Time Series Analysis

Basics of Time Series Analysis

Nicoleta Serban, Ph.D.
Professor
Stewart School of Industrial and Systems Engineering

Basic Definition and Examples

Georgia Tech

About This Lesson

Georgia Tech

Formal Definition

A *stochastic process* is a collection of random variables $\{X_t, t \in T\}$, defined on a probability space (Ω, F, P).

A *time series* is a stochastic process in which T is a set of time points, usually

$$T = \{0, \pm 1, \pm 2, \ldots\}, \{1, 2, 3, \ldots\}, [0, \infty), \text{or } (-\infty, \infty)$$

Note: The term “time series” is also used to refer to the realization of such a process (observed time series).

Example: Time Series

- Monthly sales of Australian red wine
- Monthly accidental deaths in the U.S.
- Daily Average Temperature from La Harpe station in Hancock County, Illinois
- Daily stock price of IBM stock
- US monthly interest rates
- US yearly GDP
- 1-minute intraday S&P500 return

Time Series: Characteristics

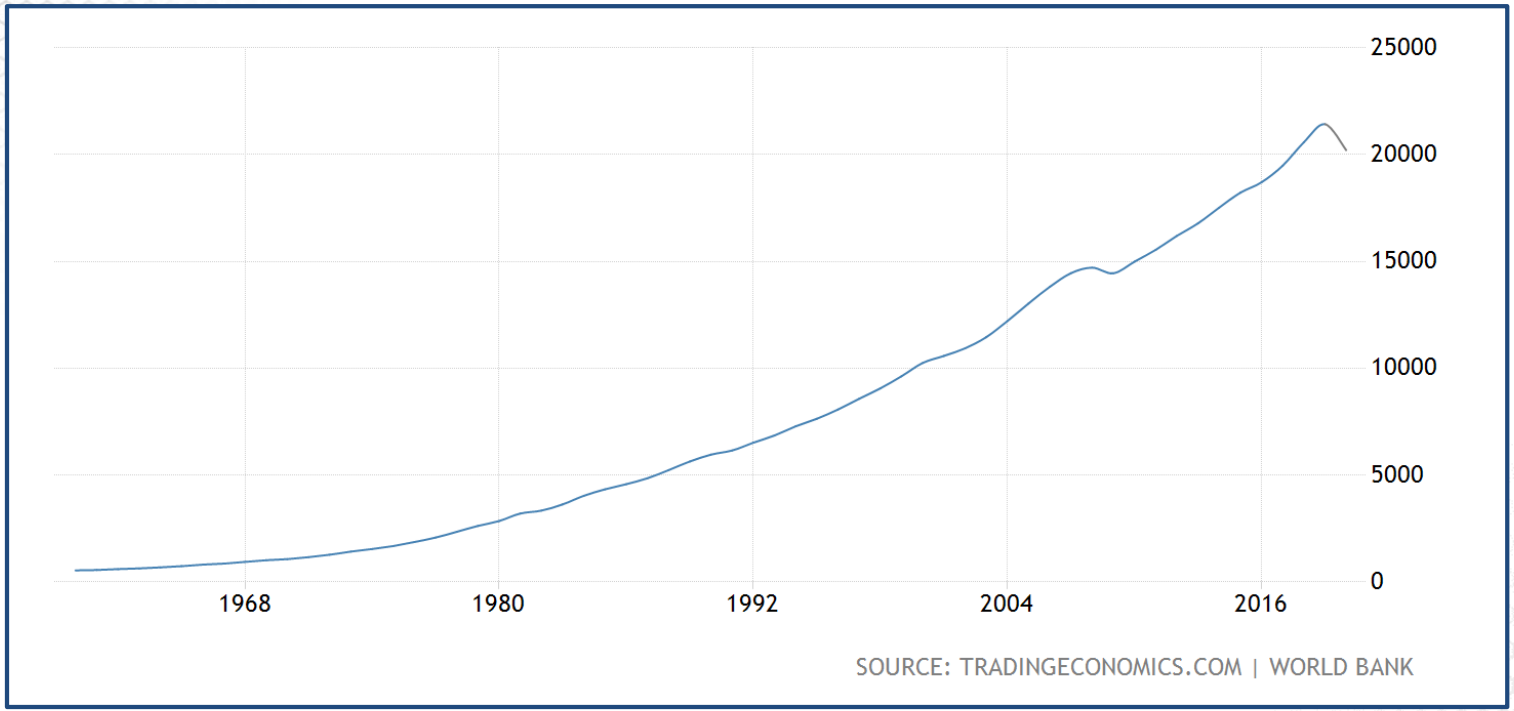
- **Trend**: long-term increase or decrease in the data over time
- **Seasonality**: influenced by seasonal factors;
- **Periodicity**: exact repetition in regular pattern
- **Cyclical trend**: rises and falls, not necessarily of a fixed/exact period
  - Seasonality vs periodicity: repeating over an exact period and modeled using seasonal models; often used interchangeably
  - Seasonality: seasonal factor ~ period of a one year
  - Periodicity: frequency of collecting measurements
  - Cyclical patterns: a dominant period, but could be very different from 'seasonality' (e.g. year) or periodicity (e.g. frequency)

Time Series: Characteristics (cont'd)

- **Varying variance** with time
  - Constant variance (also called as *homoskedasticity*)
  - Nonconstant variance (also called as *heteroskedasticity*)

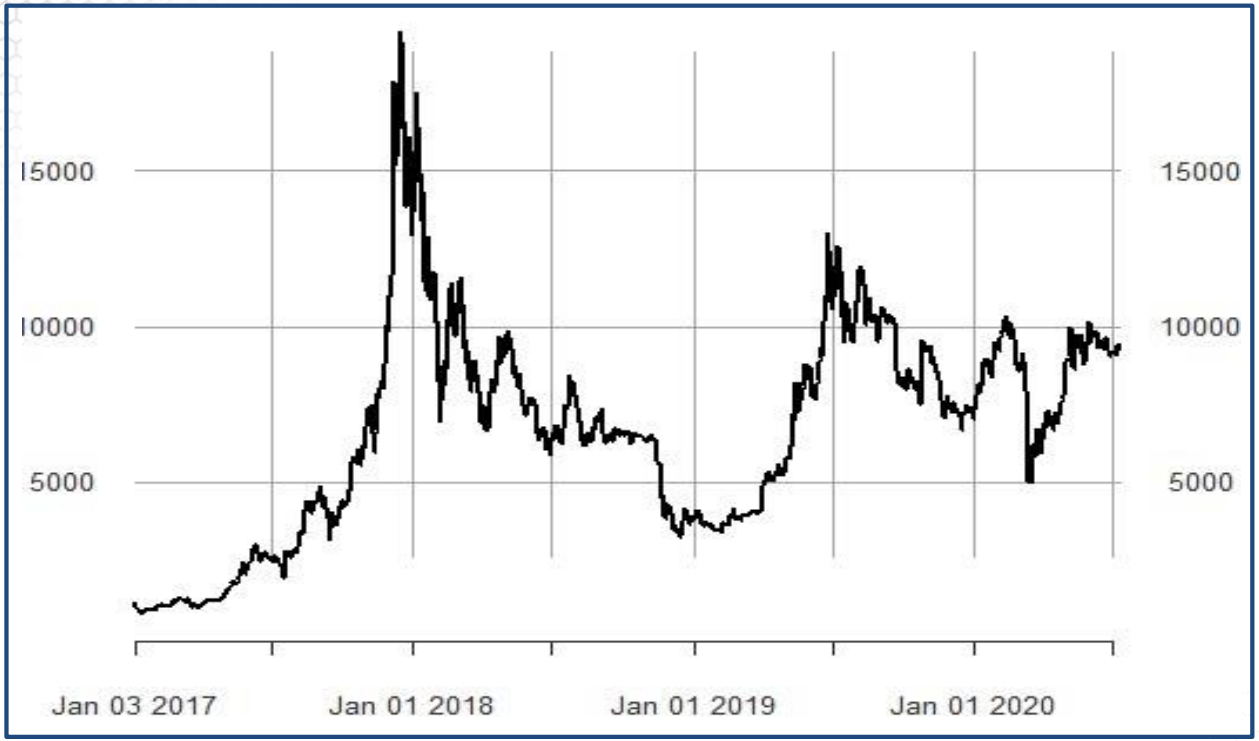
- **Dependence**: positive (successive observations are similar) or negative (successive observations are dissimilar)
  - Serial correlation commonly modeled using time series analysis
  - Dependence between time series in multivariate analysis

Example: GDP in the US



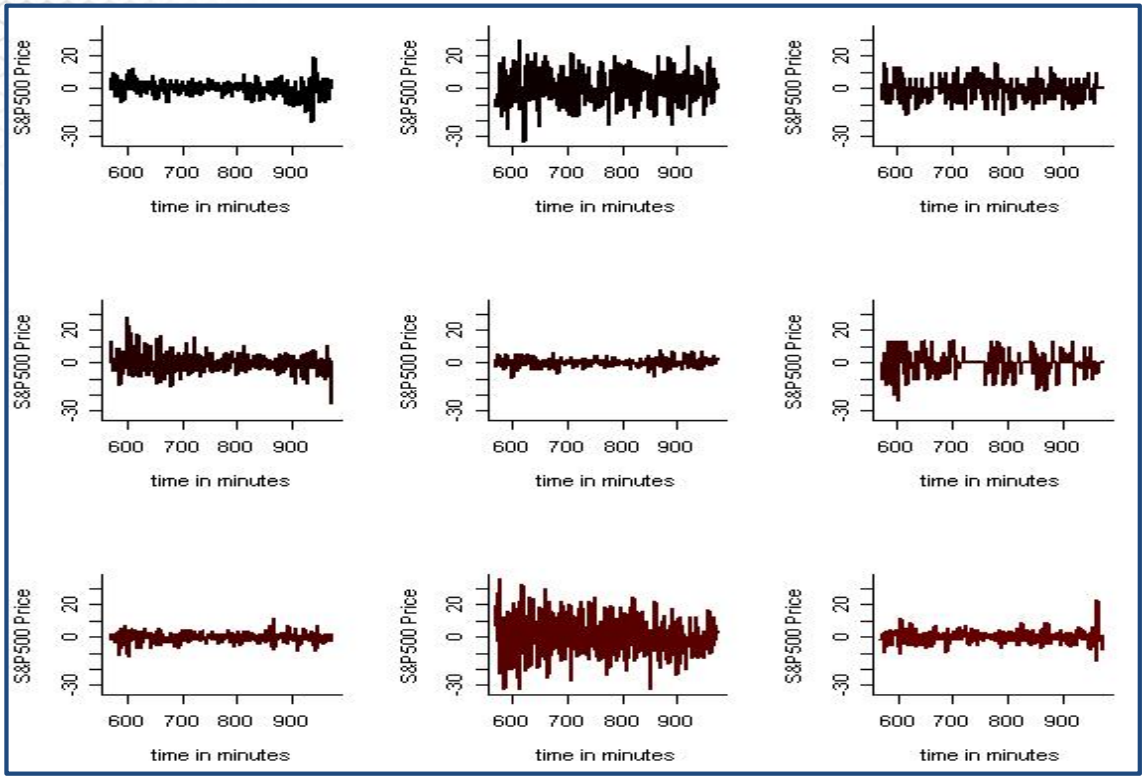
Georgia Tech

Example: Bitcoin Price



Georgia Tech

Example: S&P500 Intraday



Is Time Series Analysis Necessary?

Time Series ⇒ Dependence

- Data redundancy: number of degrees of freedom is smaller than T (T is the number of observations)
- Data sampling: $Y_t, t = 1, .., T$ concentrated about a small part of the probability space

Ignoring dependence leads to

- Inefficient estimates of regression parameters
- Poor predictions
- Standard errors unrealistically small (too narrow CI ⇒ improper inferences)

Time Series: Objectives

Description

- Plot the data and obtain simple descriptive measures of the main properties of the series.

Explanation

- Find a model to describe the time dependence in data.

Forecasting

- Given a finite sample from the series (observations), forecast the next value or the next several values.

Control/Tuning

- After forecasting, adjust various control/tune parameters.

Time Series Analysis: Approaches

Time domain approach

- Assume that correlation between adjacent points in time can be explained through dependence of the current value on past values.

Frequency domain approach

- Characteristics of interest relate to periodic (systematic) sinusoidal variations in the data, often caused by biological, physical, or environmental phenomena.

Summary

Georgia Tech