

Ch01 Study Guide  
Definitions

| Term or Concept                                     | Definition   |
|---|--|
| <b>Statistics</b>                                   | 1) The language of data.<br>2) The art and science of getting information from data.<br>3) The study of collecting, analyzing, presenting, and interpreting data.  |
| <b>The Language of Data</b>                         |  |
| <b>Data are information</b>                         | Data are information, which is comprised of characteristics about a subject of interest.   |
| <b>Who are the data about?</b>                      | Who are the subjects of the data? Subjects can be people, objects, events, etc. This includes any demographic information that describes the subject, such as country, gender, or any other group label.   |
| <b>What was measured?</b>                           | What characteristics are measured about each subject, e.g., demographic information, sales information, opinions, costs, rates, time, etc.   |
| <b>Why was the data collected?</b>                  | What business problem was the data collected to support?   |
| <b>When, where, and how was the data collected?</b> | The source of the data can make the difference between insight and nonsense. When were the data collected? Where were the data collected? How were the data collected?                                     |
| <b>Variables</b>                                    | A characteristic observed about people, objects, or events is called a variable because the values often differ in kind or degree among the various subjects.  |
| <b>Data Tables</b>                                  | The values of the characteristics are organized into a data table with each row representing a subject and each column representing a variable.  |
| <b>Databases and Data Warehouses</b>                | Data are organized and stored in a form that supports efficient movement or processing, i.e., electronically in databases and data warehouses.   |
| <b>Cross-sectional Data</b>                         | Cross-sectional data is a set of data points collected by observing many subjects (such as individuals, firms, countries, or regions) at the same point of time, or without regard to differences in time. |

Ch01 Study Guide  
Definitions

| Term or Concept  | Definition   |
|--|--|
| <b>Time Series Data</b>  | Time series data is a set of data points indexed (or listed or graphed) in time order. Most commonly, a time series is a sequence taken at successive, equally spaced points in time such as daily, weekly, monthly, quarterly, annually, etc. |
| <b>Qualitative Variable</b>  | A variable that assumes labels or names to identify the characteristic.  |
| <b>Quantitative Variable</b>   | A variable that assumes numeric values.  |
| <b>The Art and Science of Getting Information from Data</b>                  |  |
| <b>Researcher</b>  | Someone who studies a problem using statistical methods and reports or presents the information obtained.  |
| <b>Consumer of Statistics</b>  | Someone who reads statistical reports to obtain information about a problem.   |
| <b>The Study of Collecting, Analyzing, Presenting, and Interpreting Data</b> |  |
| <b>Population</b>  | All items or subjects of interest in a statistical problem.  |
| <b>Sample</b>  | A representative subset of the population.   |
| <b>Population Parameters</b>   | Data that describe a characteristic of a population.   |
| <b>Sample Statistics</b>   | 1) Data that describe a characteristic of a sample.<br>2) An estimate of a population parameter.   |
| <b>Sampling</b>  | 1) The process of selecting a subset of a population to study.<br>2) Used to estimate population parameters.<br>3) Used heavily in manufacturing and service settings to ensure high quality products and services.                            |
| <b>Descriptive Statistics</b>  | Branch of statistics concerned with numeric (averages, percentages, etc.) and graphical summaries of data.   |
| <b>Inferential Statistics</b>  | Branch of statistics concerned with the problem of estimating population parameters and testing hypotheses about the parameters.   |