

# 2

## CHAPTER

# Data, Reality, and Problem Solving

- 2.1 The Lords of Data
- 2.2 Data Classification
- 2.3 Time Series Data vs.  
Cross-Sectional Data
- CR Chapter Review

When you encounter data, ask yourself: is the data credible? Consider the following questions:

1. Is the concept under study adequately reflected by the proposed measurements?
2. Is the data measured accurately?
3. Is there a sufficient quantity of the data to draw a reasonable conclusion?

# The Scientific Method

1. Gather information about the phenomenon being studied;
2. On the basis of the data, formulate a preliminary generalization or hypothesis;
3. Collect further data to test the hypothesis;
4. If the data and other subsequent experiments support the hypothesis, it becomes a law.

**PROCEDURE**

## Decision-Making Method

1. Clearly define the problem and any influential variables.
2. Decide upon objectives and decision criteria for choosing a solution.
3. Create alternative solutions.
4. Compare alternatives using the criteria established in the second step.
5. Implement the chosen alternative.
6. Check the results to make sure the desired results are achieved.

**PROCEDURE**

## Confounding Variables

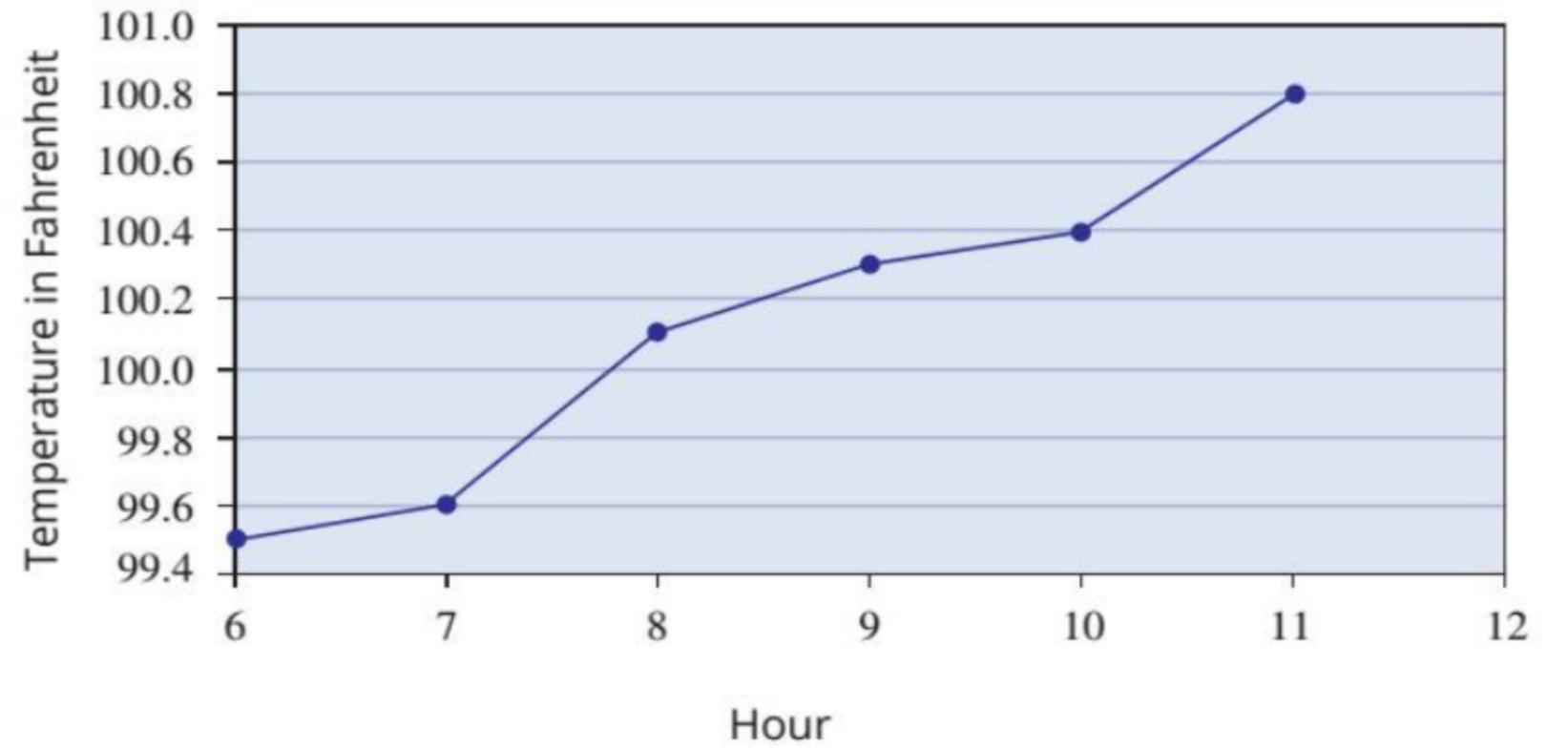
**Confounding variables** are “extra” variables that are not accounted for during experimentation and can cause results to become skewed.

**DEFINITION**



Figure 2.1.1

### Patient's Temperature History



# Collecting Data

Essentially, there are two ways to obtain data: **observation** and **controlled experiments**.

## Response Variable

A **response variable** measures the outcome of interest in a study.

DEFINITION

## Explanatory Variable

An **explanatory variable** causes or explains changes in a response variable.

DEFINITION

Suppose a new species of tomato has been genetically engineered to increase yields. The question: Does the new species produce higher yields?

<b>Plot 1</b>	<b>Plot 2</b>	<b>Plot 3</b>	<b>Plot 4</b>
<b>Plot 5</b>	<b>Plot 6</b>	<b>Plot 7</b>	<b>Plot 8</b>
<b>Plot 9</b>	<b>Plot 10</b>	<b>Plot 11</b>	<b>Plot 12</b>
<b>Plot 13</b>	<b>Plot 14</b>	<b>Plot 15</b>	<b>Plot 16</b>

## Example 2.1.2

---

Does an SAT preparation course improve performance on the SAT?

# Placebos

A **placebo** is a fake treatment that has the potential to cause a response.

**DEFINITION**

London scientists conducted a study to determine if chocolate can trigger migraines. Twelve migraine-prone subjects were given a peppermint-laced chocolate candy and eight migraine-prone subjects were given a peppermint-laced placebo made of carob, peppermint, and vegetable fat. Five subjects from the group given chocolate developed a migraine headache within one day. No one from the group given the placebo developed a migraine in the same time period.<sup>2</sup>

- a.** Which phase of the Scientific Method best describes this study?
- b.** Is this an observational study or a controlled experiment?
- c.** What is the response variable?
- d.** What is the explanatory variable?
- e.** Which group is the treatment group?
- f.** Which group is the control group?

## Discrete

Data in which the observations are restricted to a set of values (such as 1, 2, 3, 4) that possess gaps is called **discrete**.

**DEFINITION**

# Continuous

Data that can take on any value within some interval is called **continuous**.

**DEFINITION**

## Level of Measurement

The quality of data is referred to as its **level of measurement**.

**DEFINITION**

The terms used to describe the quality of data are nominal, ordinal, interval, and ratio.

## Nominal

Data that represents whether a variable possesses some characteristic is called **nominal**.

**DEFINITION**

## Ordinal Data

**Ordinal data** represents categories that have some associated order.

**DEFINITION**

*Frosty Pops taste* \_\_\_\_\_.

1–very bad    2–bad    3–fair    4–good    5–very good

## Interval

If the data can be ordered and the arithmetic difference is meaningful, the data is **interval**.

**DEFINITION**

$$48 \text{ degrees} - 45 \text{ degrees} = 3 \text{ degrees}$$

$$72 \text{ degrees} - 69 \text{ degrees} = 3 \text{ degrees}$$

## Ratio Data

**Ratio data** is similar to interval data, except that it has a meaningful zero value.

**DEFINITION**

Is money a ratio variable?

Say a friend had \$40 and you had \$20.

$$\frac{\$40}{\$20} = 2$$

According to the ratio we just computed your friend has twice as much money as you. Is this really true? Money is a ratio variable because ratios (quotients) are meaningful. If someone does have \$40 and you have \$20, they do have twice as much money as you.

## Qualitative

**Qualitative** data is measured on a nominal or ordinal scale.

**DEFINITION**

## Quantitative

**Quantitative** data is measured on an interval or ratio scale.

**DEFINITION**

Determine the level of measurement (nominal, ordinal, interval, or ratio) for each of the following variables.

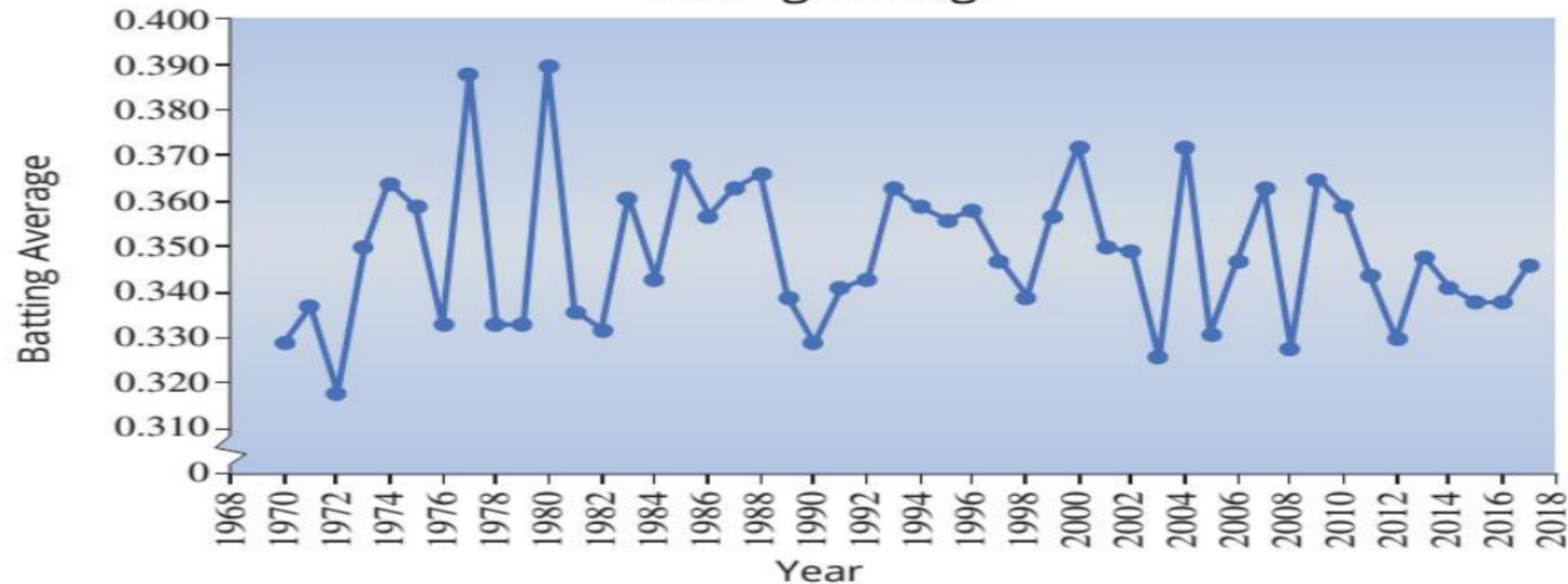
- a.** The time it takes for a student to complete an exam.
- b.** Majors of randomly selected students at a university.
- c.** The category which best describes how frequently a person eats chocolate:  
Frequently, Occasionally, Seldom, Never.
- d.** The number of pounds of snack food eaten by an individual in his or her lifetime.

# Time Series

**Time series** data originates as measurements usually taken from some process over equally spaced intervals of time.

**DEFINITION**

American League Batting Champions  
Batting Average



**Table 2.3.2 - Life Expectancy at Birth 2015**

Country	Life Expectancy
Afghanistan	61
Australia	83
Botswana	66
Egypt	71
Guatemala	72
Japan	84
Kenya	63
Sierra Leone	46
Spain	83
Sri Lanka	75
Sweden	82
United Kingdom	81
United States	79

## Cross-Sectional Data

### Cross-Sectional Data

**Cross-sectional data** are measurements created at approximately the same period of time.

**DEFINITION**