Chapter 1: Introduction to Forecasting

Rong Fan

South Dakota State University

rong.fan@sdstate.edu

1.1 The Nature and Uses of Forecasts

A forecast is a prediction of some future event or events.

Forecasting is an important problem that spans many fields including business and industry, government, economics, social sciences.....

Forecasting problems are classified as

1. Short-term
2. Medium-term
3. Long-term

1.1 The Nature and Uses of Forecasts

Time Series

A time series is a time-oriented or chronological sequence of observations on a variable of interest.

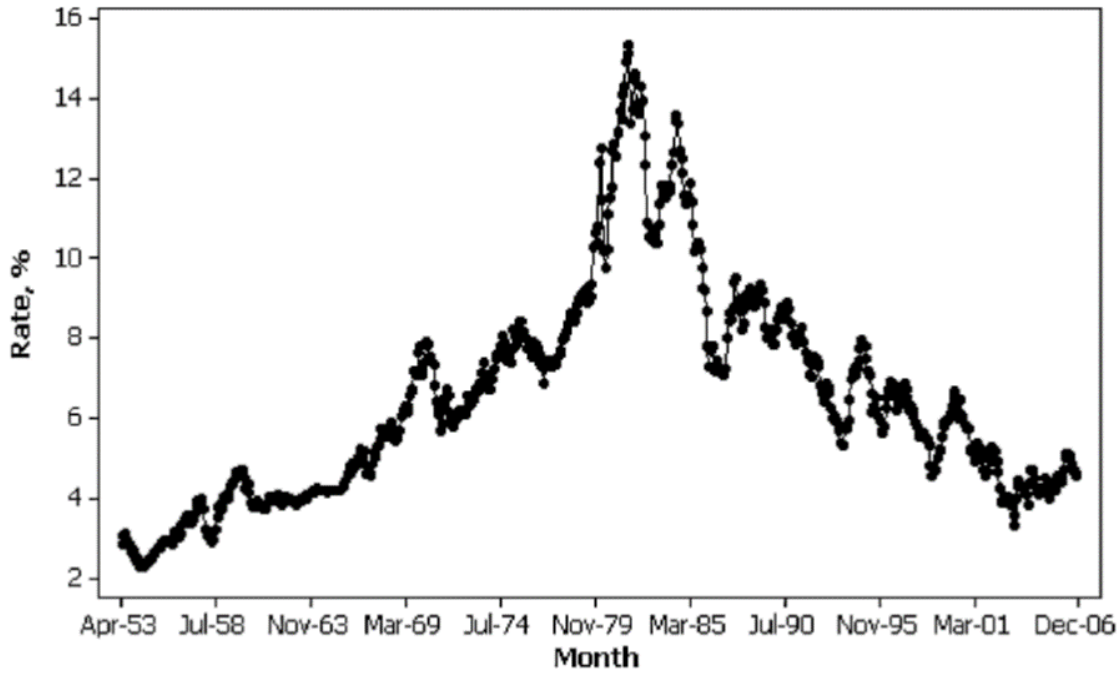

FIGURE 1.1 Time series plot of the market yield on U.S. Treasury Securities at 10-year constant maturity. (Source: U.S. Treasury.)

This graph is called a time series plot. The variable is collected at equally spaced time periods.

1.1 The Nature and Uses of Forecasts

Two broad types of methods:

1. Qualitative forecasting methods (not included in this course)
2. Quantitative forecasting methods
   a. Regression models (Chapter 3)
      — Sometimes, they are called causal forecasting models because the predictor variables are assumed to describe the forces that cause or drive the observed values of the variable of interest.
   b. Smoothing methods (Chapter 4)
   c. General time series models (Chapters 5 and 6)

1.1 The Nature and Uses of Forecasts

Terminology:

- point forecast or point estimate:
  — a single number that represents our best estimate of the future value.
- forecast error:
  — forecasts are almost always wrong, that is we experience forecast error.
- prediction interval (PI)
  — an interval of values for the future observation

1.1 The Nature and Uses of Forecasts

- forecast horizon or lead time
  — the number of future periods for which forecasts must be produced
  — the time between the initiation and completion of a production process.
- forecast interval
  — the frequency with which new forecasts are prepared
- rolling or moving horizon forecasts
  — System updates the forecasts for T-1 of the periods in the horizon and computes a forecast of the newest period T.

A forecast refers to a calculation or an estimation which uses data from previous events, combined with recent trends to come up a future event outcome. Forecast implies **time series** and **future**, while prediction does not. A prediction is a statement which tries to explain a possible **outcome** or future **event**.

1.1 The Nature and Uses of Forecasts

Example:
In a production planning, we might forecast demand on a monthly basis, for up to 3 months in the future and prepare a new forecast each month.

1. forecast horizon or lead time: 3 months
2. The forecast interval is 1 month, same as the basic period of time for which each forecast is made.
3. If the forecast lead time is always the same length, say, 3 months, and the forecast is revised each month, then we are employing a rolling or moving horizon forecasting approach. This system updates or revise the forecasts for 2 months in the horizon and computes a forecast for the newest 3rd month.

1.2 Some Examples of Time Series

Uncorrelated data, constant process model

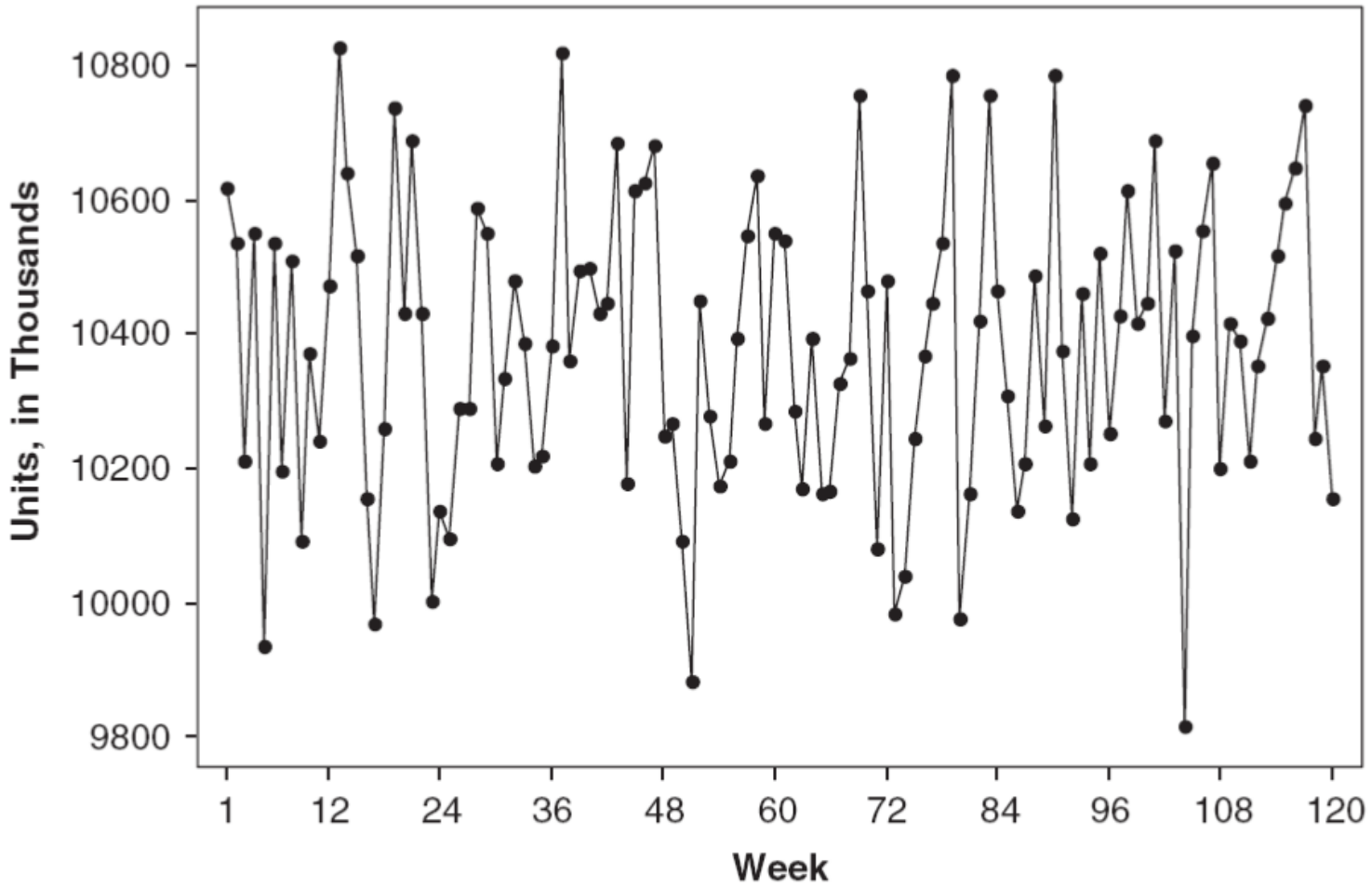

FIGURE 1.2 Pharmaceutical product sales.

1.2 Some Examples

Autocorrelated data: Autocorrelation, also known as serial correlation, is the correlation of a signal with a delayed copy of itself as a function of delay. Informally, it is the similarity between observations as a function of the time lag between them.

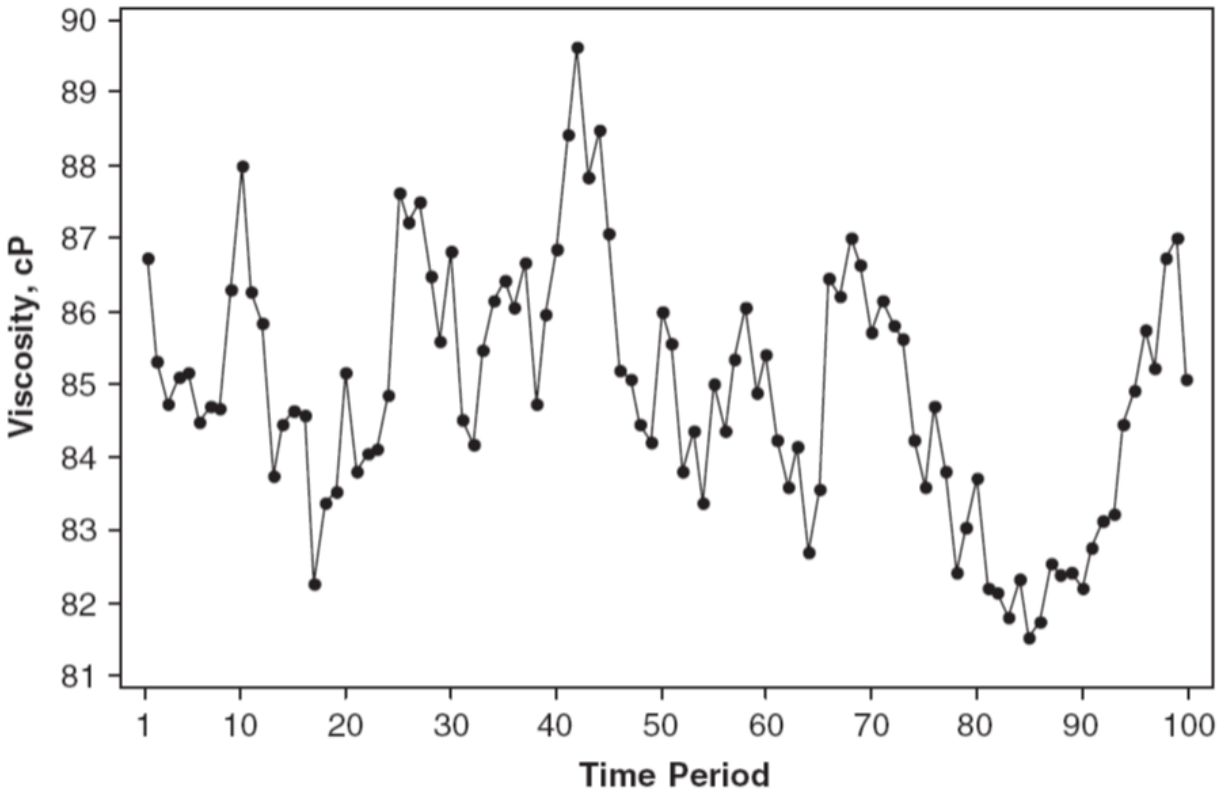

FIGURE 1.3 Chemical process viscosity readings.

1.2 Some Examples

Trend:A pattern of gradual change in a condition, output, or process, or an average or general tendency of a series of data points to move in a certain direction over time, represented by a line or curve on a graph.

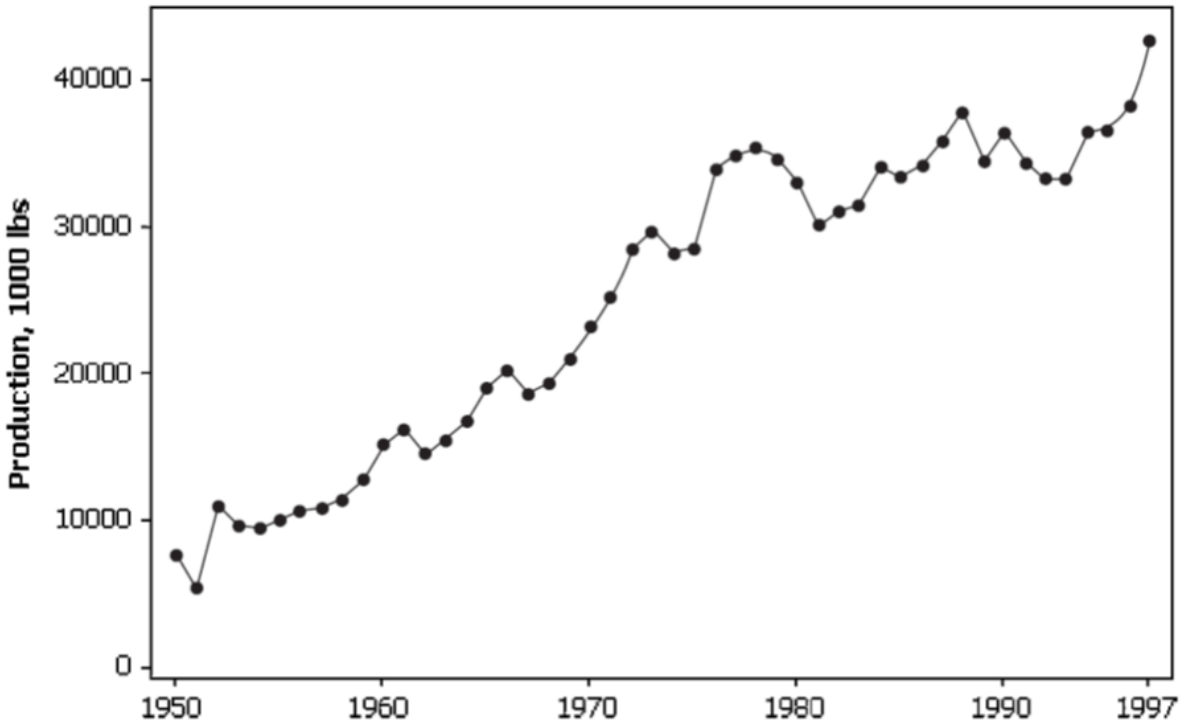

FIGURE 1.4 The U.S. annual production of blue and gorgonzola cheeses. (Source: USDA–NASS.)

1.2 Some Examples

Seasonal data

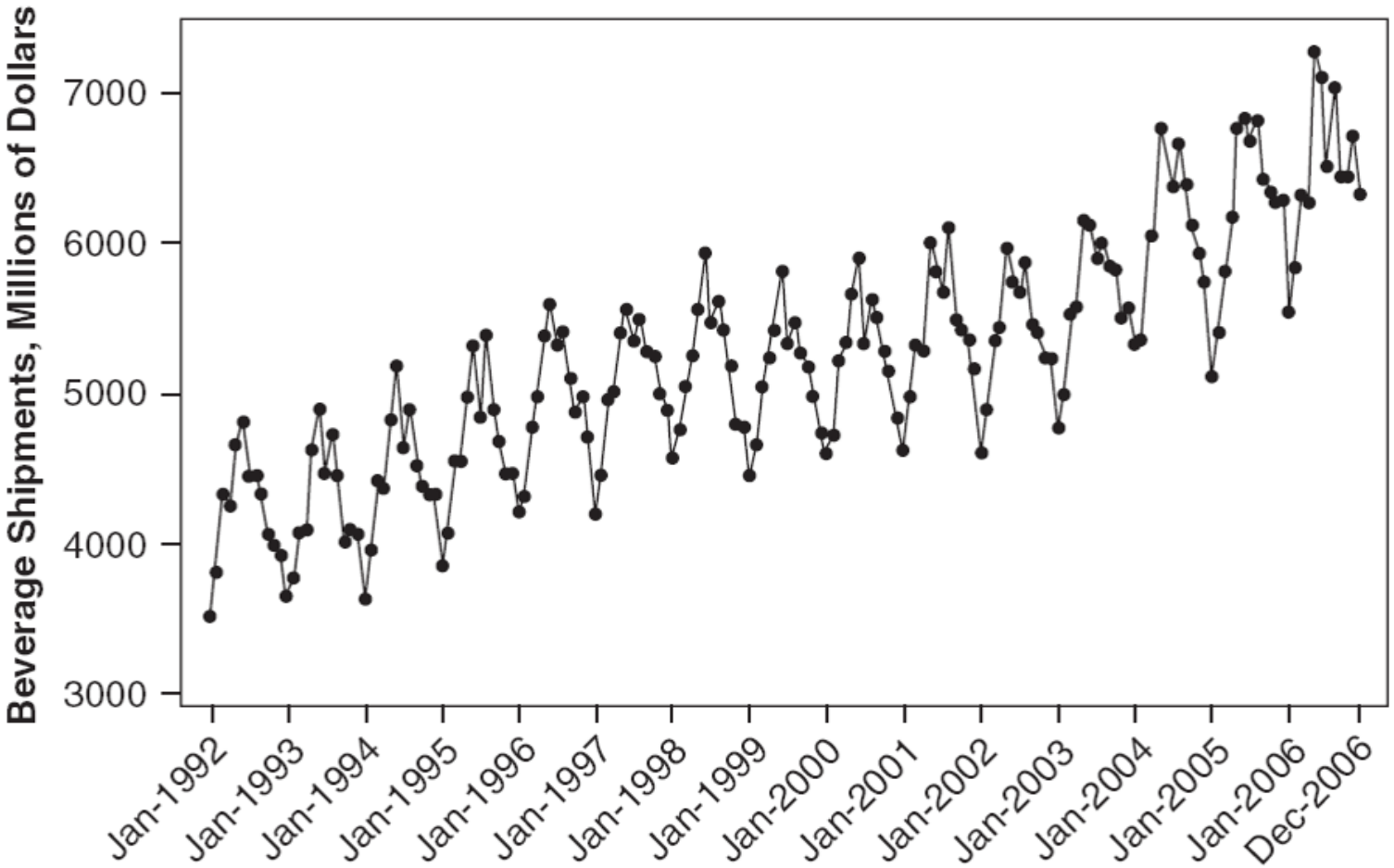

FIGURE 1.5 The U.S. beverage manufacturer monthly product shipments, unadjusted. (Source: U.S. Census Bureau.)

1.2 Some Examples

Nonstationary data

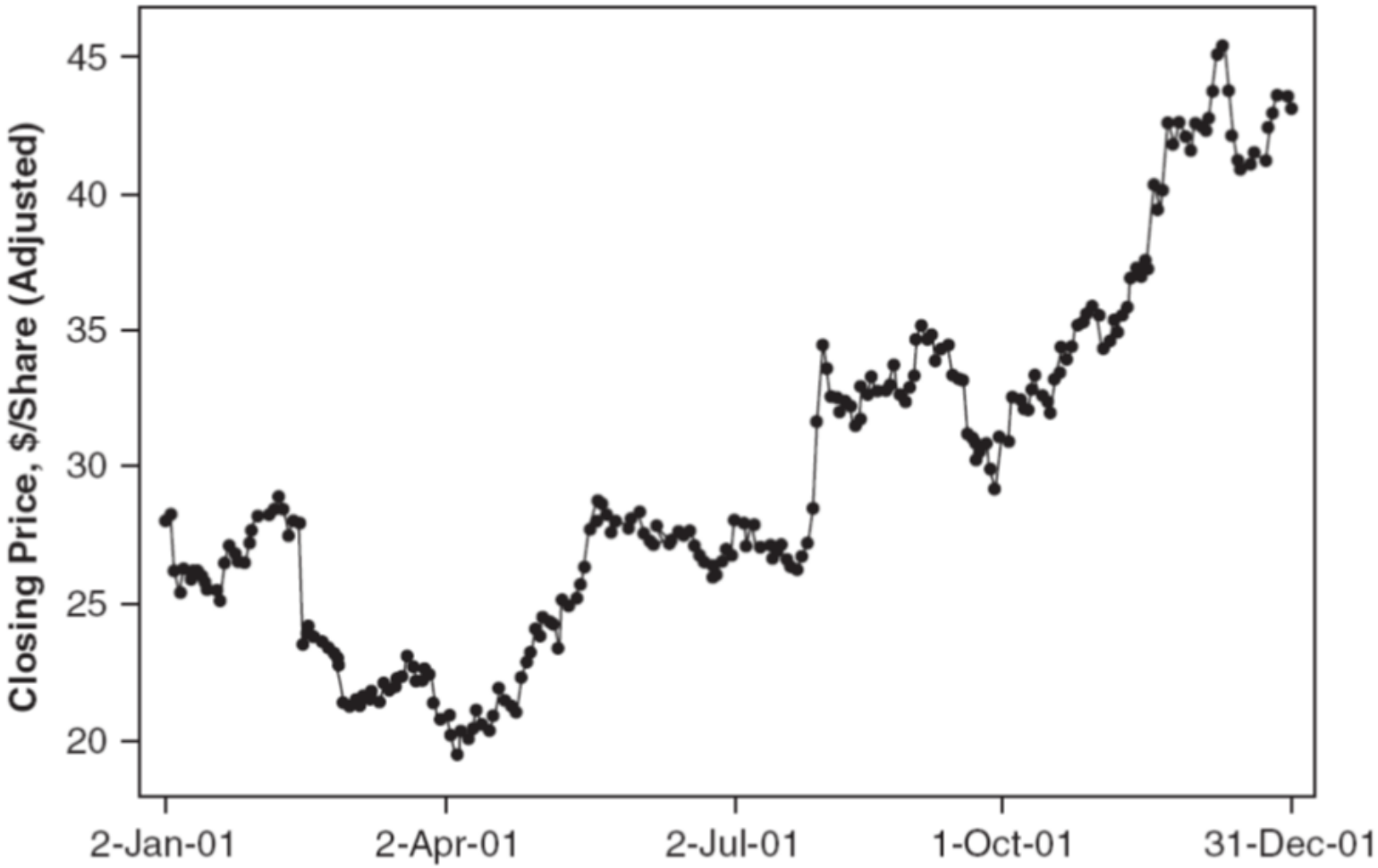

FIGURE 1.7 Whole Foods Market stock price, daily closing adjusted for splits.

1.2 Some Examples

A mixture of patterns

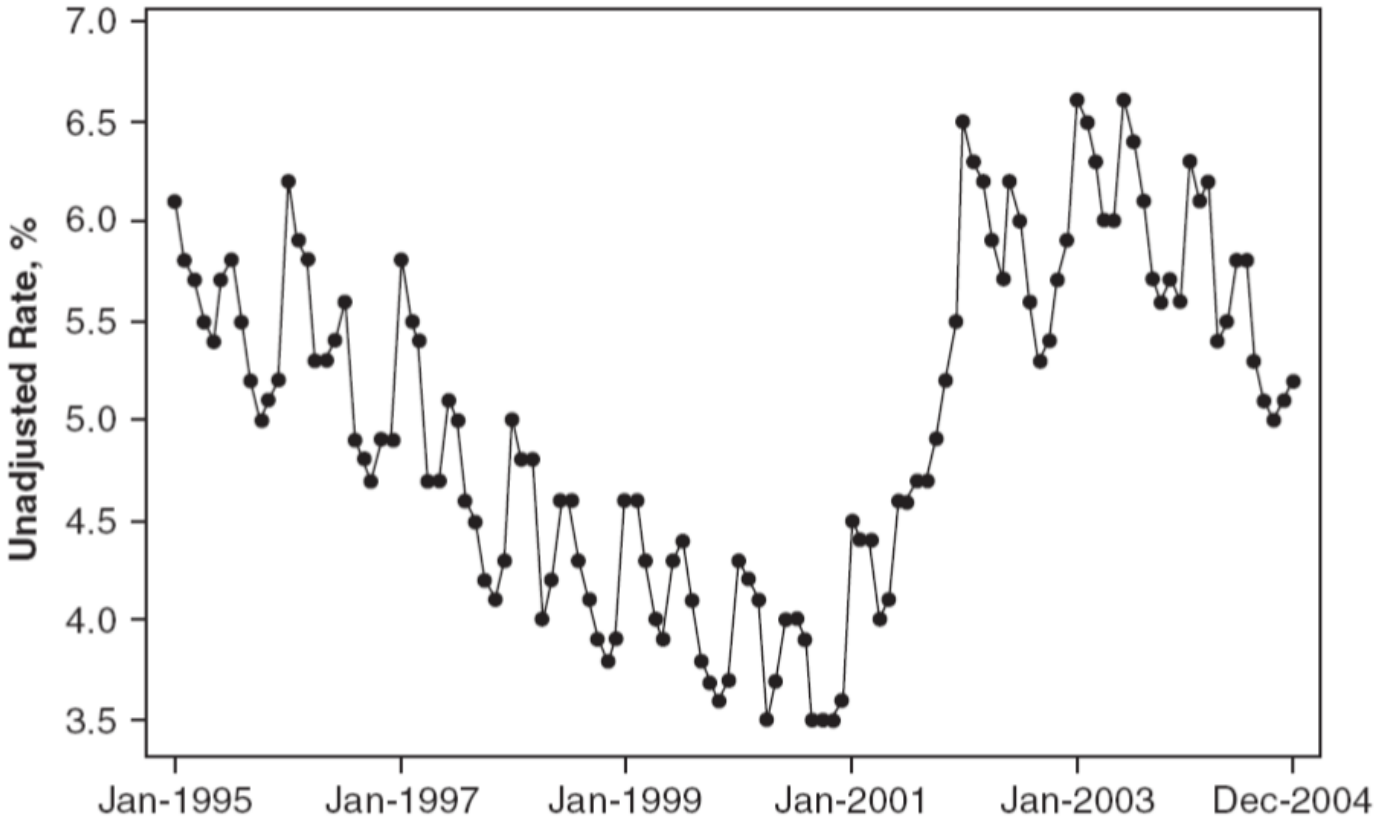

FIGURE 1.8 Monthly unemployment rate—full-time labor force, unadjusted. (Source: U.S. Department of Labor–BLS.)

1.2 Some Examples

Atypical events

Weekly sales dropped due to limited availability resulting from a fire at one of the four production facilities.

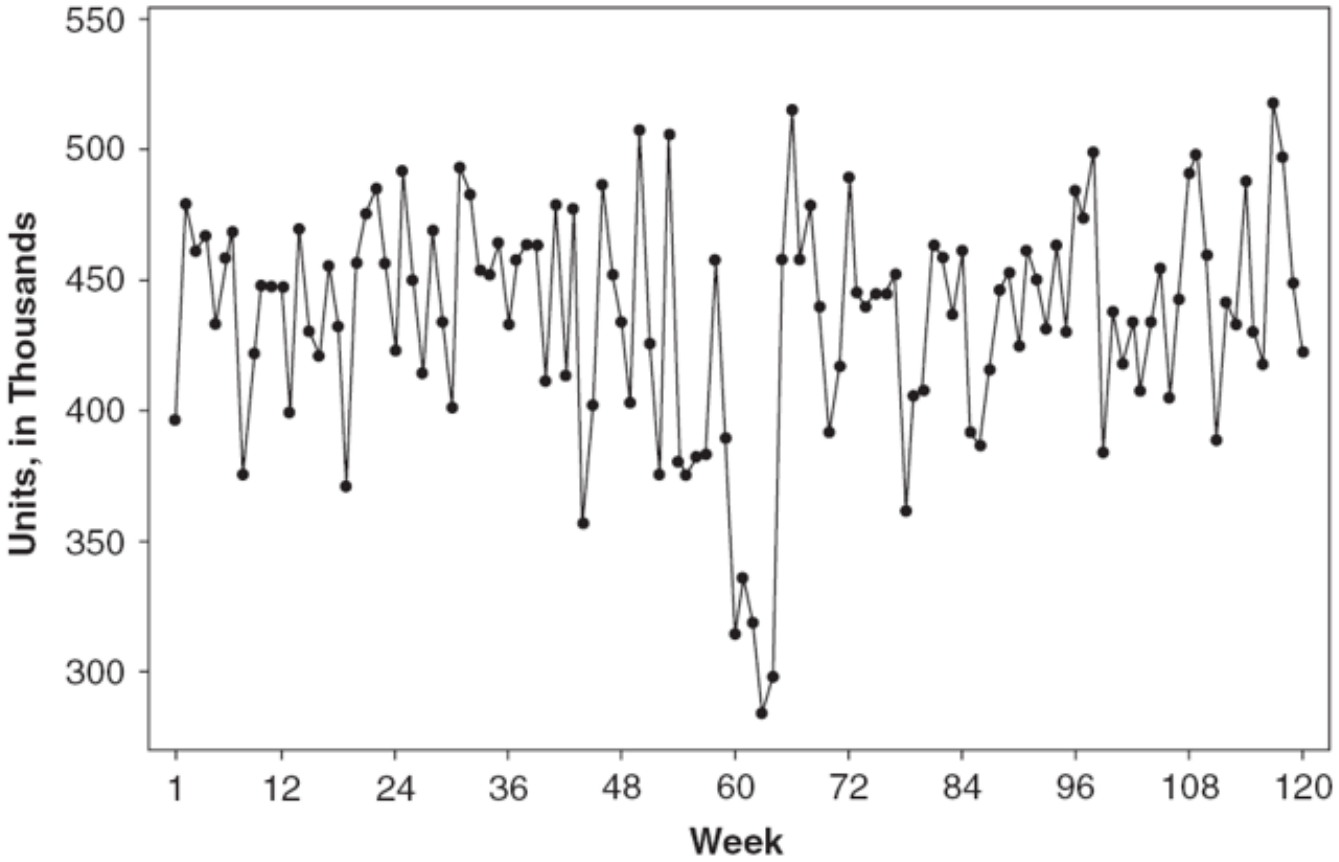

FIGURE 1.10 Pharmaceutical product sales.

1.2 Some Examples

Atypical events

The measurement system was determined to be out of calibration at time period 70.

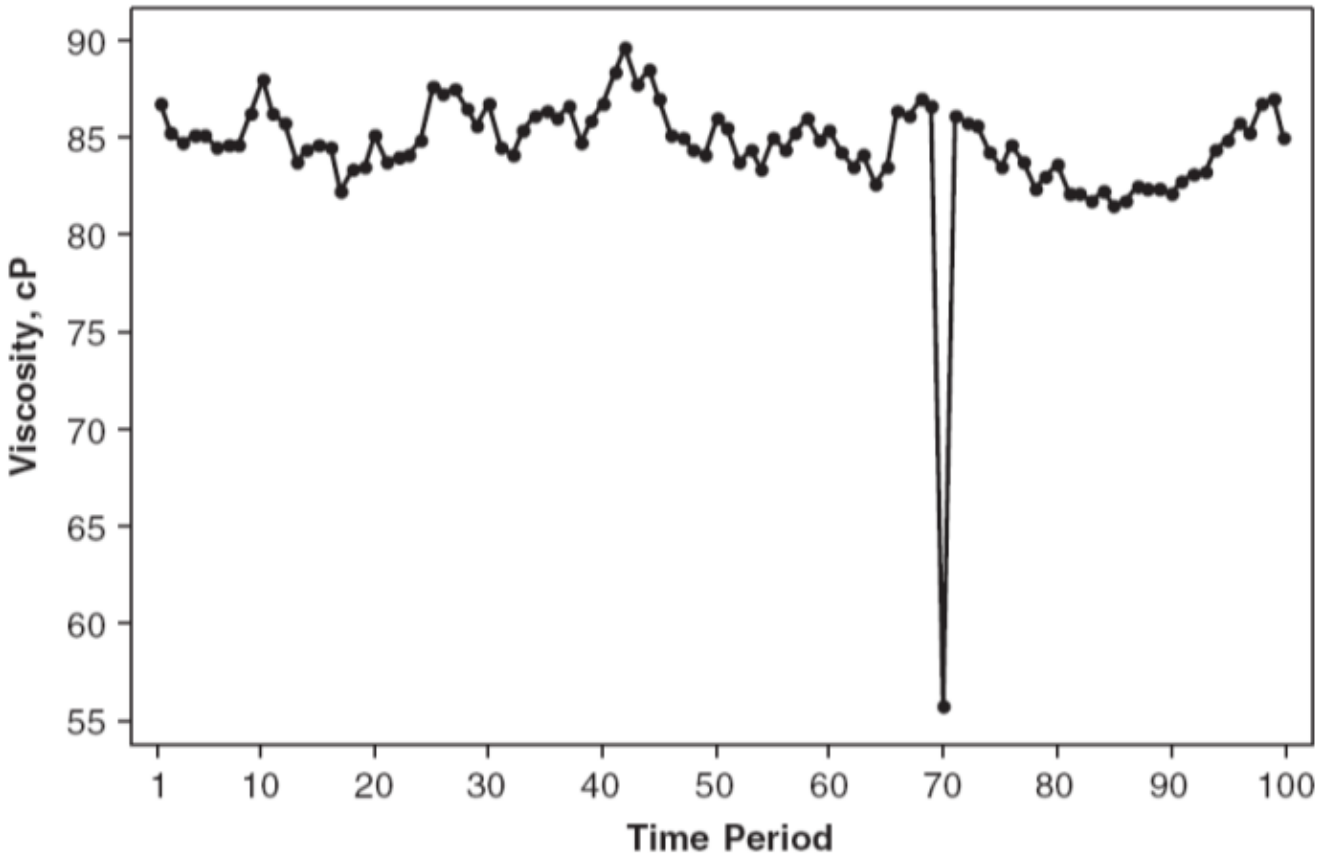

FIGURE 1.11 Chemical process viscosity readings, with sensor malfunction.

1.3 Forecasting Process

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

1.4 Data

Data cleaning for missing data and outliers.

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

1.4 Data

Data imputation is the process of correcting missing data or replacing outliers with an estimation process. For example,

1. Mean value imputation
2. Stochastic mean value imputation
3. Mean value imputation using a subset of data

$$y_j^* = \frac{1}{2k}\left(\sum_{t=j-k}^{j-1} y_t + \sum_{t=j+1}^{j+k} y_t\right)$$

4. Regression imputation: the imputed value is computed from a model
5. Hot deck imputation: uses the data currently available
6. Cold deck imputation: uses other data not currently in use

Homework Format Requirement

Format: Using R-markdown or Word to submit a pdf version.

Acknowledgement

Thank Dr.Zheng Hao, a former instructor at SDSU, for his contribution to the note.