Chapter 1

What is statistics?
- Numerical facts:
- Average income of Kuwaiti families.
- Your monthly expenses.
- Wedding cost.
- A group of methods used to collect, organize, present, analyze, and interpret data to make more effective decisions (educated guess vs. pure guess).
- Opening a business without assessing the need for it may affect its success.
- Two fields of study:
  - Mathematical statistics.
  - Applied statistics.

Types of statistics
- Applied statistics can be divided into two areas: descriptive statistics and inferential statistics.
Chapter 1

Types of statistics

- A population is a collection of all possible individuals, objects, or measurements of interest.
- A parameter is a summary measure computed to describe a characteristic of the population.
- A sample is a portion, or part, of the population of interest.
- A statistic is a summary measure computed to describe a characteristic of the sample.
- Descriptive statistics consists of methods for organizing, displaying, and describing data in an informative way by using tables, graphs, and summary measures.
- According to bank reports, 20% of the investors in the KSE declared bankruptcy during 2007. The statistic 20 describes the number of bankruptcies out of every 100 KSE investors.

Chapter 1: An Introduction to Business Statistics

Types of statistics

- A Gallup poll found that 49% of the people in a survey knew the name of the first president of the USA. The statistic 49 describes the number out of every 100 persons who knew the answer.

Inferential statistics consists of methods that use sample results to help make decisions or predictions about a population.

- One can make some decisions about the political view of all KU students (around 15000) based on a sample of 500 students.
- The accounting department of a large firm will select a sample of invoices to check the accuracy of all the invoices of the company.
Chapter 1

Descriptive Statistics

- Collect data
  - e.g., Survey, Observation, Experiments

- Present data
  - e.g., Charts and graphs

- Characterize data
  - e.g., Sample average

\[ \bar{x} = \frac{\sum x_i}{n} \]

Chapter 1

Inferential Statistics

- Making statements about a population by examining sample results

Sample statistics \(\rightarrow\) Inference \(\rightarrow\) Population parameters

Sample statistics \(\text{(known)}\) \(\rightarrow\) Inference \(\text{(unknown, but can be estimated from sample evidence)}\)
Chapter 1

Remember

- **Descriptive statistics**
  - Collecting, presenting, and describing data

- **Inferential statistics**
  - Drawing conclusions and/or making decisions concerning a population based only on sample data

**Chapter 1**

- **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

QM-120, M. Zainal
Population versus sample

- Suppose a statistician is interested in knowing:
  - The 2004 gross sale of all companies in Kuwait.
  - The prices of all houses in Mishrif.
- All companies and all houses are the target population for each case.
- We can make our decision based on a portion of the population (sample).
- USA presidential election polls are based on few hundred voters instead of 205,018,000 voters.

Why we 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.
Chapter 1

Population versus sample

- **Census and sample survey**
  - A survey that includes every member of the population is called a *census*. The technique of collecting information from a portion of the population is called a *sample survey*.

- **Representative sample**
  - A sample that represents the characteristics of the population as closely as possible is called a *representative sample*.

- **Random Sample**
  - A sample drawn in such a way that each element of the population has the same chance of being selected is called a *random sample*.
  - If the chance of being selected is the same for each element of the population, it is called a *simple random sample* (SRS).

QM-120, M. Zainal

Chapter 1

Applications in business and economics

- Successful managers and decision-makers should understand and use statistics effectively.

- Examples of the uses of statistics in business and economics are:
  - Accounting: Sample of balance sheets to audit.
  - Finance: Comparing stocks.
  - Marketing: Understanding the relationship between promotions and sales.
  - Production: Quality control charts.
  - Economics: Forecasting unemployment rate.
  - Insurance: Finding premiums of a policy

QM-120, M. Zainal
Chapter 1

Basic terms

- Element, Variable, observation, and data set.
  - An element or member of a sample or population is a specific subject or object (person, firm, item, country... etc.).
  - A variable is a characteristic under study that assumes different values for different element.
  - The value of a variable for an element is called observation or measurement.

<table>
<thead>
<tr>
<th>Company</th>
<th>2001 Sales (million of dollars)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IBM</td>
<td>85,866</td>
</tr>
<tr>
<td>Dell Computer</td>
<td>31,168</td>
</tr>
<tr>
<td>GM</td>
<td>177,260</td>
</tr>
</tbody>
</table>

An element

2001 Sales of three U.S. Companies

Data set

Chapter 1

Types of variables

- Quantitative Variable:
  - When the characteristic being studied can be reported numerically, it is called a quantitative variable.
  - Example: The number of students in each section of the Business Statistics course; the distance students travel from home to CBA; the number of children in a family.

- Qualitative or categorical Variable:
  - When the characteristic being studied is nonnumeric, it is called a qualitative variable.
  - Example: A classification of university students by gender or by program (Business, Education, Arts, etc.) is an example of a qualitative variable. Type of automobile owned, and eye color are also qualitative.
Chapter 1

Types of variables

- **Quantitative Variables** (data that can be reported numerically) can be classified as either Discrete or Continuous.
  - **Discrete Variable**
    - A quantitative variable that can only assume certain values.
    - *Example*: The number of bedrooms in a house, the number of hamburgers sold at Burger King today, the number of children in a family. Usually discrete variables result from counting. The number of children in a family can be 2 or 3 but not 2.45
  - **Continuous Variable**
    - A quantitative variable that can assume any value within a specified range.

- **Example**: The amount of rain in Kuwait last winter (it could be 20.55 cm), the height of students in a class, tire pressure, time to do an assignment, a persons weight.
Levels of variables

Another way to classify data is by the way it is measured.

Nominal level of measurement

- A nominal level of measurement data strictly with qualitative data.
- Observations are simply assigned to predetermined categories.
- This data type does not allow us to perform any mathematical operations, such as adding or multiplying.
- Number can be used at the nominal level but they can’t be added or placed in a meaningful order of greater than or less than.
- This type is considered the lowest level of data.

Ordinal level of measurement

- Ordinal is the next level up.
- It has all the properties of nominal data with the added feature that we can rank-order the values from highest to lowest.
- An example is if you were to have a cooking race. Let’s say the finishing order was Scott, Tom, and Bob. We still can’t perform mathematical operations on this data, but we can say that Scott’s cooking was faster than Bob’s.
- However, we cannot say how much faster.
- Ordinal data does not allow us to make measurements between the categories and to say, for instance, that Scott’s cooking is twice as good as Bob’s.
- Ordinal data can be either qualitative or quantitative.
Chapter 1

Levels of variables

- Interval level of measurement
  - Moving up the scale of data, we find ourselves at the interval level, which is strictly quantitative data.
  - Now we can get to work with the mathematical operations of addition and subtraction when comparing values.
  - For this data, the difference between the different categories can be measured with actual numbers and also provides meaningful information.
  - Temperature measurement in degrees Fahrenheit is a common example here. For instance, 70 degrees is 5 degrees warmer than 65 degrees.
  - However, multiplication and division can’t be performed on this data. Why not? Simply because we cannot argue that 100 degrees is twice as warm as 50 degrees.

QM-120, M. Zainal

---

Chapter 1

Levels of variables

- Ratio level of measurement
  - The king of data types is the ratio level.
  - This is as good as it gets as far as data is concerned.
  - Now we can perform all four mathematical operations to compare values with absolutely no feelings of guilt.
  - Examples of this type of data are age, weight, height

QM-120, M. Zainal
Chapter 1

Data Measurement Levels

- **Measurements**
  - Ratio/Interval Data
    - Highest Level
      - Complete Analysis

- **Rankings**
  - Ordinal Data
    - Higher Level
      - Mid-level Analysis

- **Categorical Codes**
  - Nominal Data
    - Lowest Level
      - Basic Analysis

Chapter 1

Cross-section versus time-series data

- **Cross-section data:**
  - Data collected on different elements at the same point of time or for the same period of time are called cross-section data.

<table>
<thead>
<tr>
<th>2001 Sales of three U.S. Companies</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Company</strong></td>
</tr>
<tr>
<td>IBM</td>
</tr>
<tr>
<td>Dell Computer</td>
</tr>
<tr>
<td>GM</td>
</tr>
</tbody>
</table>

QM-120, M. Zainal
Cross-section versus time-series data

- **Time-series data:**
  - Data collected on the same element for the same variable at different points in time or for different periods of time are called *time-series data*.

<table>
<thead>
<tr>
<th>Year</th>
<th>Super Bowl</th>
<th>Price of a 30-second TV commercial (million of dollars)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1999</td>
<td>XXXIII</td>
<td>1.60</td>
</tr>
<tr>
<td>2000</td>
<td>XXIV</td>
<td>2.10</td>
</tr>
<tr>
<td>2001</td>
<td>XXXV</td>
<td>2.05</td>
</tr>
<tr>
<td>2002</td>
<td>XXXVI</td>
<td>1.90</td>
</tr>
<tr>
<td>2003</td>
<td>XXXVII</td>
<td>2.20</td>
</tr>
</tbody>
</table>

Price of a 30-second TV commercial during Super Bowl telecast

---

**Chapter 1**

Cross-section versus time-series data

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

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

<table>
<thead>
<tr>
<th></th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atlanta</td>
<td>435</td>
<td>460</td>
<td>475</td>
<td>490</td>
</tr>
<tr>
<td>Boston</td>
<td>320</td>
<td>345</td>
<td>375</td>
<td>395</td>
</tr>
<tr>
<td>Cleveland</td>
<td>405</td>
<td>390</td>
<td>410</td>
<td>395</td>
</tr>
<tr>
<td>Denver</td>
<td>260</td>
<td>270</td>
<td>285</td>
<td>280</td>
</tr>
</tbody>
</table>
Chapter 1

Sources of data

- Internal sources.
  - Company’s own personnel file.
  - Accounting records.
    - Example: A company that wants to forecast the future sale of its product may use the data of past periods from its own record.

- External sources.
  - Governmental publications.
  - Private publications.
  - Survey or experiment.

Survey or experiment

Experiments

Written questionnaires

Telephone surveys

Direct observation and personal interview
Statistics and Ethics

- People often use statistics to persuade others to their opinions, it can lead to the misuse of statistics in several ways:
  - Choosing a sample that ensures results in favoring your desired outcome (*biased sample*)
  - Making the differences seems greater than they actually look in graphs

Internet polls are unreliable because those websites have no control over the respondents or how many times they respond.