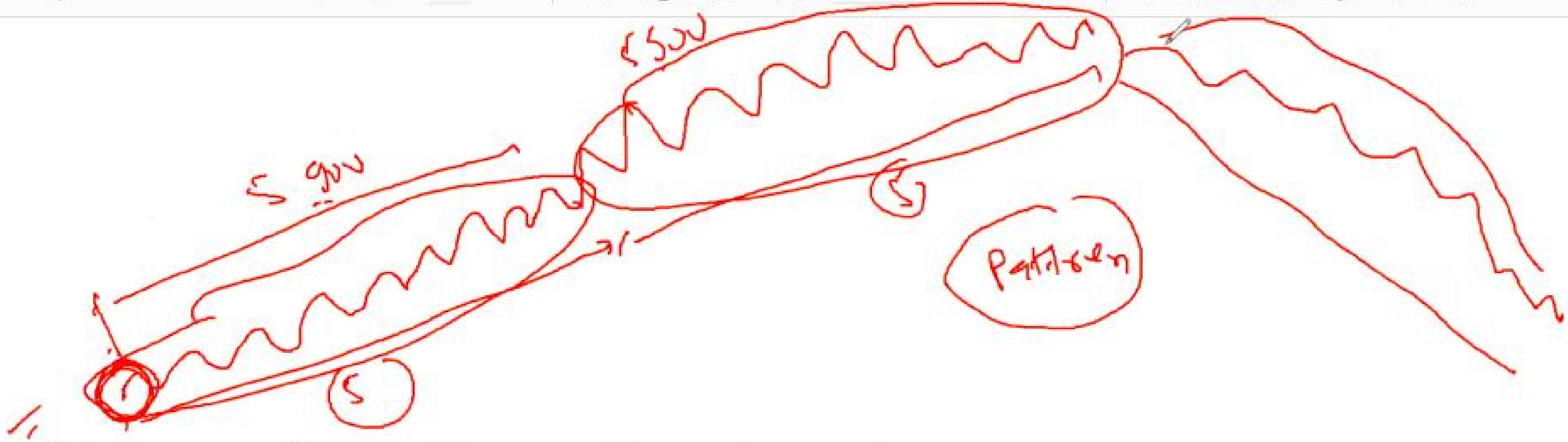
SP500



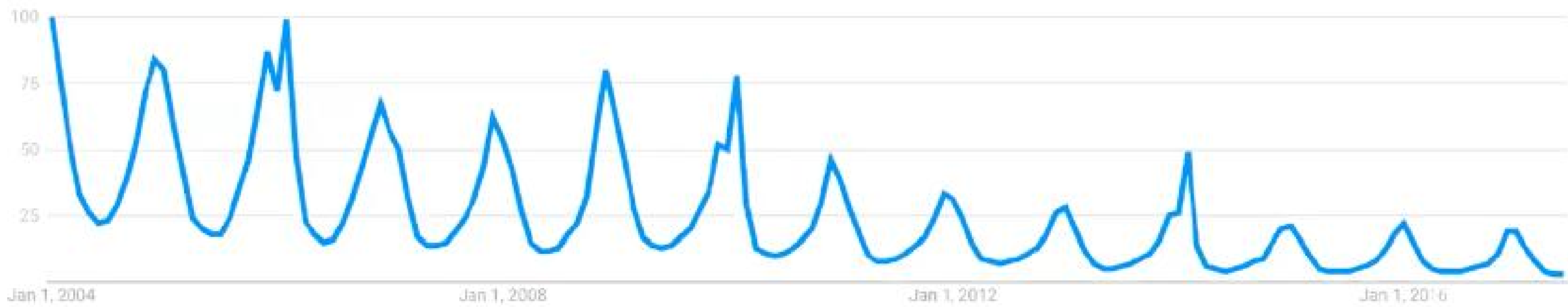
- Cyclical - Trends with no set repetition.

SP500





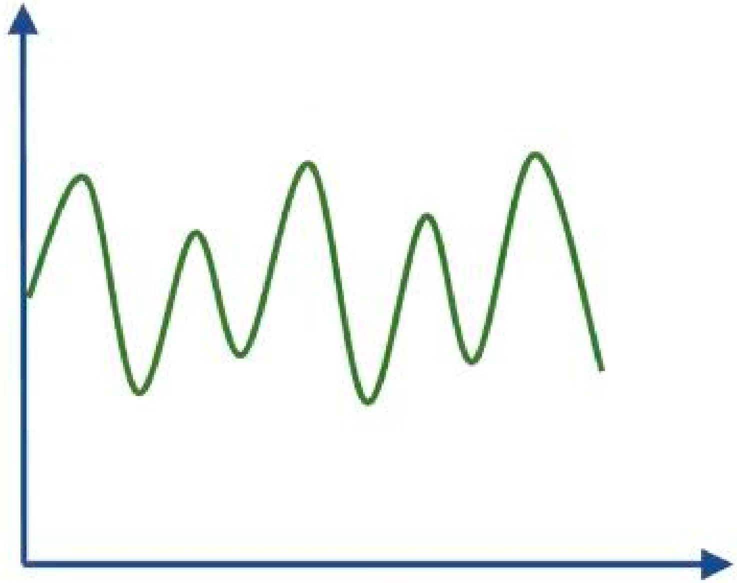
- Seasonality - Repeating trends

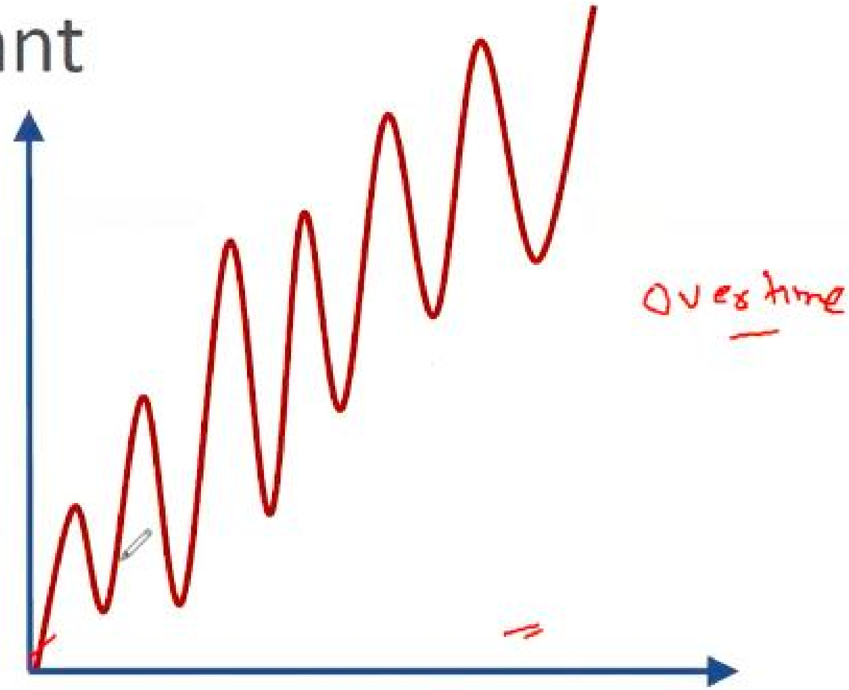


Google Trends - “Snowboarding”

- Stationary vs Non-Stationary Data
    - To effectively use ARIMA, we need to understand Stationarity in our data.
    - So what makes a data set Stationary?
        - A Stationary series has constant mean and variance over time.

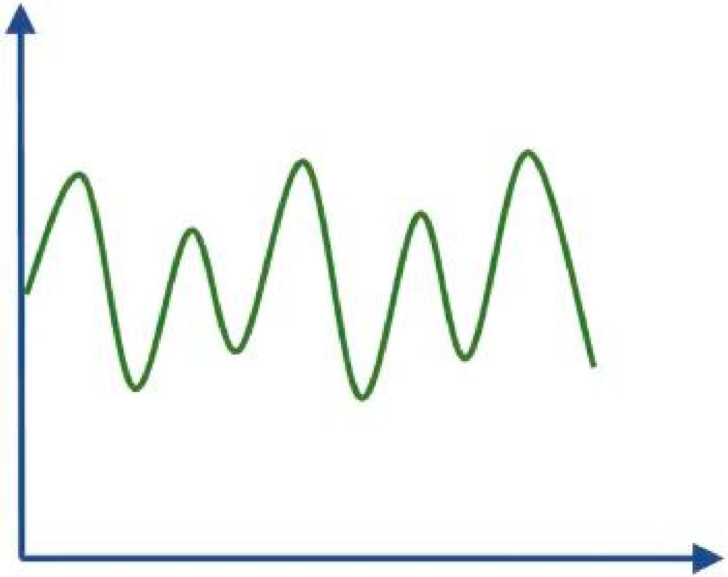
- Mean needs to be constant


Stationary

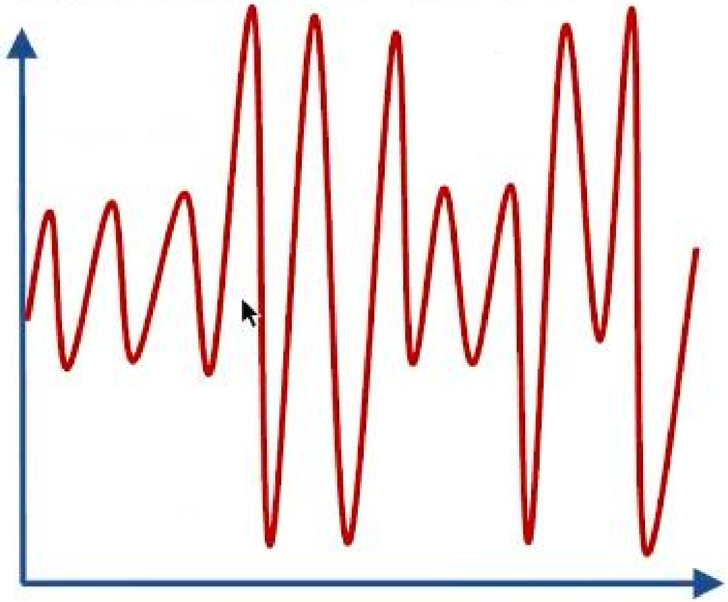


Non-Stationary

- Variance should not be a function of time



Stationary



Non-Stationary

- A Stationary data set will allow our model to predict that the mean and variance will be the same in future periods.

- There are also mathematical tests you can use to test for stationarity in your data.
- A common one is the Augmented Dickey–Fuller test

- If we’ve determined your data is not stationary (either visually or mathematically), we will then need to transform it to be stationary in order to evaluate it and what type of ARIMA terms you will use.

- One simple way to do this is through “differencing”.

Original Data

Time 1	10
Time 2	12
Time 3	8
Time 4	14
5	

First Difference

Time 1	NA
Time 2	2
Time 3	-4
Time 4	6
Time 5	-7

Second Difference

Time 1	NA
Time 2	NA
Time 3	-6
Time 4	10
Time 5	-13

Original Data

Time 1	10
Time 2	12
Time 3	8
Time 4	14
5	7

12-10

First Difference

Time 1	NA
Time 2	2
Time 3	-4
Time 4	6
Time 5	-7

-4-2

6+4

-7-6
= -13

Second Difference

Time 1	NA
Time 2	NA
Time 3	-6
Time 4	10
Time 5	-13

- You can continue differencing until you reach stationarity (which you can check visually and mathematically)
- Each differencing step comes at the cost of losing a row of data.

- For seasonal data, we can also difference by a season.
- For example, if we had monthly data with yearly seasonality, we could difference by a time unit of 12, instead of just 1.

- With our data now stationary it is time the p,d,q terms and how we choose them.
- A big part of this are AutoCorrelation Plots and Partial AutoCorrelation Plots.

Trend

- From the plots it is obvious that there is some kind of increasing trend in the series along with seasonal variation.

- Stationarity is a vital assumption we need to verify if our time series follows a stationary process or not.

Trend

- We can do by
  - Plots: review the time series plot of our data and visually check if there are any obvious trends or seasonality
  - Statistical tests: use statistical tests to check if the expectations of stationarity are met or have been violated.

Trend using MAs

- Moving averages over time
  - One way to identify a trend pattern is to use moving averages over a specific window of past observations.
  - This smoothens the curve by averaging adjacent values over the specified time horizon (window).

Seasonality

- People tend to go on vacation mainly during summer holidays.

- At some time periods during the year people tend to use aircrafts more frequently. We can check the hypothesis of a seasonal effect

Noise

- To understand the underlying pattern in the number of international airline passengers, we assume a multiplicative time series decomposition model

- Purpose is to understand underlying patterns in temporal data to use in more sophisticated analysis like Holt-Winters seasonal method or ARIMA.

Noise

- Noise - is the residual series left after removing the trend and seasonality components

Stationarize a Time series

- Before models forecasting can be applied, the series must be transformed into a stationary time series.

- The Augmented-Dickey Fuller Test can be used to test whether or not a given time series is stationary.

Stationarize a Time series

- If the test statistic is smaller than the critical value, the hypothesis is rejected, the series would be stationary, and no further transformations of the data would be required.

Residuals Serial Correlation

- When the residuals (errors) in a time series are correlated with each other it is said to exhibit serial correlation.
- Autocorrelation is a better measurement for the dependency structure, because the autocovariacne will be affected by the underlying units of measurement for the observation.