


Business Statistics:
A Decision-Making Approach
6th Edition

Chapter 1
The Where, Why, and How of
Data Collection

Business Statistics: A Decision-Making Approach, 6e © 2005 Prentice-Hall, Inc. Chap 1-1




Chapter Goals

After completing this chapter, you should be able to:

- Describe key data collection methods
- Know key definitions:
 - ◆ Population vs. Sample
 - ◆ Primary vs. Secondary data types
 - ◆ Qualitative vs. Quantitative data
 - ◆ Time Series vs. Cross-Sectional data
- Explain the difference between descriptive and inferential statistics
- Describe different sampling methods

Business Statistics: A Decision-Making Approach, 6e © 2005 Prentice-Hall, Inc. Chap 1-2



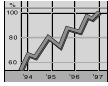

Tools of Business Statistics

- **Descriptive statistics**
 - Collecting, presenting, and describing data
- **Inferential statistics**
 - Drawing conclusions and/or making decisions concerning a population based only on sample data

Business Statistics: A Decision-Making Approach, 6e © 2005 Prentice-Hall, Inc. Chap 1-3

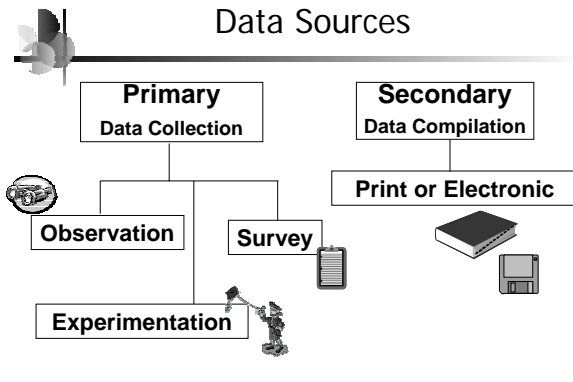
Descriptive Statistics

- **Collect data**
 - e.g. Survey, Observation, Experiments
- **Present data**
 - e.g. Charts and graphs
- **Characterize data**
 - e.g. Sample mean = $\frac{\sum x_i}{n}$



Business Statistics: A Decision-Making Approach, 6e © 2005 Prentice-Hall, Inc. Chap 1-4

Data Sources




```
graph TD; A[Primary Data Collection] --> B[Observation]; A --> C[Experimentation]; A --> D[Survey]; E[Secondary Data Compilation] --> F[Print or Electronic];
```

Business Statistics: A Decision-Making Approach, 6e © 2005 Prentice-Hall, Inc. Chap 1-5

Survey Design Steps

- **Define the issue**
 - what are the purpose and objectives of the survey?
- **Define the population of interest**
- **Formulate survey questions**
 - make questions clear and unambiguous
 - use universally-accepted definitions
 - limit the number of questions

Business Statistics: A Decision-Making Approach, 6e © 2005 Prentice-Hall, Inc. Chap 1-6




Survey Design Steps

(continued)

- Pre-test the survey
 - pilot test with a small group of participants
 - assess clarity and length
- Determine the sample size and sampling method
- Select Sample and administer the survey

Business Statistics: A Decision-Making Approach, 6e © 2005 Prentice-Hall, Inc. Chap 1-7



Types of Questions


- **Closed-end Questions**
 - Select from a short list of defined choices

Example: Major: business liberal arts
 science other
- **Open-end Questions**
 - Respondents are free to respond with any value, words, or statement

Example: What did you like best about this course?
- **Demographic Questions**
 - Questions about the respondents' personal characteristics

Example: Gender: Female Male

Business Statistics: A Decision-Making Approach, 6e © 2005 Prentice-Hall, Inc. Chap 1-8



Populations and Samples

- A **Population** is the set of all items or individuals of interest

- Examples: All likely voters in the next election
 - All parts produced today
 - All sales receipts for November
- A **Sample** is a subset of the population

- Examples: 1000 voters selected at random for interview
 - A few parts selected for destructive testing
 - Every 100th receipt selected for audit

Business Statistics: A Decision-Making Approach, 6e © 2005 Prentice-Hall, Inc. Chap 1-9

Population vs. Sample

Population	Sample
a b c d	b c
ef gh i jkl m n	g i n
o p q rs t u v w	o r u
x y z	y

Business Statistics: A Decision-Making Approach, 6e © 2005 Prentice-Hall, Inc. Chap 1-10

Why Sample?

- Less time consuming than a census
- Less costly to administer than a census
- It is possible to obtain statistical results of a sufficiently high precision based on samples.

Business Statistics: A Decision-Making Approach, 6e © 2005 Prentice-Hall, Inc. Chap 1-11

Sampling Techniques

```
graph TD; Samples[Samples] --> NonProbability[Non-Probability Samples]; Samples --> Probability[Probability Samples]; NonProbability --> Judgement[Judgement]; NonProbability --> Convenience[Convenience]; Probability --> SimpleRandom[Simple Random]; Probability --> Systematic[Systematic]; Probability --> Stratified[Stratified]; Probability --> Cluster[Cluster];
```

Business Statistics: A Decision-Making Approach, 6e © 2005 Prentice-Hall, Inc. Chap 1-12

Statistical Sampling

- Items of the sample are chosen based on known or calculable probabilities

```
graph TD; A[Probability Samples] --> B[Simple Random]; A --> C[Stratified]; A --> D[Systematic]; A --> E[Cluster]
```

Business Statistics: A Decision-Making Approach, 6e © 2005 Prentice-Hall, Inc. Chap 1-13

Simple Random Samples

- Every individual or item from the population has an equal chance of being selected
- Selection may be with replacement or without replacement
- Samples can be obtained from a table of random numbers or computer random number generators

Business Statistics: A Decision-Making Approach, 6e © 2005 Prentice-Hall, Inc. Chap 1-14

Stratified Samples

- Population divided into subgroups (called *strata*) according to some common characteristic
- Simple random sample selected from each subgroup
- Samples from subgroups are combined into one

Business Statistics: A Decision-Making Approach, 6e © 2005 Prentice-Hall, Inc. Chap 1-15

Systematic Samples

- Decide on sample size: n
- Divide frame of N individuals into groups of k individuals: $k=N/n$
- Randomly select one individual from the 1st group
- Select every k^{th} individual thereafter

N = 64
n = 8
k = 8

Business Statistics: A Decision-Making Approach, 6e © 2005 Prentice-Hall, Inc. Chap 1-16

Cluster Samples

- Population is divided into several “clusters,” each representative of the population
- A simple random sample of clusters is selected
 - All items in the selected clusters can be used, or items can be chosen from a cluster using another probability sampling technique

Population divided into 16 clusters.

Business Statistics: A Decision-Making Approach, 6e © 2005 Prentice-Hall, Inc. Chap 1-17

Key Definitions

- A **population** is the entire collection of things under consideration
 - A **parameter** is a summary measure computed to describe a characteristic of the population
- A **sample** is a portion of the population selected for analysis
 - A **statistic** is a summary measure computed to describe a characteristic of the sample

Business Statistics: A Decision-Making Approach, 6e © 2005 Prentice-Hall, Inc. Chap 1-18

Inferential Statistics

- Making statements about a population by examining sample results

Sample statistics (known) **Inference** Population parameters (unknown, but can be estimated from sample evidence)

Business Statistics: A Decision-Making Approach, 6e © 2005 Prentice-Hall, Inc. Chap 1-19

Inferential Statistics

Drawing conclusions and/or making decisions concerning a population based on sample results.

- **Estimation**
 - e.g.: Estimate the population mean weight using the sample mean weight
- **Hypothesis Testing**
 - e.g.: Use sample evidence to test the claim that the population mean weight is 120 pounds

Business Statistics: A Decision-Making Approach, 6e © 2005 Prentice-Hall, Inc. Chap 1-20

Data Types

```

    graph TD
      Data[Data] --> Qual[Qualitative (Categorical)]
      Data --> Quant[Quantitative (Numerical)]
      Qual --- QualEx[Examples: Marital Status, Political Party, Eye Color (Defined categories)]
      Quant --> Discrete[Discrete]
      Quant --> Continuous[Continuous]
      Discrete --- DiscreteEx[Examples: Number of Children, Defects per hour (Counted items)]
      Continuous --- ContinuousEx[Examples: Weight, Voltage (Measured characteristics)]
    
```

Business Statistics: A Decision-Making Approach, 6e © 2005 Prentice-Hall, Inc. Chap 1-21

Data Types

- **Time Series Data**
 - Ordered data values observed over time

- **Cross Section Data**
 - Data values observed at a fixed point in time

Business Statistics: A Decision-Making Approach, 6e © 2005 Prentice-Hall, Inc. Chap 1-22

Data Types

	Sales (in \$1000's)				
	2003	2004	2005	2006	
Atlanta	435	460	475	490	Time Series Data
Boston	320	345	375	395	
Cleveland	405	390	410	395	
Denver	260	270	285	280	

Cross Section Data

Business Statistics: A Decision-Making Approach, 6e © 2005 Prentice-Hall, Inc. Chap 1-23

Data Measurement Levels

Measurements

Ratio/Interval Data

Highest Level
Complete Analysis

↑

Rankings
Ordered Categories

Ordinal Data

Higher Level
Mid-level Analysis


↑

Categorical Codes
ID Numbers
Category Names

Nominal Data

Lowest Level
Basic Analysis

Business Statistics: A Decision-Making Approach, 6e © 2005 Prentice-Hall, Inc.



Chapter Summary

- Reviewed key data collection methods
- Introduced key definitions:
 - ◆ Population vs. Sample
 - ◆ Primary vs. Secondary data types
 - ◆ Qualitative vs. Quantitative data
 - ◆ Time Series vs. Cross-Sectional data
- Examined descriptive vs. inferential statistics
- Described different sampling techniques
- Reviewed data types and measurement levels

Business Statistics: A Decision-Making Approach, 6e © 2005 Prentice-Hall, Inc. Chap 1-25
