

---

# Statistics Success Stories and Cautionary Tales

---

1

## 1.1 What is Statistics?

*Statistics* is a collection of procedures and principles for gathering data and analyzing information in order to help people make decisions when faced with uncertainty.

2

## 1.2 Seven Statistical Stories With Morals



- **Case Study 1.1: Who Are Those Speedy Drivers?**
- **Case Study 1.2: Safety in the Skies**
- **Case Study 1.3: Did Anyone Ask Whom You've Been Dating?**
- **Case Study 1.4: Who Are Those Angry Women?**
- **Case Study 1.5: Does Prayer Lower Blood Pressure?**
- **Case Study 1.6: Does Aspirin Reduce Heart Attack Rates?**
- **Case Study 1.7: Does the Internet Increase Loneliness and Depression?**

3

## First, let's take a survey!



Please take a few minutes to take the survey given at <http://math.fullerton.edu/mori/Math120/survey1.pdf>

We will use the result of this survey in learning some of our course material.

Let's enter the data in Minitab. The instructions are given in the Minitab Manual that is contained in the CD that comes with your textbook.

Jamshidian, 2005

4

## Case Study 1.1 *Who Are Those Speedy Drivers?*

**Question:** What's the fastest you have ever driven a car? \_\_\_\_\_ mph.

**Data:** 87 male and 102 female students from large statistics class at University.

**Males:** 110 109 90 140 105 150 120 110 110 90 115 95 145 140 110 105 85 95 100 115 124 95 100 125 140 85 120 115 105 125 102 85 120 110 120 115 94 125 80 85 140 120 92 130 125 110 90 110 110 95 95 110 105 80 100 110 130 105 105 120 90 100 105 100 120 100 100 80 100 120 105 60 125 120 100 115 95 110 101 80 112 120 110 115 125 55 90

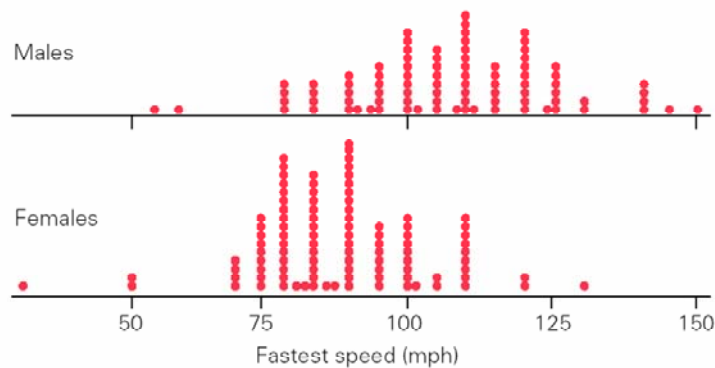
**Females:** 80 75 83 80 100 100 90 75 95 85 90 85 90 90 120 85 100 120 75 85 80 70 85 110 85 75 105 95 75 70 90 70 82 85 100 90 75 90 110 80 80 110 110 95 75 130 95 110 110 80 90 105 90 110 75 100 90 110 85 90 80 80 85 50 80 100 80 80 80 95 100 90 100 95 80 80 50 88 90 90 85 70 90 30 85 85 87 85 90 85 75 90 102 80 100 95 110 80 95 90 80 90

*Which gender has driven faster? How to summarize data?*

5

## Case Study 1.1 *Who Are Those Speedy Drivers?*

**Dotplot:** each dot represents the response of an individual student.



6

## Case Study 1.1 *Who Are Those Speedy Drivers?*

**Five-number summary:** the lowest value, the cutoff points for  $\frac{1}{4}$ ,  $\frac{1}{2}$ , and  $\frac{3}{4}$  of the data, and the highest value.

	Males (87 Students)		Females (102 Students)	
<b>Median</b>	110		89	
<b>Quartiles</b>	95	120	80	95
<b>Extremes</b>	55	150	30	130

Note:  $\frac{3}{4}$  of men have driven 95 mph or more, only  $\frac{1}{4}$  of women have done so.

**Moral:** *Simple summaries of data can tell an interesting story and are easier to digest than long lists.*

7

## Case Study 1.2 *Safety in the Skies?*

**“Planes get closer in midair as traffic control errors rise”**

**“Errors by air traffic controllers climbed from 746 in fiscal 1997 to 878 in fiscal 1998, an 18% increase.”**

*USA Today, Levin, 1999*

Sounds ominous, but “The errors per million flights handled by controllers climbed from 4.8 to 5.5.”

So the original **rate** of errors in 1997, from which the 18% increase was calculated, was less than 5 errors per million flights.

**Moral:** *When discussing the change in the rate or risk of occurrence of something, make sure you also include the base rate or baseline risk.*

8

### Case Study 1.3 *Did Anyone Ask Whom You've Been Dating?*

“According to a new *USA Today*/Gallup Poll of teenagers across the country, 57 percent of teens who go out on dates say they’ve been out with someone of another race or ethnic group.” (Peterson, 1997).

*Sacramento Bee* headline read:  
 “Interracial dates common among today’s teenagers.”

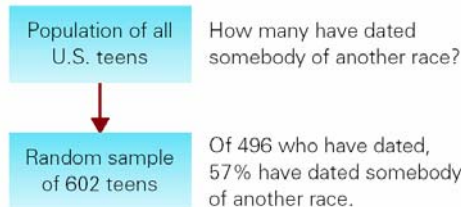
- Millions of teenagers in U.S. -- Did polltakers ask all of them? **No.**
- The article states “the results of the new poll of 602 teens, conducted Oct. 13–20, reflect the ubiquity of interracial dating today...”
- Asked only 602 teens. Could such a small sample tell us anything about the millions of teenagers in the U.S.? **Yes...**  
 if those teens constituted a *random sample* from the *population*.



### Case Study 1.3 *Did Anyone Ask Whom You've Been Dating?*

How *accurate* could this sample be?

**Margin of error** is about 5%.



The percent of all teenagers in the US who date that would say they have dated interracially is likely to be in the range  $57\% \pm 5\%$ , or between 52% and 62%.

The 5% given above is the *margin of error*. A quick estimate of the margin error is  $1/\sqrt{n}$ . For example,  $1/\sqrt{496} = .045$ .





### Case Study 1.3 *Did Anyone Ask Whom You've Been Dating?*



**Moral:** *A representative sample of only a few thousand, or perhaps even a few hundred, can give reasonably accurate information about a population of millions.*

11



### Case Study 1.4 *Who Are Those Angry Women?*



Shere Hite sent questionnaires to 100,000 women asking about love, sex, and relationships. Only 4.5% of the women responded, and Hite used those responses to write her book, *Women and Love*.

“The women who responded were fed up with men and eager to fight them. For example, 91% of those who were divorced said that they had initiated the divorce. The anger of women toward men became the theme of the book.” Moore (1997, p. 11).

Extensive **nonresponse** from a random sample, or the use of a **self-selected** (i.e., **all-volunteer**) sample, will probably produce **biased** results.

**Moral:** *An unrepresentative sample, even a large one, tells you almost nothing about the population.*

12

## Case Study 1.5 *Does Prayer Lower Blood Pressure?*

“Attending religious services lowers blood pressure more than tuning into religious TV or radio, a new study says”

*USA Today* headline read:

“Prayer can lower blood pressure.” (Davis, 1998)

Based on *observational study*, followed 2391 people 6 years.

“People who attended a religious service once a week and prayed or studied the Bible once a day were 40% less likely to have high blood pressure than those who don’t go to church every week and prayed and studied the Bible less.”

Researchers *did observe a relationship*, but it’s a **mistake to conclude** prayer actually *causes* lower blood pressure.

13

## Case Study 1.5 *Does Prayer Lower Blood Pressure?*

In observational studies, **groups can differ** by important ways that may contribute to the observed relationship.

People who attended church regularly may have ...

- been less likely to smoke or drink alcohol;
- had a better social network;
- been somewhat healthier and able to go to church.

These other factors are possible *confounding variables*.

**Moral:** *Cause-and-effect conclusions cannot generally be made based on an observational study.*

14

## Case Study 1.6 *Does Aspirin Reduce Heart Attack Rates?*

### Physician's Health Study (1988)

5-year *randomized experiment* ...

- 22,071 male physicians of age 40 - 84;
- *randomly assigned* to one of two *treatment* groups;
- Group 1 = aspirin every other day; Group 2 = *placebo*;
- Physicians *blinded* as to which group they were in.

15

## Case Study 1.6 *Does Aspirin Reduce Heart Attack Rates?*

TABLE 1.1 The Effect of Aspirin on Heart Attacks

Treatment	Heart Attacks	Doctors in Group	Attacks Per 1000 Doctors
Aspirin	104	11,037	9.42
Placebo	189	11,034	17.13

Aspirin group: 9.42 heart attacks per 1000 participants

Placebo group: 17.13 heart attacks per 1000 participants

**Randomization** => other important factors (age, diet, etc) should have been similar for both groups. Only important difference should be whether they took aspirin or placebo.

**Moral:** *Unlike with observational studies, cause-and-effect conclusions can generally be made on the basis of randomized experiments.*

16

## Case Study 1.7 *Does the Internet Increase Loneliness and Depression?*

**“greater use of the Internet was associated with declines in participants’ communication with family members in the household, declines in size of their social circle, and increases in their depression and loneliness” (Kraut et al., 1998, p. 1017)**

***New York Times* headline read:**

**“Sad, Lonely World discovered in Cyberspace.” (Harmon, 1998)**

Study included 169 individuals from 73 households in Pittsburgh given free computers and internet service.

Participants answered questions at beginning and either 1-2 years later on social contacts, stress, loneliness, depression.

17

## Case Study 1.7 *Does the Internet Increase Loneliness and Depression?*

Changes were quite small, although significant ...

- # of people in “local social network” decreased from average of 23.94 to 22.90 people;
- on scale 1 to 5, self-reported loneliness decreased from average of 1.99 to 1.89 (lower => greater loneliness).
- on a scale 0 to 3, self-reported depression dropped from average of 0.73 to 0.62 (lower => higher depression).

**Follow-up study: “Using the Internet at home doesn’t make people more depressed and lonely after all.” (Elias, 2001)**

Whether the link ever existed will never be known, but it is not surprising, given the small magnitude of the original finding, that it subsequently disappeared.

18

### Case Study 1.7 *Does the Internet Increase Loneliness and Depression?*



**Moral:** A “statistically significant” finding does not necessarily have practical importance. When a study reports a statistically significant finding, find out the magnitude of the relationship or difference.

A secondary moral to this story is that the implied direction of cause and effect may be wrong. In this case, it could be that people who were more lonely and depressed were more prone to use the Internet.

And, as the follow-up research makes clear, remember that “truth” doesn’t necessarily remain fixed across time.

19

### 1.3 The Common Elements in the Seven Stories



In every story, *data are used to make a judgment about a situation.*

This is what *statistics* is all about.

20

## The Discovery of Knowledge



1. *Asking the right question(s).*
2. *Collecting useful data, which includes deciding how much is needed.*
3. *Summarizing and analyzing data, with the goal of answering the questions.*
4. *Making decisions and generalizations based on the observed data.*
5. *Turning the data and subsequent decisions into new knowledge.*

21

- Class exercises
  - Page 10, #22
  - Page 11, #26

22