CHAPTER 1

Introduction to Forecasting

It is difficult to make predictions, especially about the future.

NEILS BOHR, *Danish physicist*

Nature and Uses of Forecasts

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.

Forecasting problems occur in many fields:

- Business and industry
- Economics
- Finance
- Environmental sciences
- Social sciences
- Political sciences

Forecasting Problems

- Short-term
  - Predicting only a few periods ahead (hours, days, weeks)
  - Typically based on modeling and extrapolating patterns in the data
- Medium-term
  - One to two years into the future, typically
- Long-term
  - Several years into the future

Most forecasting problems involve a time series:

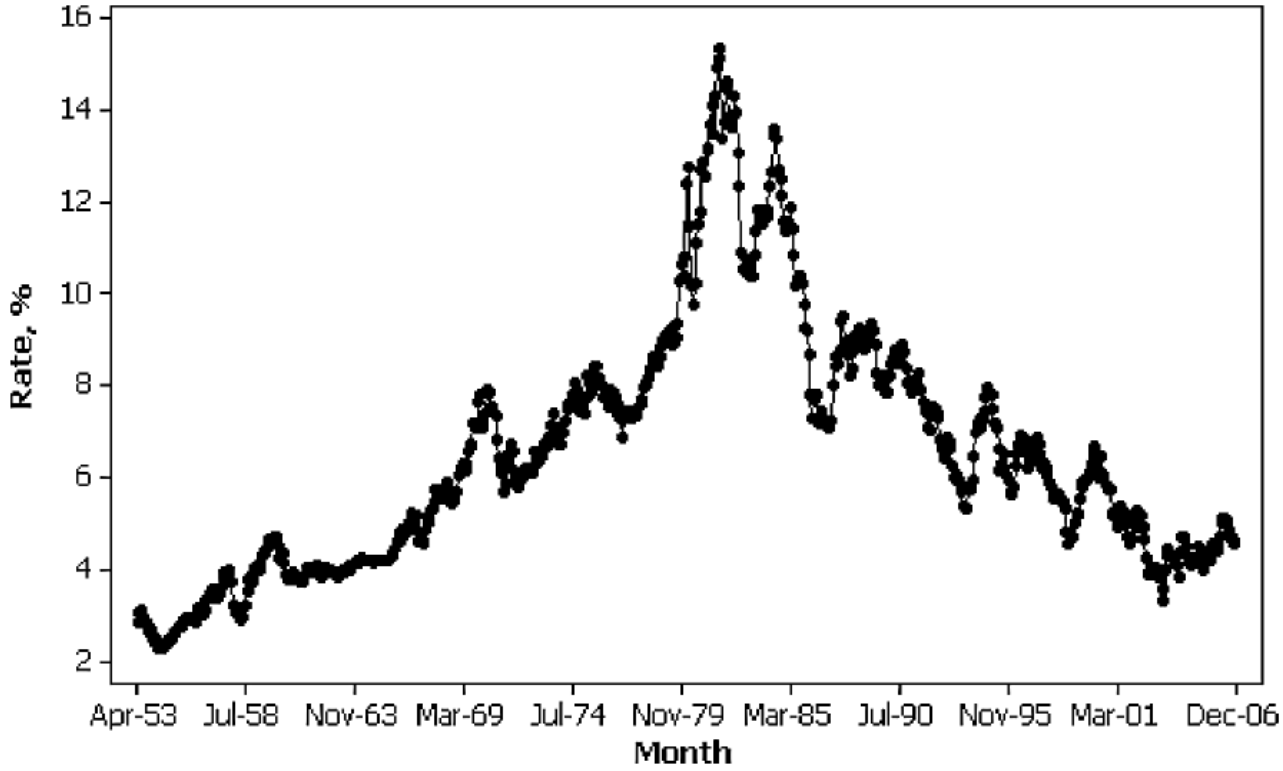

FIGURE 1.1 Time series plot of the market yield on U.S. Treasury Securities at 10-year constant maturity. (Source: U.S. Treasury.)

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.

The reason that forecasting is so important is that prediction of future events is a critical input into many types of planning and decision making processes, with application to areas such as the following:

1. *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.

2. *Marketing*. Forecasting is important in many marketing decisions. Forecasts of sales response to advertising expenditures, new promotions, or changes in pricing polices enable businesses to evaluate their effectiveness, determine whether goals are being met, and make adjustments.
3. *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.
4. *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.

5. *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.
6. *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.

Two broad types of methods:

- Quantitative forecasting methods
  - Makes formal use of historical data
  - A mathematical/statistical model
  - Past patterns are modeled and projected into the future
- Qualitative forecasting methods
  - Subjective
  - Little available data (new product introduction)
  - Expert opinion often used
  - The Delphi method

Quantitative Forecasting Methods

- Regression methods
  - Sometimes called causal methods
  - Chapter 3
- Smoothing methods
  - Often justified empirically
  - Chapter 4
- Formal time series analysis methods
  - Chapters 5 and 6
  - Some other related methods are discussed in Chapter 7

Terminology

- Point forecast or point estimate
- Prediction interval
- Forecast horizon or lead time
- Forecasting interval
- Moving horizon forecasts

Examples of time series: Uncorrelated data, constant process model

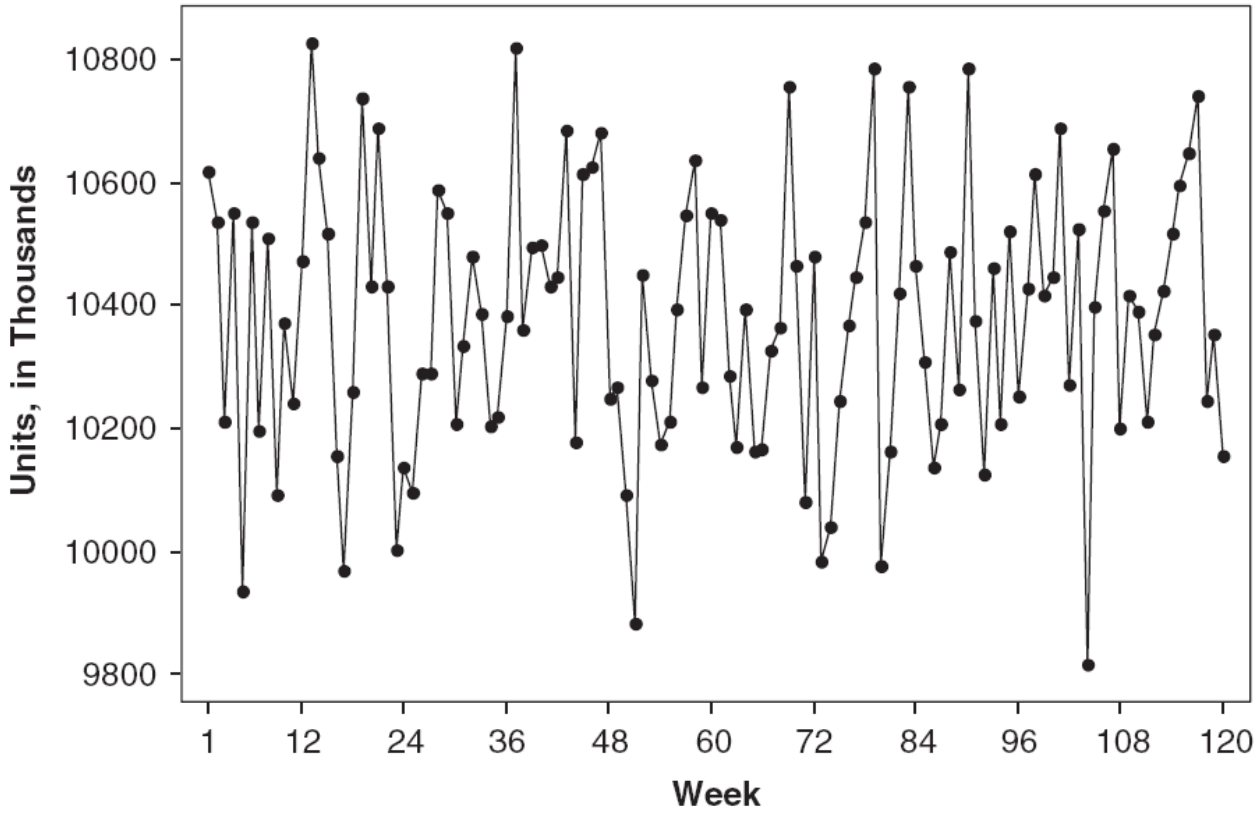

FIGURE 1.2 Pharmaceutical product sales.

Autocorrelated data

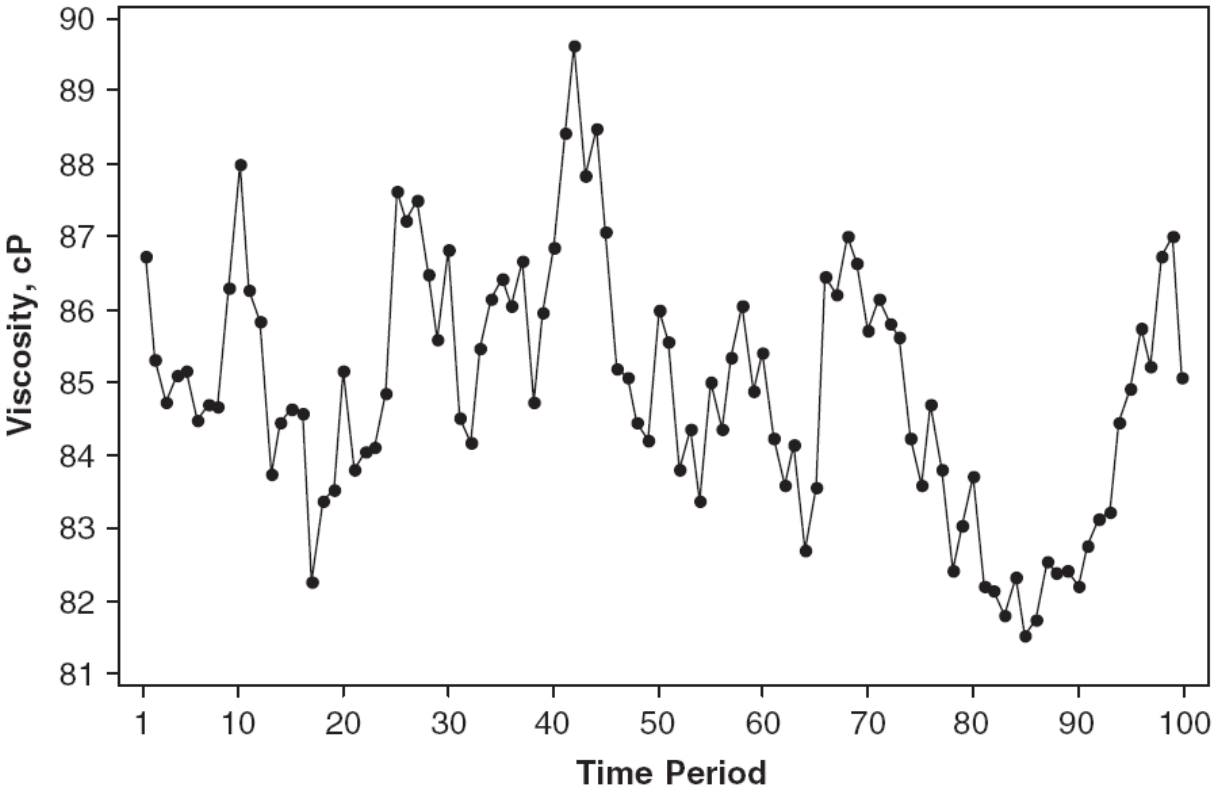

FIGURE 1.3 Chemical process viscosity readings.

Trend

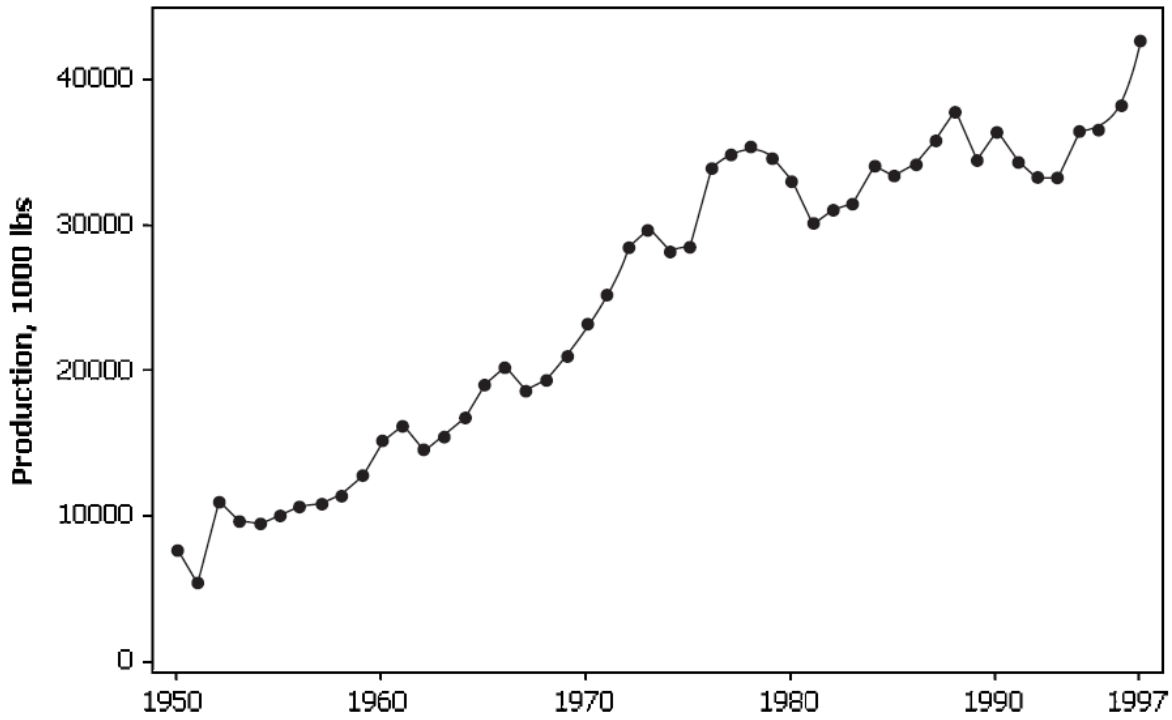

FIGURE 1.4 The U.S. annual production of blue and gorgonzola cheeses. (Source: USDA–NASS.)

Cyclic or seasonal data

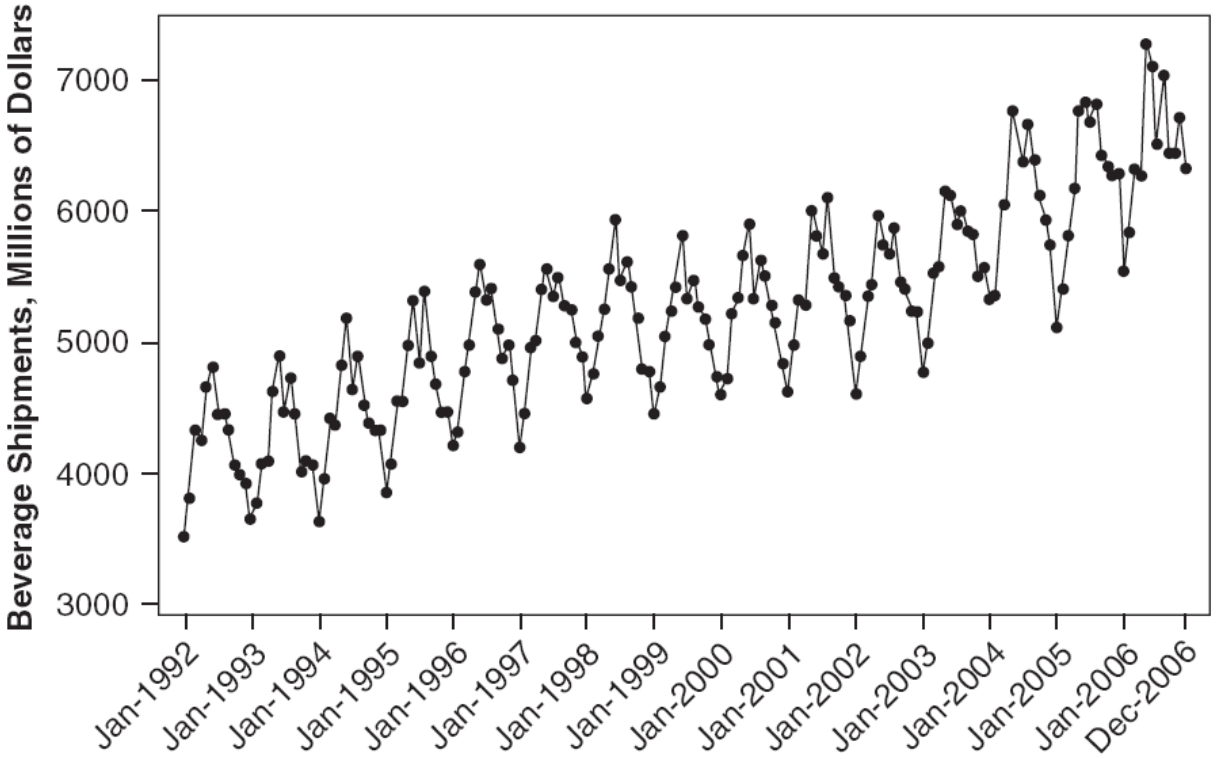

FIGURE 1.5 The U.S. beverage manufacturer monthly product shipments, unadjusted. (Source: U.S. Census Bureau.)

Nonstationary data

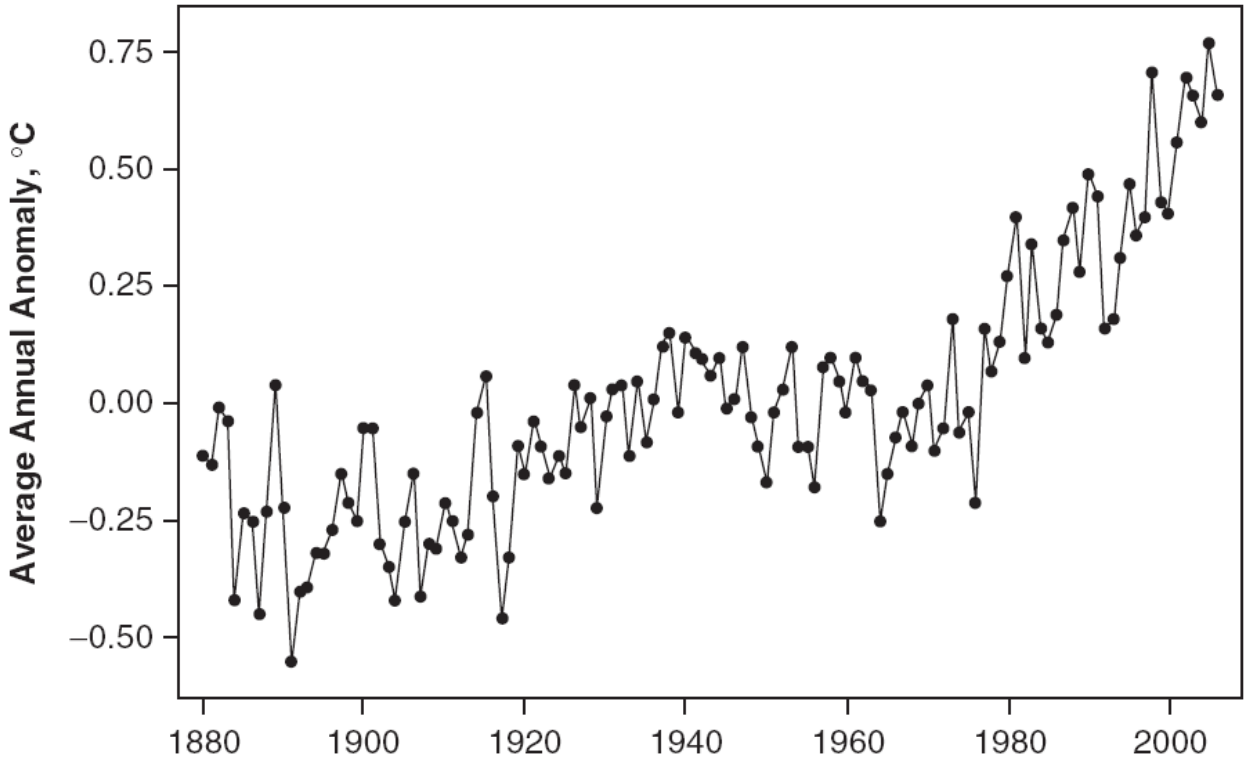

FIGURE 1.6 Global mean surface air temperature annual anomaly. (Source: NASA–GISS.)

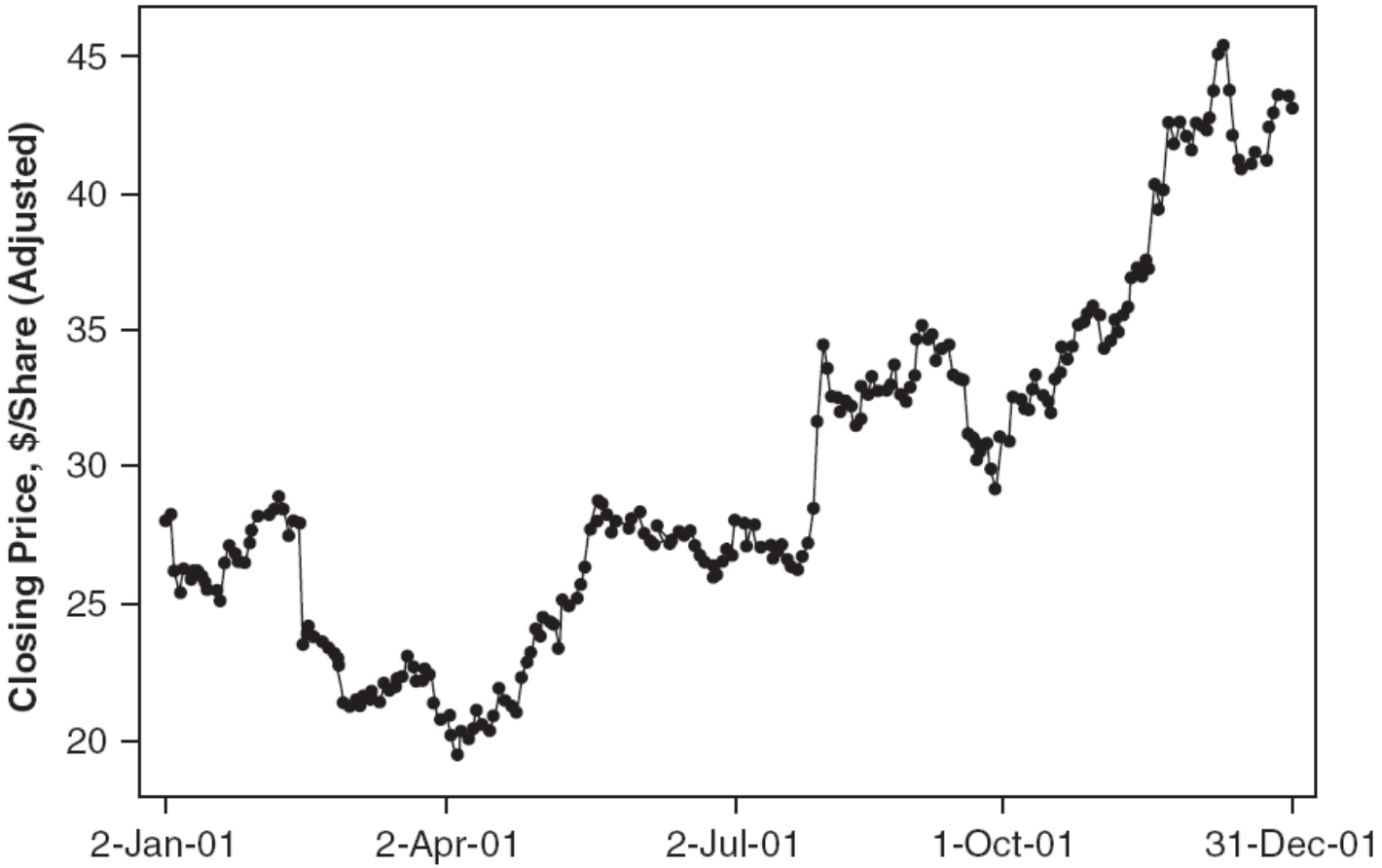

FIGURE 1.7 Whole Foods Market stock price, daily closing adjusted for splits.

A mixture of patterns

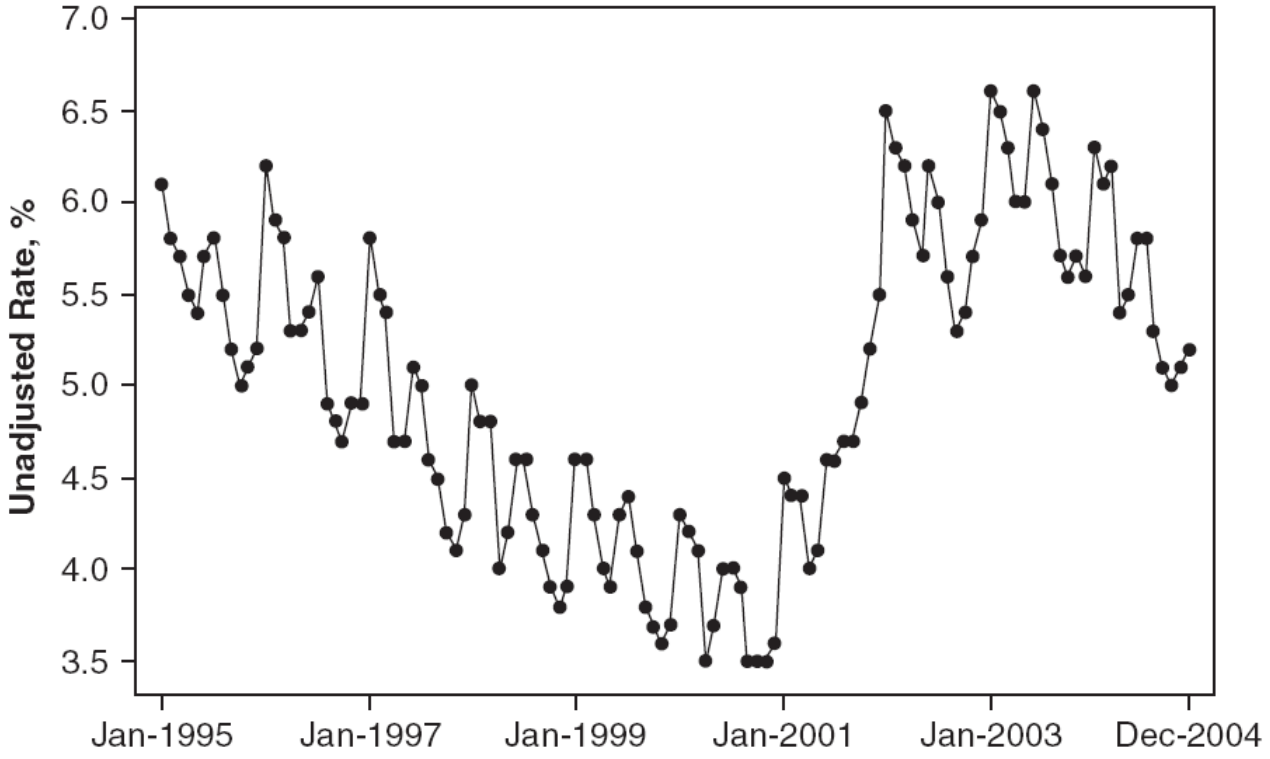

FIGURE 1.8 Monthly unemployment rate—full-time labor force, unadjusted. (Source: U.S. Department of Labor–BLS.)

Cyclic patterns of different magnitudes

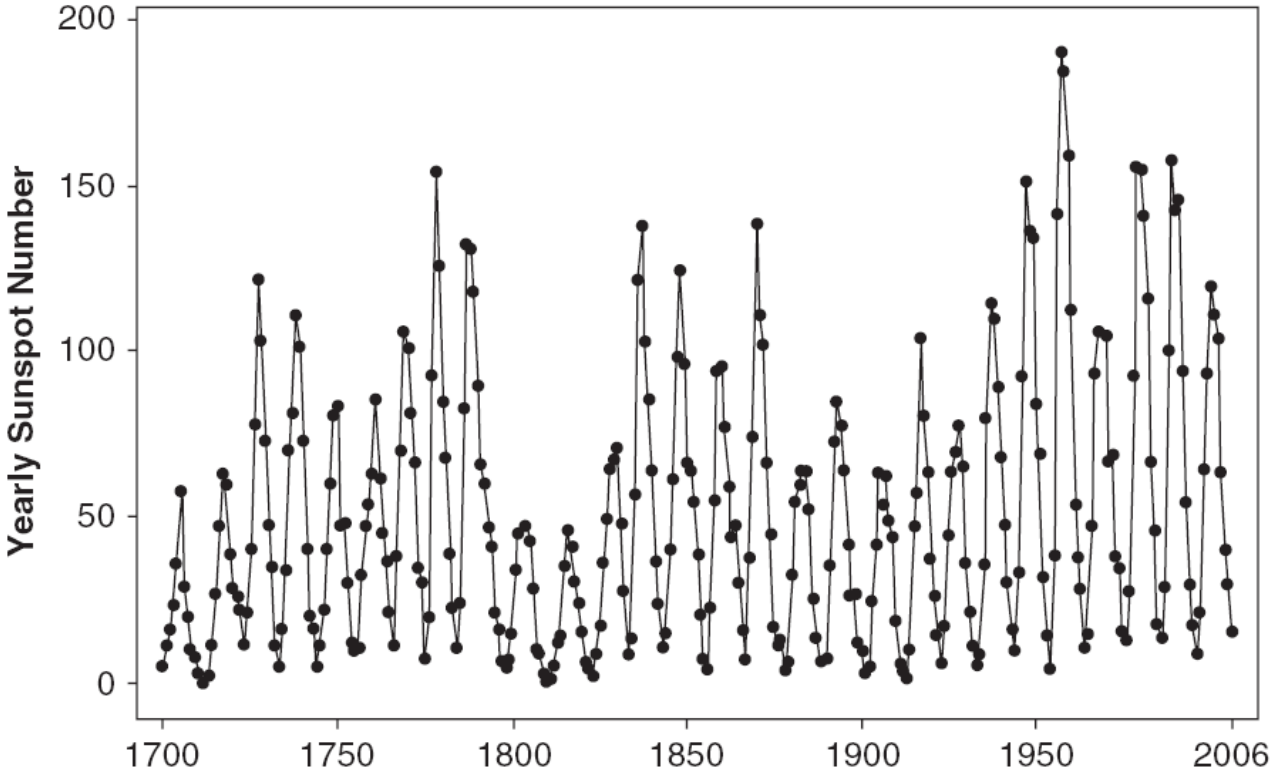

FIGURE 1.9 The International Sunspot Number. (Source: SIDC.)

Atypical events

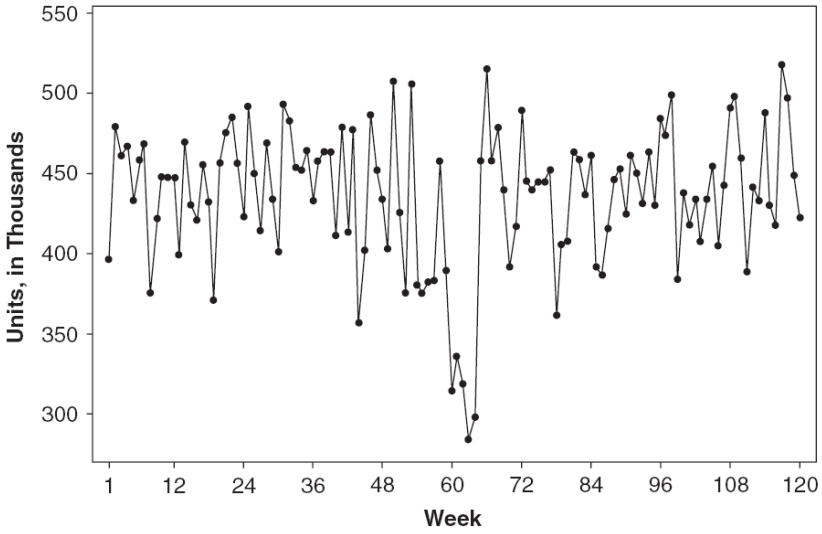

FIGURE 1.10 Pharmaceutical product sales.

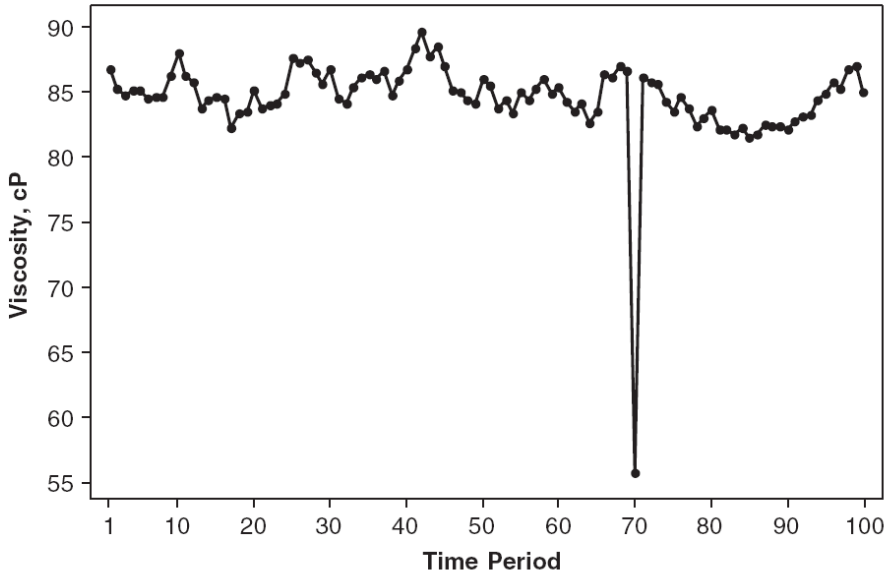

FIGURE 1.11 Chemical process viscosity readings, with sensor malfunction.

The Forecasting Process

A process is a series of connected activities that transform one or more inputs into one or more outputs. All work activities are performed in processes, and forecasting is no exception. The activities in the forecasting process are:

1. Problem definition
2. Data collection
3. Data analysis
4. Model selection and fitting
5. Model validation
6. Forecasting model deployment
7. Monitoring forecasting model performance

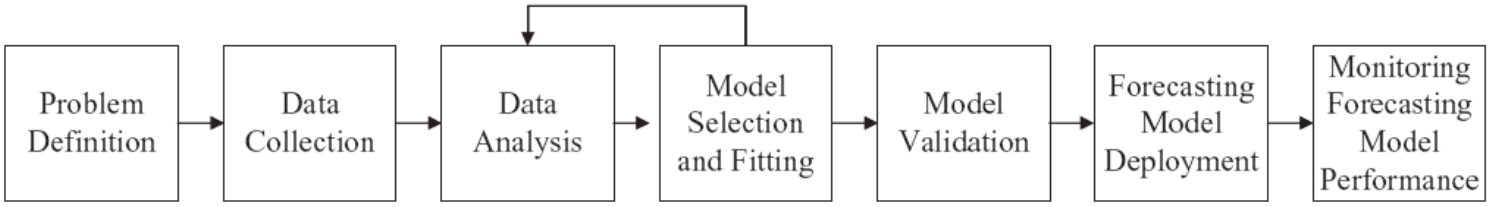

FIGURE 1.12 The forecasting process.

Data for Forecasting

- The Data Warehouse
- Data versus Information
- In most modern organizations data regarding sales, transactions, company financial and business performance, supplier performance, and customer activity and relations are stored in a repository known as a data warehouse. Sometimes this is a single data storage system; but as the volume of data handled by modern organizations grows rapidly, the data warehouse has become an integrated system comprised of components that are physically and often geographically distributed, such as cloud data storage. The data warehouse must be able to organize, manipulate, and integrate data from multiple sources and different organizational information systems.

Required Functionality

- Data Extraction
- Data Transformation
  - Duplication of records
  - Missing data
  - Other problems
  - Sometimes called **data cleaning**
- Data Loading

Data Quality

- Accuracy
- Timeliness
- Completeness
- Representativeness
- Consistency
- Jones-Farmer et al. (2014) describe how statistical quality control methods (specifically control charts) can be used to enhance data quality in the data production process.

Data Cleaning

- Looking for and fixing potential errors, missing data, outliers, inconsistencies
- Some common “automatic” checks include:
  - Is data missing?
  - Does the data fall within expected ranges?
  - Are there outliers or unusual values?
- Graphical as well as analytical methods can be useful

Data Imputation

- Correcting missing values
- Mean value imputation
- Stochastic mean value imputation
- Mean value imputation using a subset of the data:

Data Imputation

- Regression imputation – the imputed value is computed from a model
- “Hot deck” imputation – uses the data currently available
- “Cold deck” imputation – uses other data not currently in use