Introduction to Time Series and Forecasting

APM1215 : TIME SERIES ANALYSIS

Forecast

A forecast is a prediction of some future event or events. As suggested by Neils Bohr, making good predictions is not always easy.

Famously "bad" forecasts include the following from the book Bad Predictions:
+ "The population is constant in size and will remain so right up to the end of mankind." L'Encyclopedie, 1756.
"1930 will be a splendid employment year." U.S. Department of Labor, New Year's Forecast in 1929, just before the market crash on October 29.
"Computers are multiplying at a rapid rate. By the turn of the century there will be 220,000 in the U.S." Wall Street Journal, 1966.

Classifications of forecasting problems

+ Short-term forecasting problems involve predicting events only a few time periods (days, weeks, months) into the future.

+ Medium-term forecasts extend from one to two years into the future, and

+ Long-term forecasting problems can extend beyond that by many years.

Remarks

+ Short- and medium-term forecasts are required for activities that range from operations management to budgeting and selecting new research and development projects.

+ Long-term forecasts impact issues such as strategic planning.

+ Short- and medium-term forecasting is typically based on identifying, modeling, and extrapolating the patterns found in historical data. Because these historical data usually exhibit inertia and do not change dramatically very quickly, statistical methods are very useful for short- and medium-term forecasting.
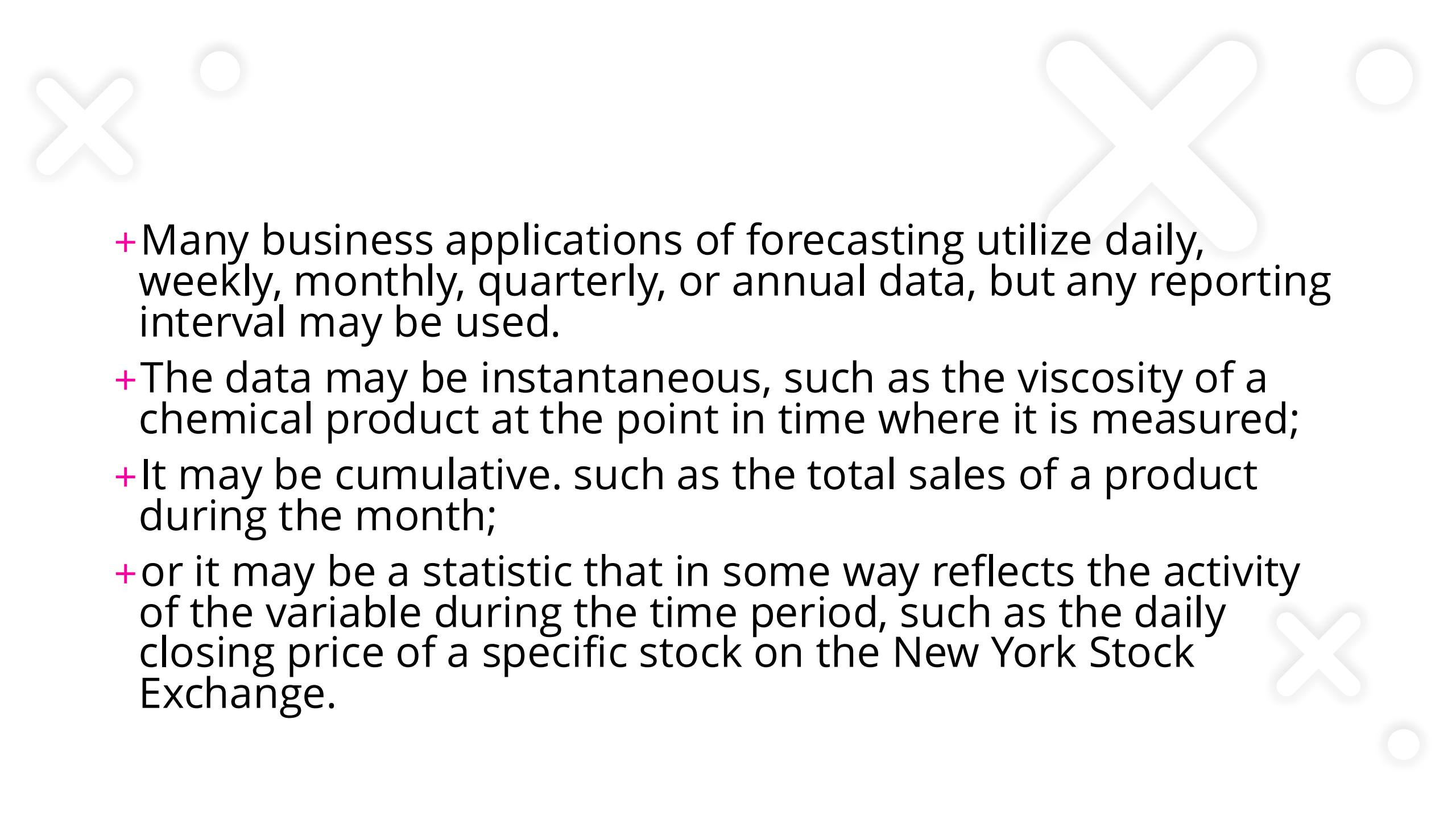
Time Series

A time series is a time-oriented or chronological sequence of observations on a variable of interest.
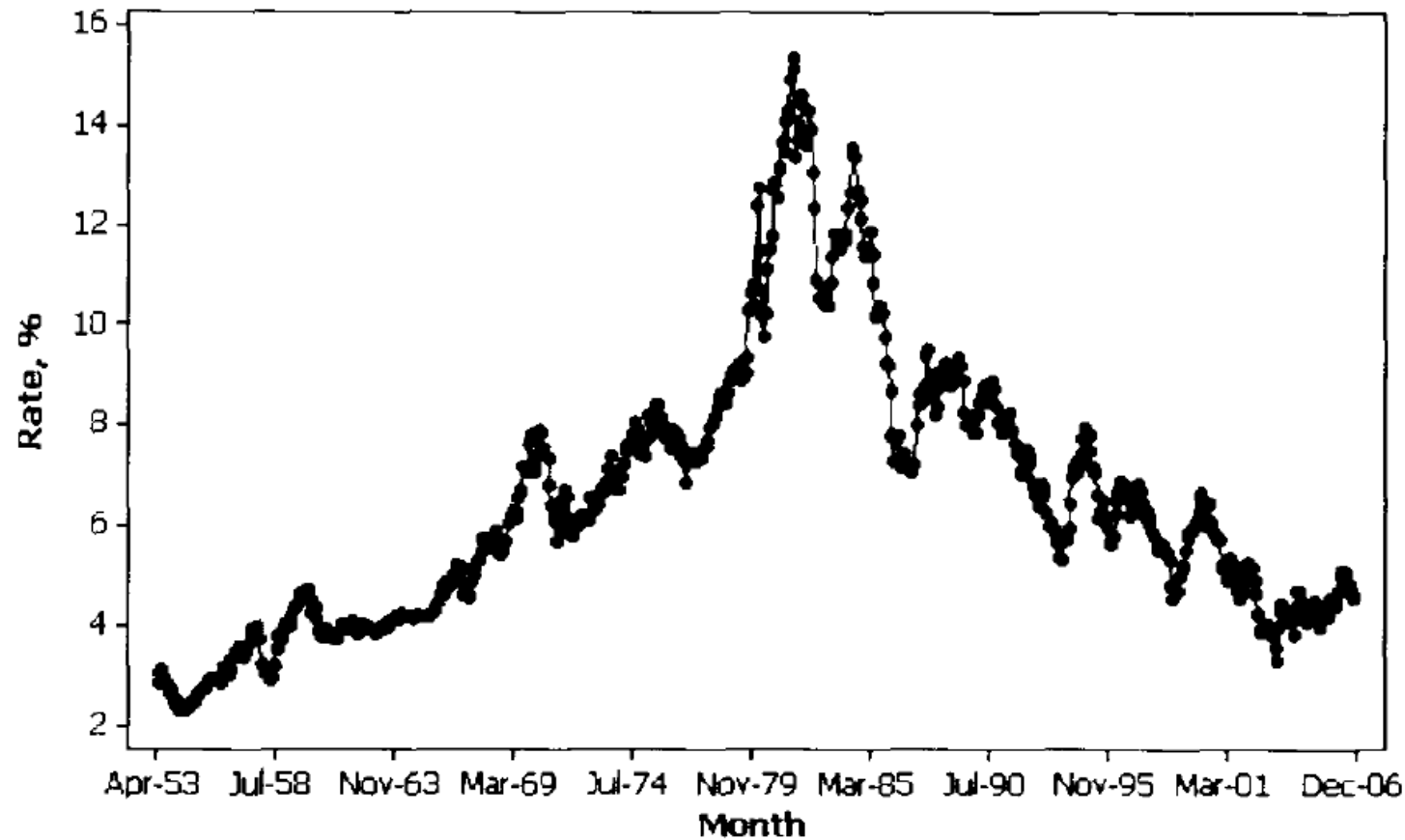
market yield on U.S. Treasury Securities at 10-year constant maturity from April 1953 through December. This graph is called a **time series plot**. The rate variable is collected at equally spaced time periods.

+ Many business applications of forecasting utilize daily, weekly, monthly, quarterly, or annual data, but any reporting interval may be used.

+ The data may be instantaneous, such as the viscosity of a chemical product at the point in time where it is measured;

+ It may be cumulative. such as the total sales of a product during the month;

+ or it may be a statistic that in some way reflects the activity of the variable during the time period, such as the daily closing price of a specific stock on the New York Stock Exchange.

Applications

Operations Management

Business organizations routinely use forecasts of product sales or demand for services in order to schedule production. Control inventories, manage the supply chain, determine staffing requirements. And plan capacity. Forecasts may also be used to determine the mix of products or services to be offered and the locations at which products are to be produced.

Marketing

Forecasting is important in many marketing decisions. Forecasts of sales response to advertising expenditures, new promotions, or changes in pricing policies enable businesses to evaluate their effectiveness, determine whether goals are being met, and make adjustments.

Finance and Risk Management

Investors in financial assets are interested in forecasting the returns from their investments. These assets include but are not limited to stocks, bonds, and commodities; other investment decisions can be made relative to forecasts of interest rates, options, and currency exchange rates.

Financial risk management requires forecasts of the volatility of asset returns so that the risks associated with investment portfolios can be evaluated and insured, and so that financial derivatives can be properly priced.

Economics

Governments, financial institutions, and policy organizations require forecasts of major economic variables, such as gross domestic product, population growth, unemployment, interest rates, inflation, job growth, production, and consumption.

These forecasts are an integral part of the guidance behind monetary and fiscal policy and budgeting plans and decisions made by governments. They are also instrumental in the strategic planning decisions made by business organizations and financial institutions.

Industrial Process Control

Forecasts of the future values of critical quality characteristics of a production process can help determine when important controllable variables in the process should be changed, or if the process should be shut down and overhauled.

Feedback and feedforward control schemes are widely used in monitoring and adjustment of industrial processes, and predictions of the process output are an integral part of these schemes.

Demography

Forecasts of population by country and regions are made routinely, often stratified by variables such as gender, age, and race.

Demographers also forecast births, deaths, and migration patterns of populations.

Governments use these forecasts for planning policy and social service actions, such as spending on health care, retirement programs, and antipoverty programs.

Many businesses use forecasts of populations by age groups to make strategic plans regarding developing new product lines or the types of services that will be offered.
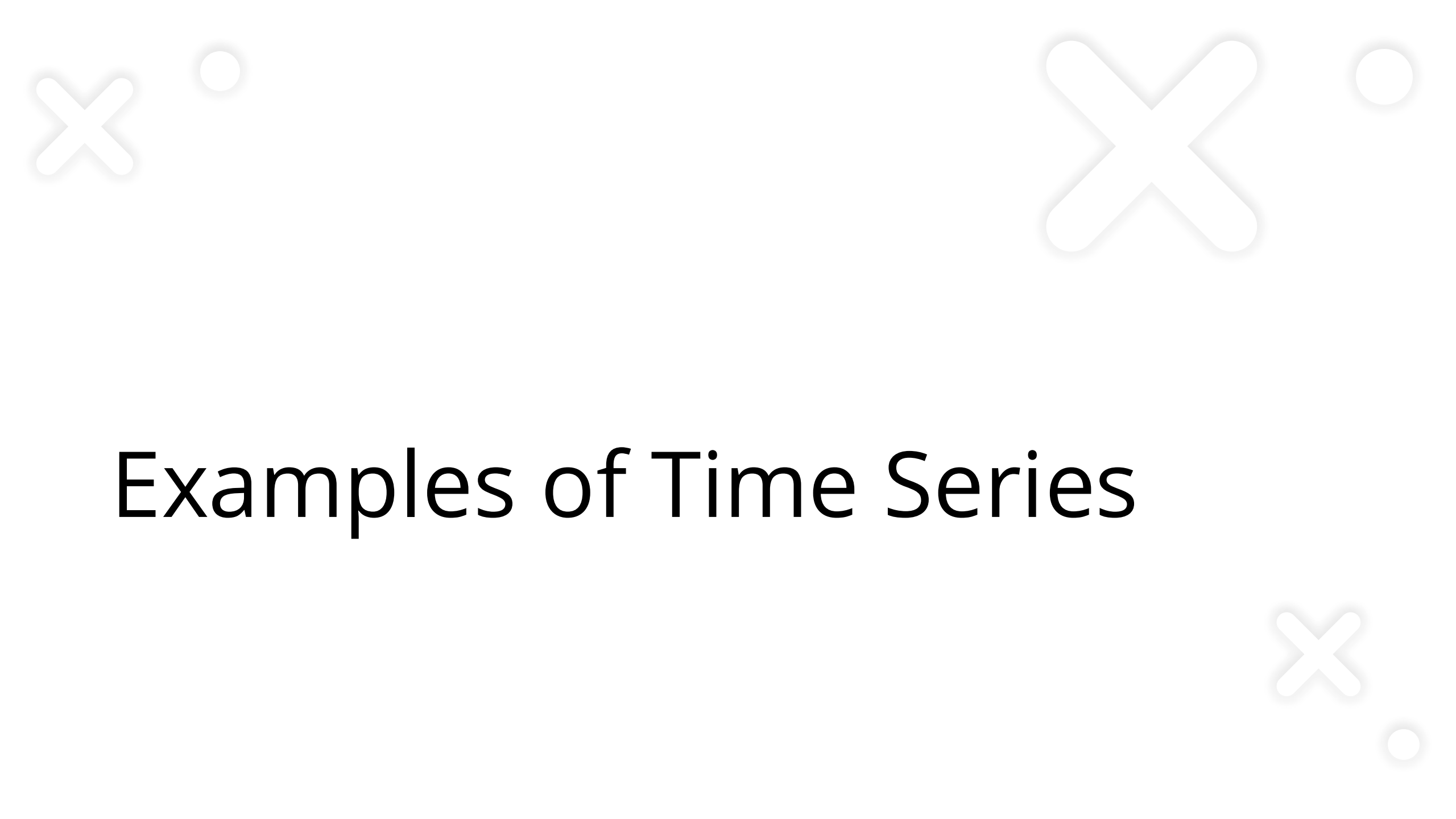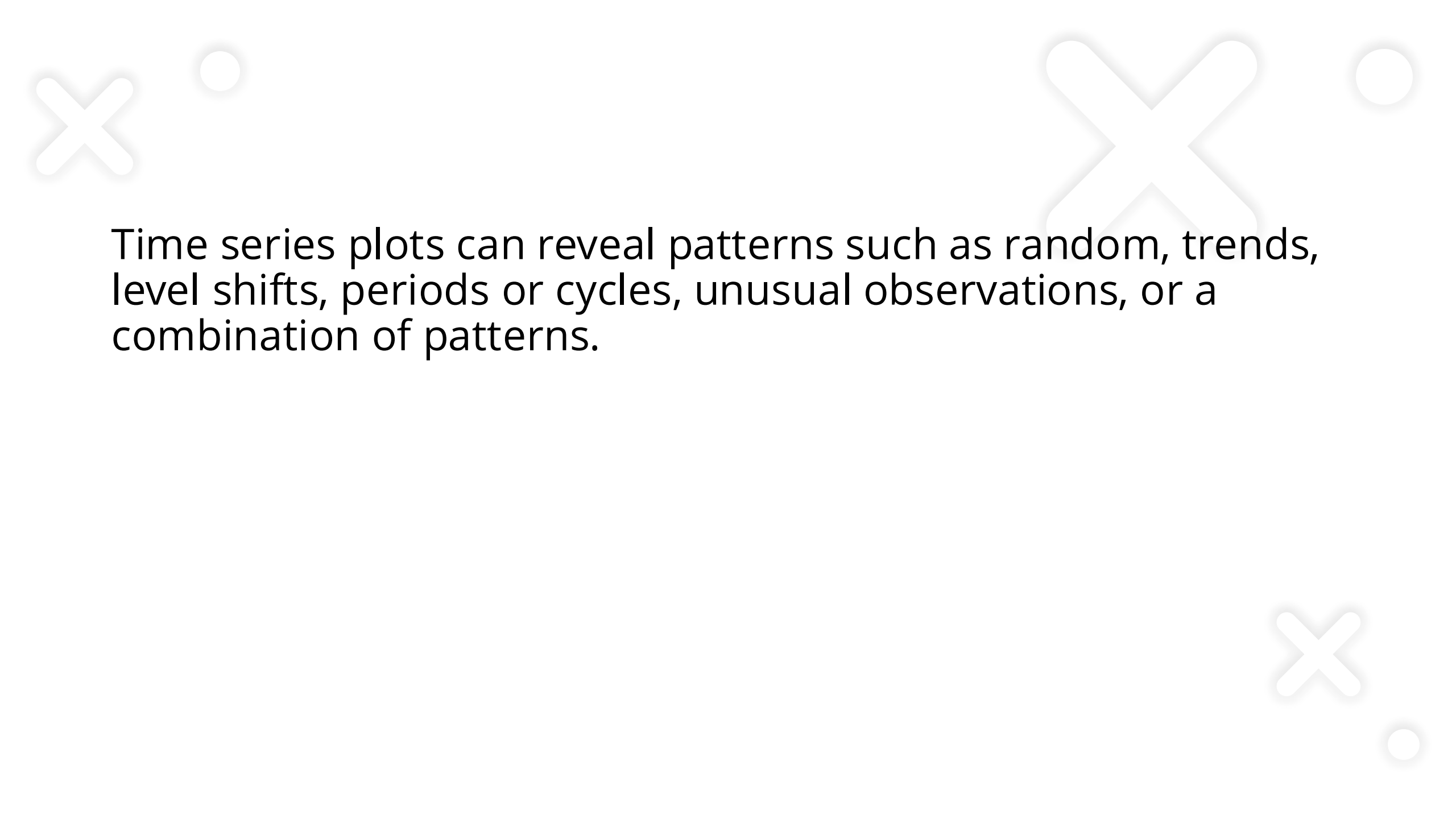
Terms

+ **Quantitative forecasting techniques** make formal use of historical data and a forecasting model. The forecasting model is use to extrapolate past and current behavior into the future.

- o **Regression models** – uses relationships between the variable of interest and one or more predictor variables. Sometimes used for **causal forecasting models.**
- o **Smoothing models** - employ a simple function of previous observations to provide a forecast of the variable of interest.
- o **General time series models** - employ the statistical properties of the historical data to specify a formal model and then estimate the unknown parameters of this model (usually) by least squares.

+ **Point estimate or point forecast** – single number that represents our best estimate of the future value of the variable of interest.

+ **Forecast error** – usually accompany a forecast to estimate how large an error might be experienced.

+ **Prediction interval** – range of values for the future observations.

+ **Forecast horizon (forecast lead time)** – number of future periods for which the forecasts must be produced.

+ **Forecast interval** – frequency with which new forecasts are prepared.

Examples of Time Series

Time series plots can reveal patterns such as random, trends, level shifts, periods or cycles, unusual observations, or a combination of patterns.
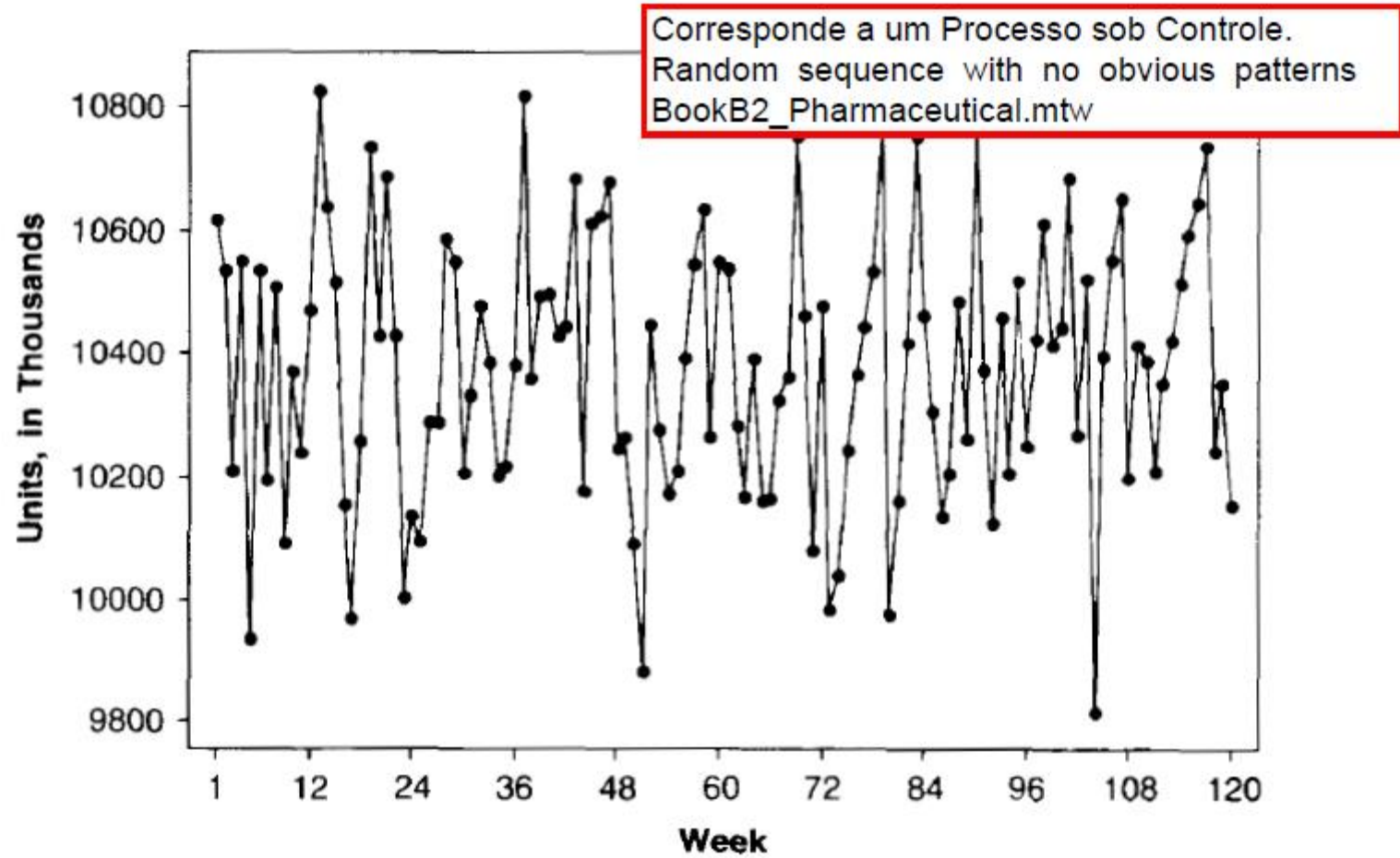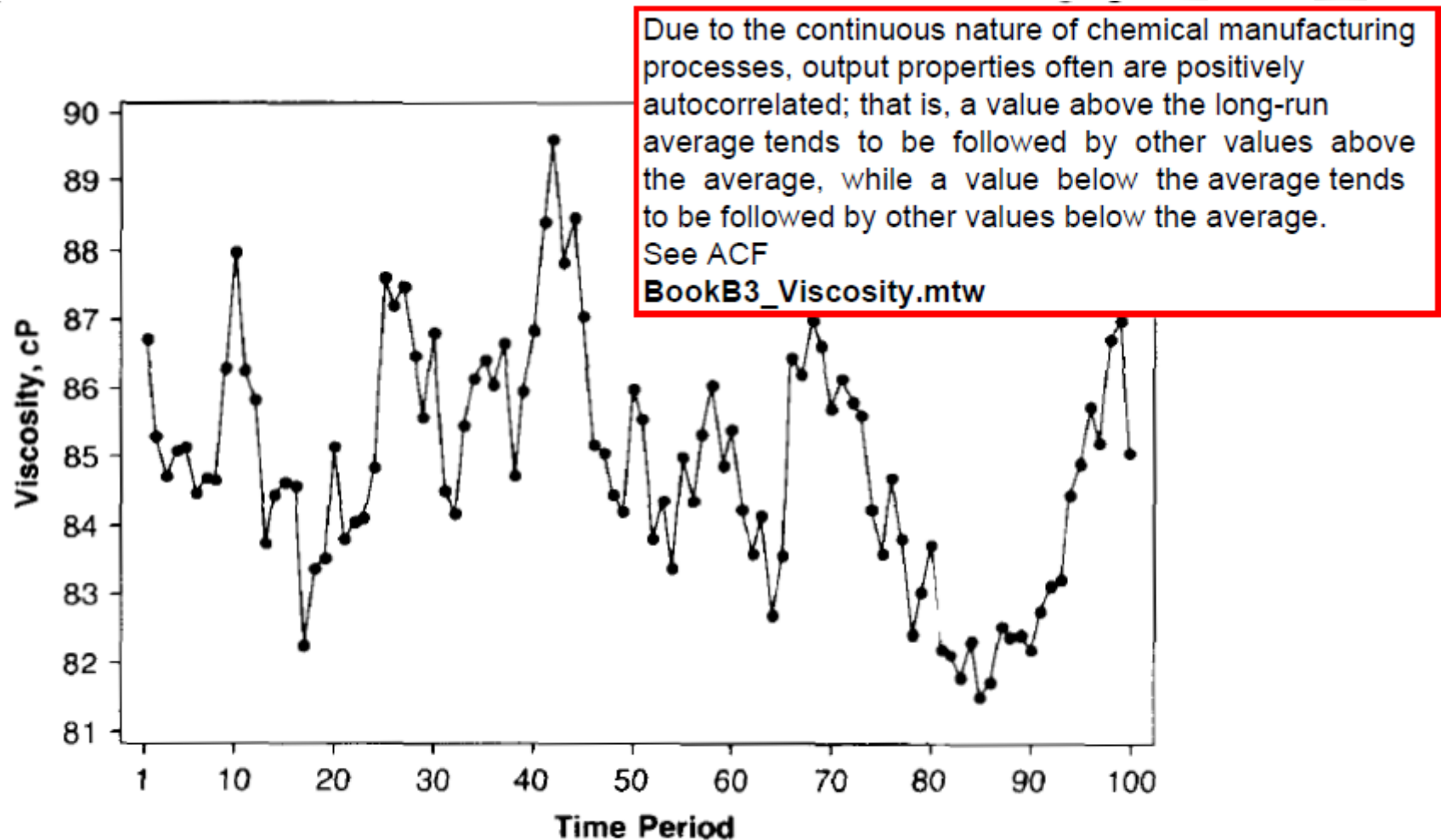
FIGURE 1.2 Pharmaceutical product sales.

Due to the continuous nature of chemical manufacturing processes, output properties often are positively autocorrelated; that is, a value above the long-run average tends to be followed by other values above the average, while a value below the average tends to be followed by other values below the average.
See ACF
BookB3_Viscosity.mtw

FIGURE 1.3 Chemical process viscosity readings.

The linear trend has a constant positive slope with random, year-to-year variation.
BookB4_Gorgonzola.mtw
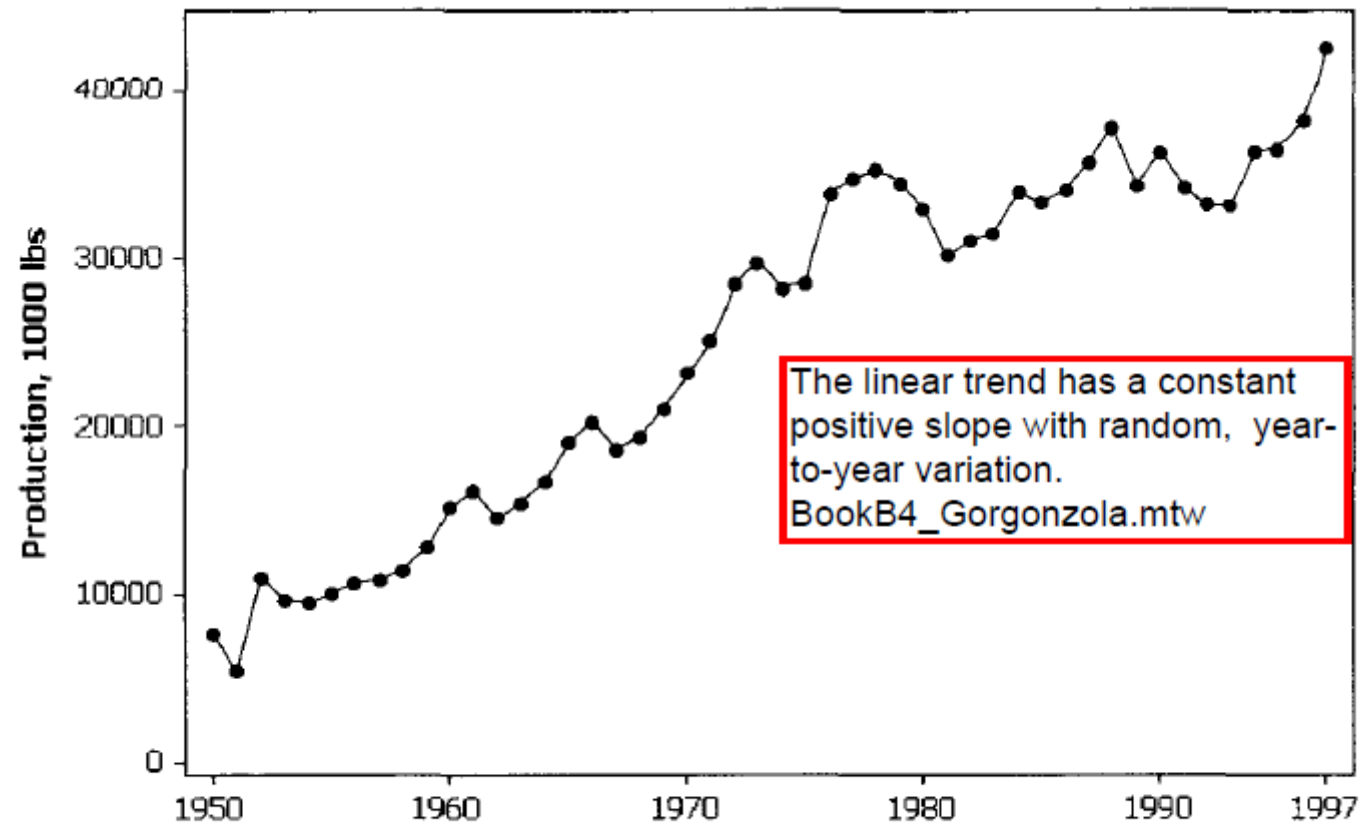
FIGURE 1.4 The U.S. annual production of blue and gorgonzola cheeses. (Source: USDA–NASS.)

The plot reveals overall increasing trend, with a distinct cyclic pattern that is repeated within each year.
Seazonal é geralmente igual a ciclic. Em alguns textos, ciclo/tendência são tratados juntos.
BookB5_BeverageShipment.mtw
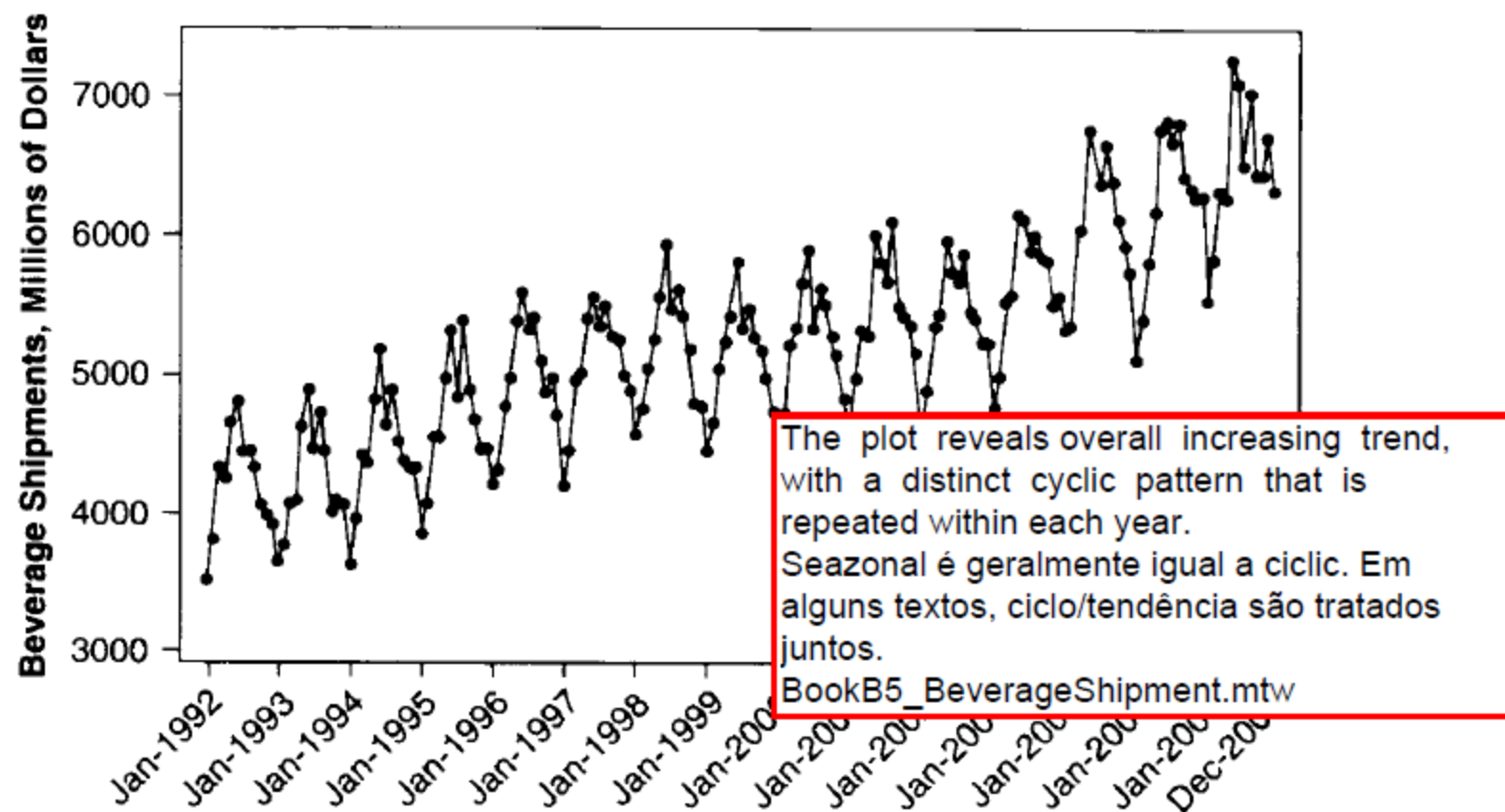
FIGURE 1.5 The U.S. beverage manufacturer monthly product shipments, unadjusted. (Source: U.S. Census Bureau.)

The plot of the annual mean anomaly in global surface air temperature shows an increasing trend since 1880.
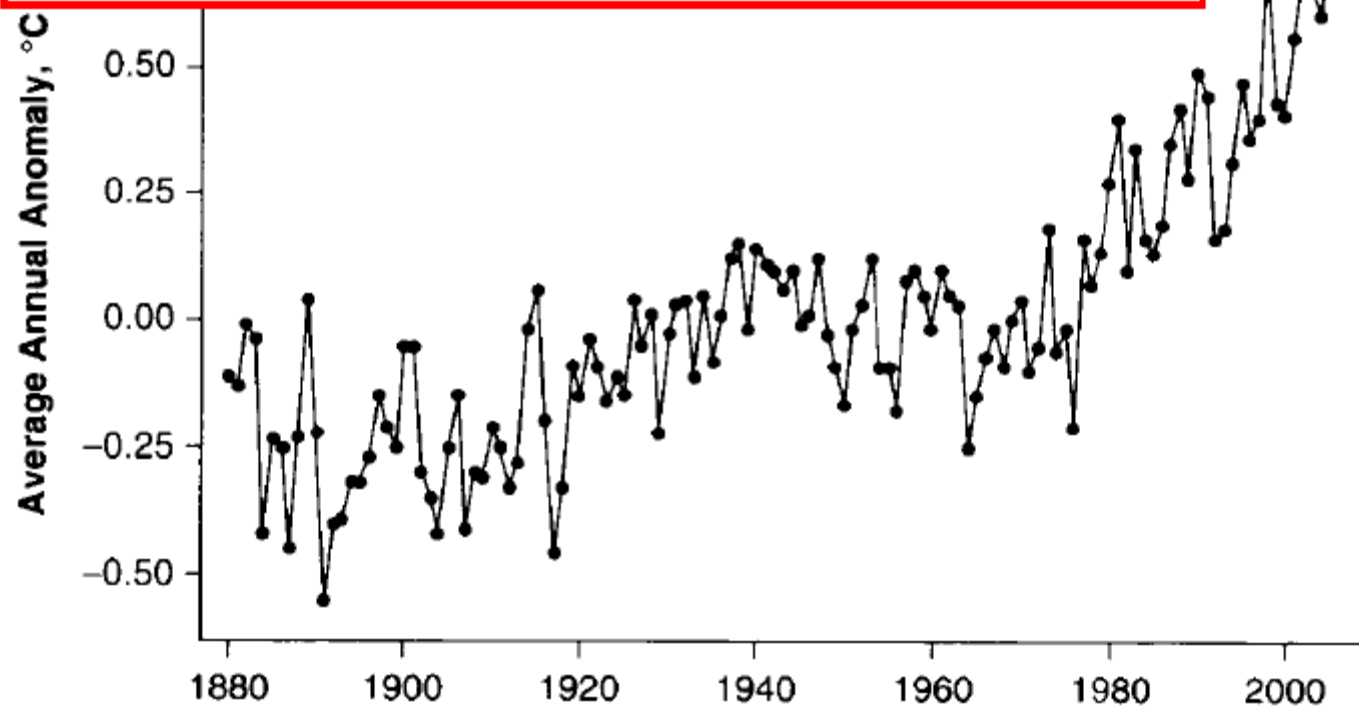BookB6_AirTempAnomalyCO2.mtw

FIGURE 1.6 Global mean surface air temperature annual anomaly. (Source: NASA–GISS.)

Business data such as stock prices and interest rates often exhibit nonstationary behavior; that is, the time series has no natural mean.
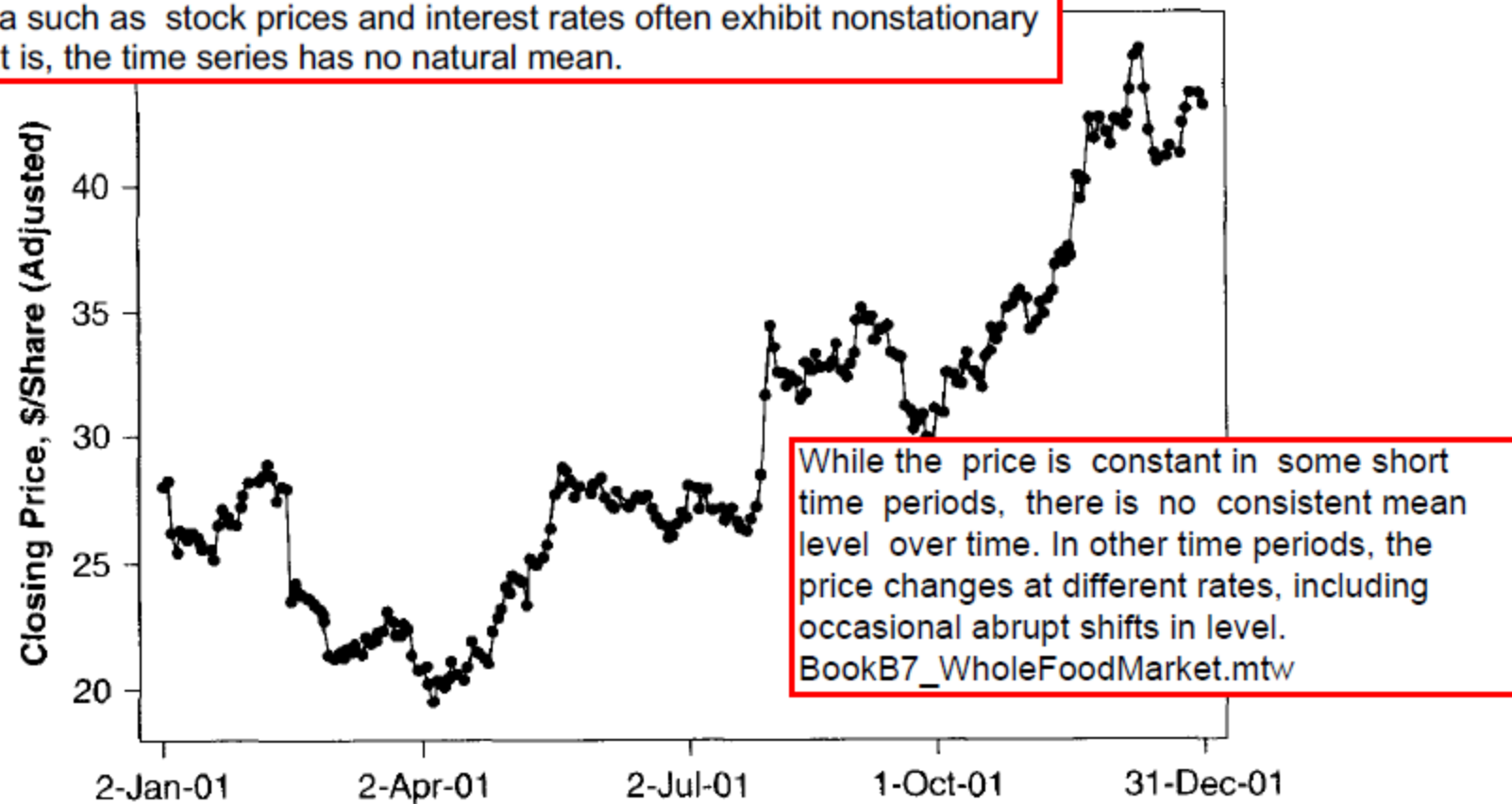
While the price is constant in some short time periods, there is no consistent mean level over time. In other time periods, the price changes at different rates, including occasional abrupt shifts in level. BookB7_WholeFoodMarket.mtw

FIGURE 1.7 Whole Foods Market stock price, daily closing adjusted for splits.

10

The plot exhibits a mixture of patterns. There is a distinct cyclic pattern within a year; January, February, and March generally have the highest unemployment rates. The overall level is also changing, from a gradual decrease, to a steep increase, followed by a gradual decrease.
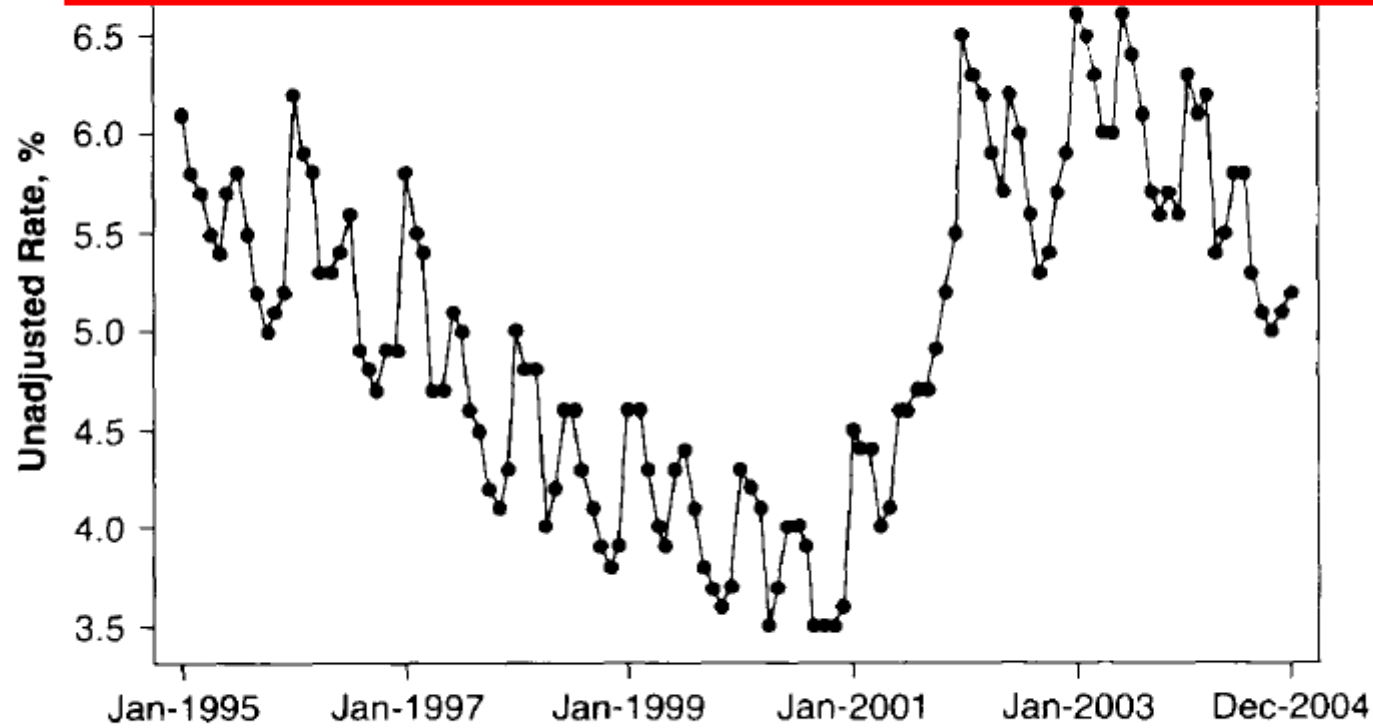BookB8_UnemploymentRate.mtw

FIGURE 1.8 Monthly unemployment rate—full-time labor force, unadjusted. (Source: U.S. Department of Labor–BLS.)
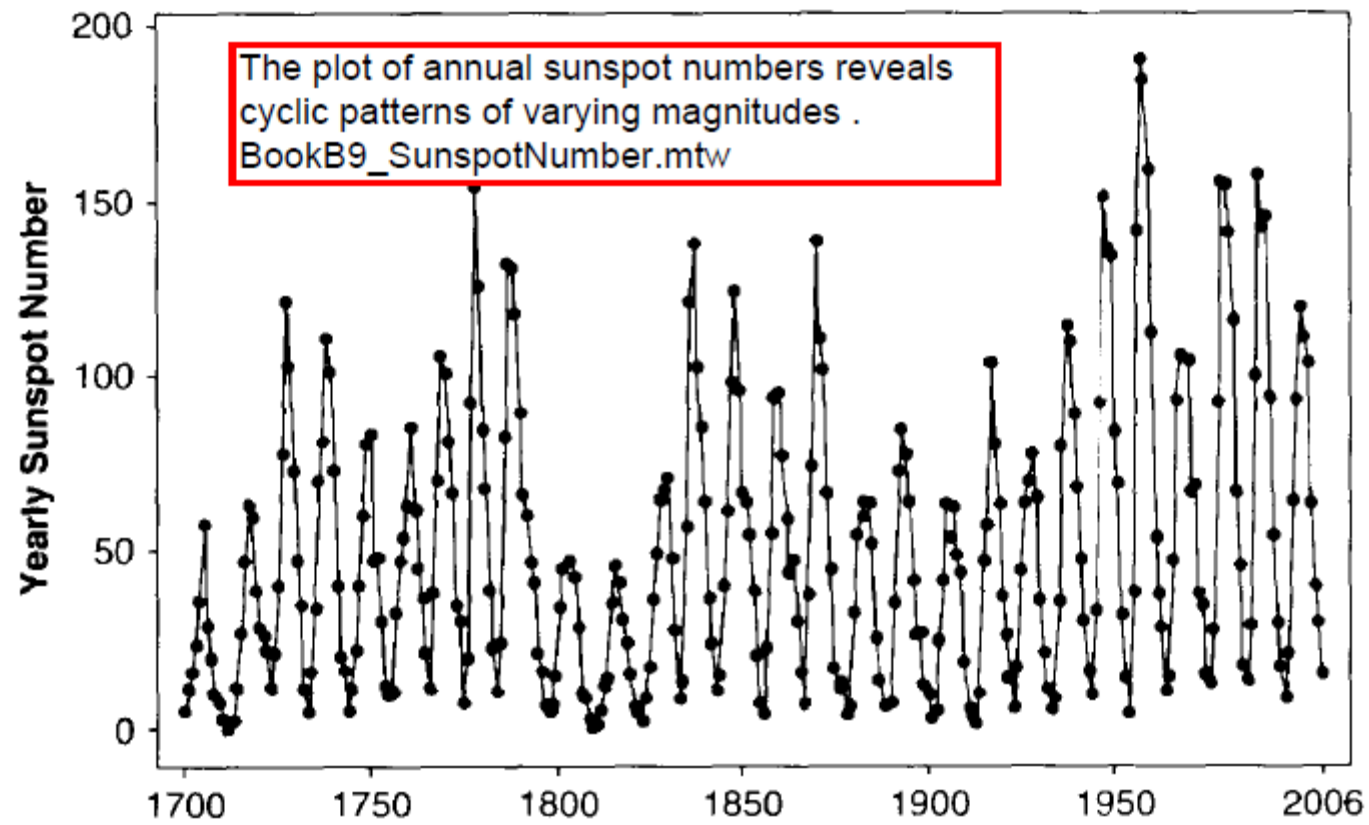
The plot of annual sunspot numbers reveals cyclic patterns of varying magnitudes .
BookB9_SunspotNumber.mtw

FIGURE 1.9   The International Sunspot Number. (Source: SIDC.)

Weekly sales of a generic
pharmaceutical product dropped due to
limited availability resulting from
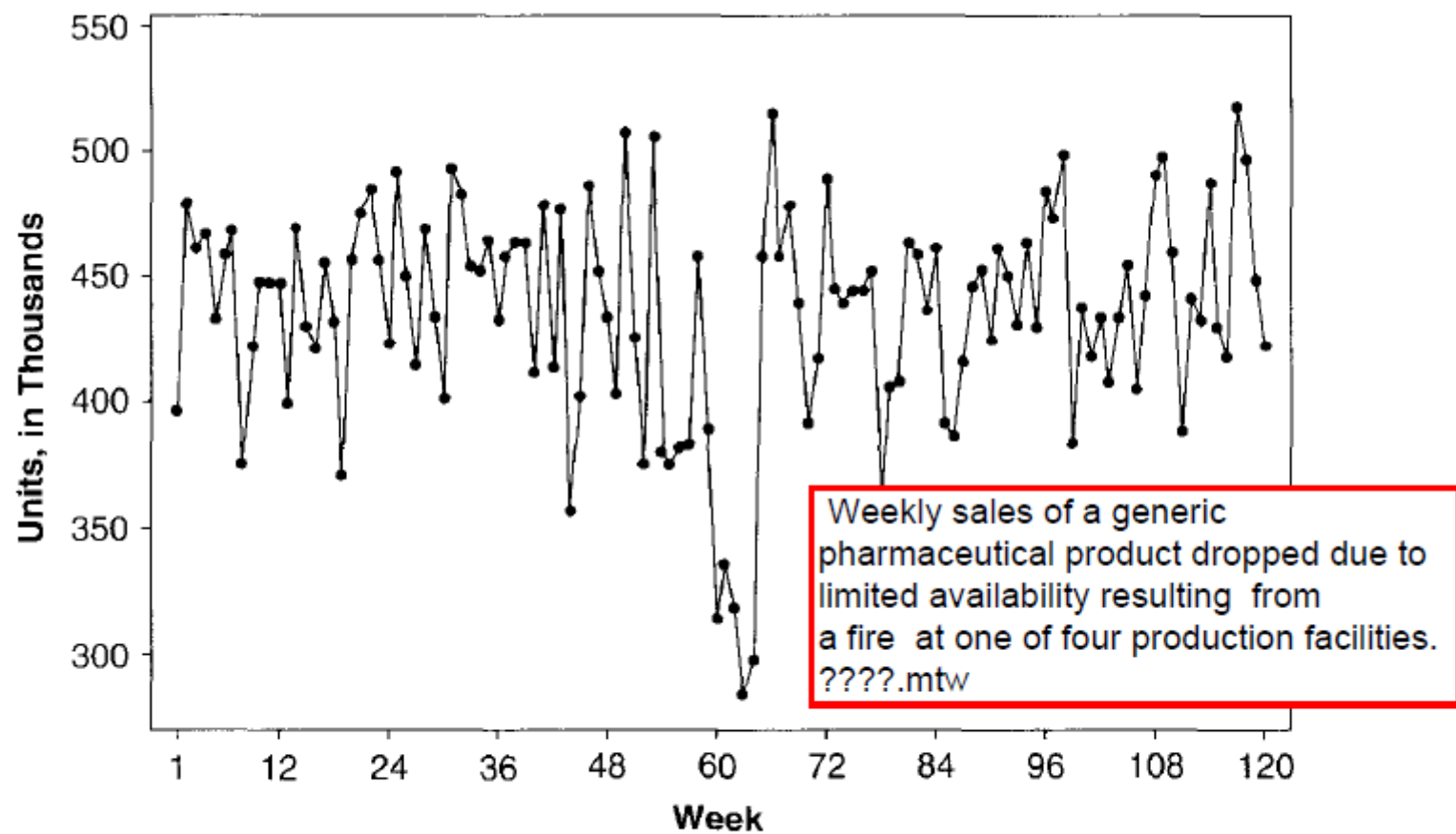a fire at one of four production facilities.
????.mtw

FIGURE 1.10 Pharmaceutical product sales.
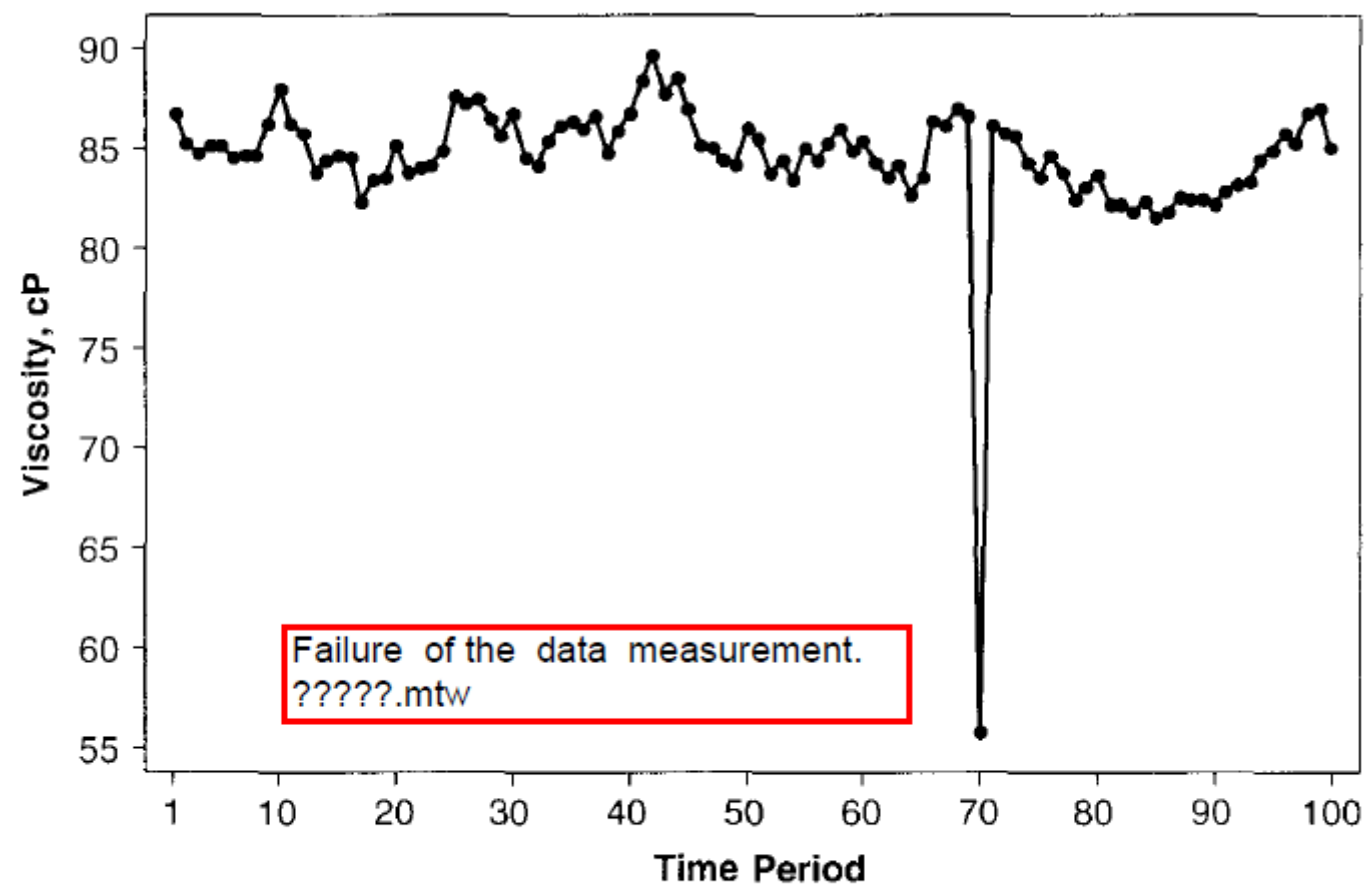
FIGURE 1.11 Chemical process viscosity readings, with sensor malfunction.

The Forecasting Process
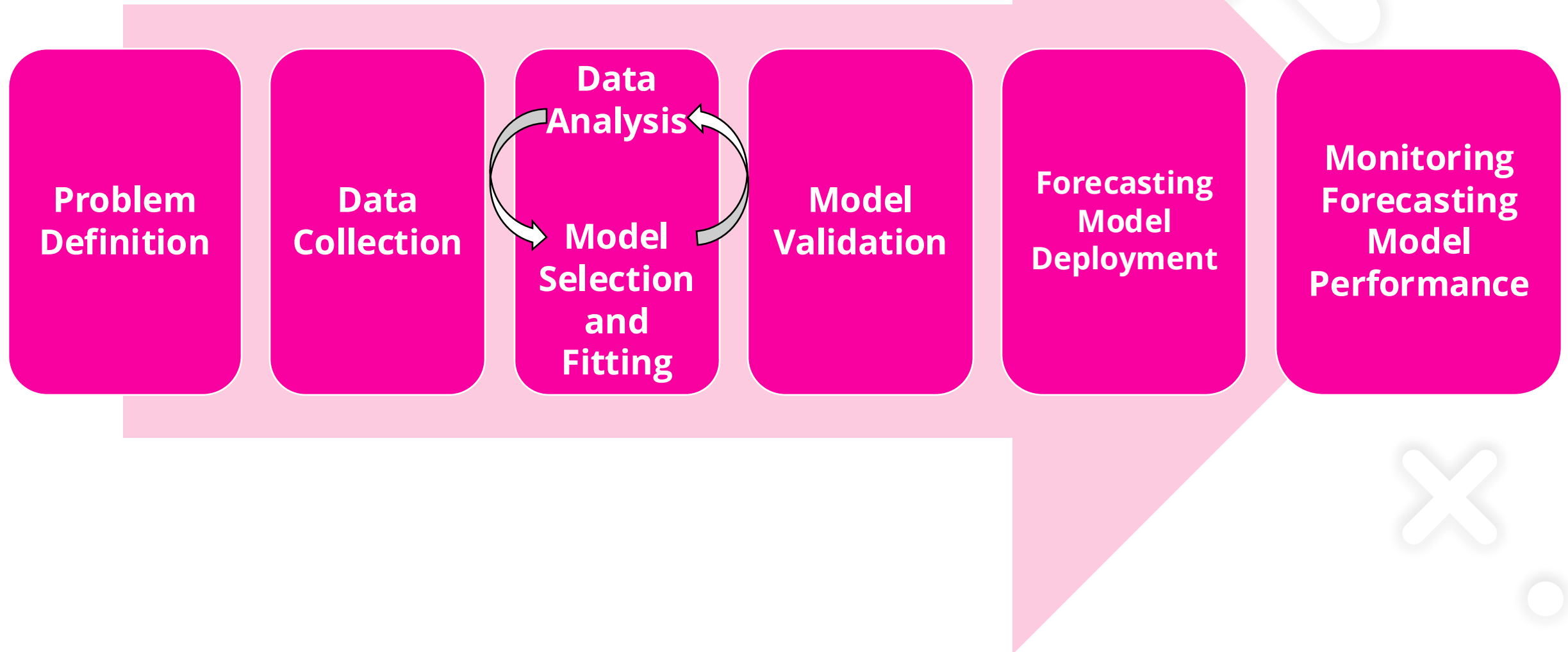
Problem Definition

+ involves developing understanding of how the forecast will be used along with the expectations of the "customer" (the user of the forecast).

+ Questions that must be addressed during this phase include
  - the desired form of the forecast (e.g., are monthly forecasts required).
  - the forecast horizon or lead time.
  - how often the forecasts need to be revised (the forecast interval). and
  - what level of forecast accuracy is required in order to make good business decisions.

+ This is also an opportunity to introduce the decision makers to the use of prediction intervals as a measure of the risk associated with forecasts.

Data Collection

+ consists of obtaining the relevant history for the variable(s) that are to be forecast, including historical information on potential predictor variables.

+ necessary to deal with missing values of some variables, potential outliers, or other data-related problems that have occurred in the past.

+ begin planning how the data collection and storage issues in the future will be handled so that the reliability and integrity of the data will be preserved.

Data Analysis

+ Time series plots of the data should be constructed and visually inspected for recognizable patterns, such as trends and seasonal or other cyclical components.

+ Sometimes we will smooth the data to make identification of the patterns more obvious.

+ Numerical summaries of the data, such as the sample mean, standard deviation, percentiles, and autocorrelations, should also be computed and evaluated.

+ If potential predictor variables are available, scatter plots of each pair of variables should be examined.

+ Unusual data points or potential outliers should be identified and flagged for possible further study.

Model Selection and Fitting

+ consists of choosing one or more forecasting models and fitting the model to the data. By fitting, we mean estimating the unknown model parameters, usually by the method of least squares.

Model Validation

+ consists of an evaluation of the forecasting model to determine how it is likely to perform in the intended application.

+ must examine what magnitude of forecast errors will be experienced when the model is used to forecast "fresh" or new data.

Forecasting Model Deployment

+ involves getting the model and the resulting forecasts in use by the customer.

+ ensure that the customer understands how to use the model and that generating timely forecasts from the model becomes as routine as possible.

+ Model maintenance, including making sure that data sources and other required information will continue to be available to the customer.

Monitoring Forecasting Model Performance

+ should be an ongoing activity after the model has been deployed to ensure that it is still performing satisfactorily.

+ Control charts of forecast errors are a simple but effective way to routinely monitor the performance of a forecasting model.