

STATISTICS: THE ART & SCIENCE OF LEARNING FROM DATA

Chapter 1

PART

1

Gathering and Exploring Data

Chapter 1
Statistics: The Art
and Science of
Learning from Data



Chapter 2
Exploring Data
with Graphs
and Numerical
Summaries



Chapter 3
Association:
Contingency,
Correlation, and
Regression



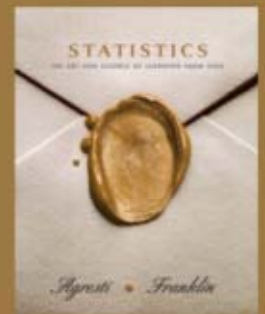
Chapter 4
Gathering Data



- 1.1 How Can You Investigate Using Data?
- 1.2 We Learn about Populations Using Samples
- 1.3 What Role Do Computers Play in Statistics?



Math 1530
Elementary Statistics
Austin Peay State University



EXAMPLE 1

How Can Statistics Help Us Learn about the World?



How can we evaluate evidence against global warming?

Are cell phones dangerous?

What are the chances of a tax return being audited?

How likely are we to win the lottery?

Is there bias in appointing female managers?

EXAMPLE 1

How Can Statistics Help Us Learn about the World?



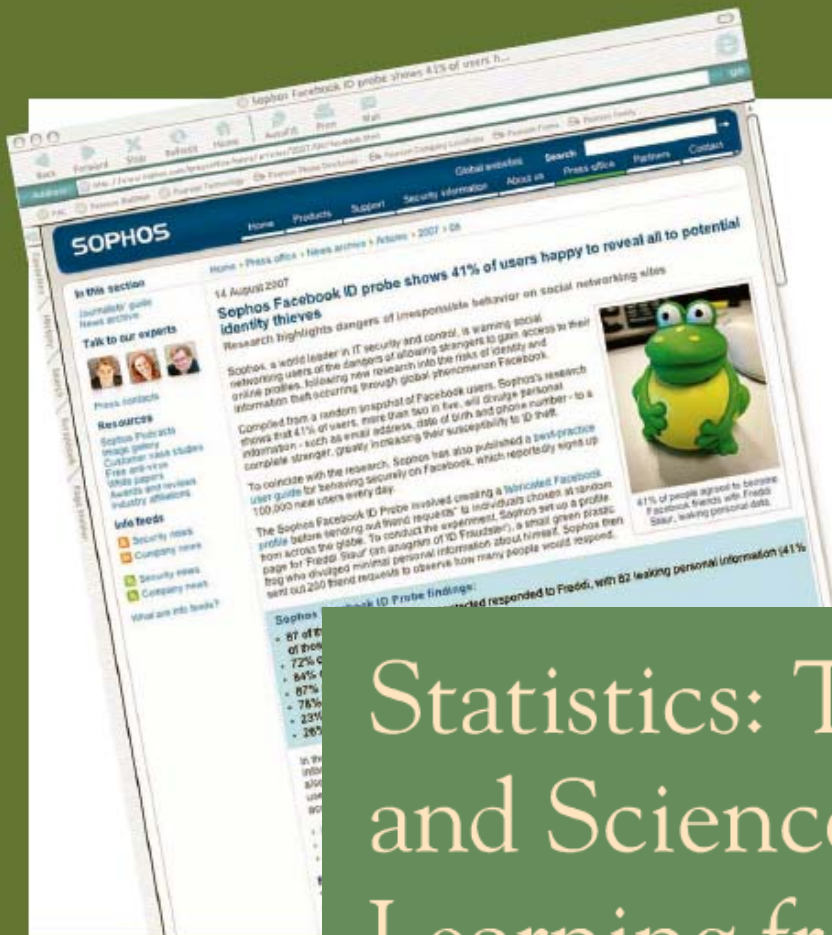
What hot streaks should we expect in basketball?

How can we analyze whether a diet works?

Are astrology predictions more accurate than guessing?

Do more people believe in heaven or hell?

What predicts whether or not we are happy?



Statistics: The Art and Science of Learning from Data

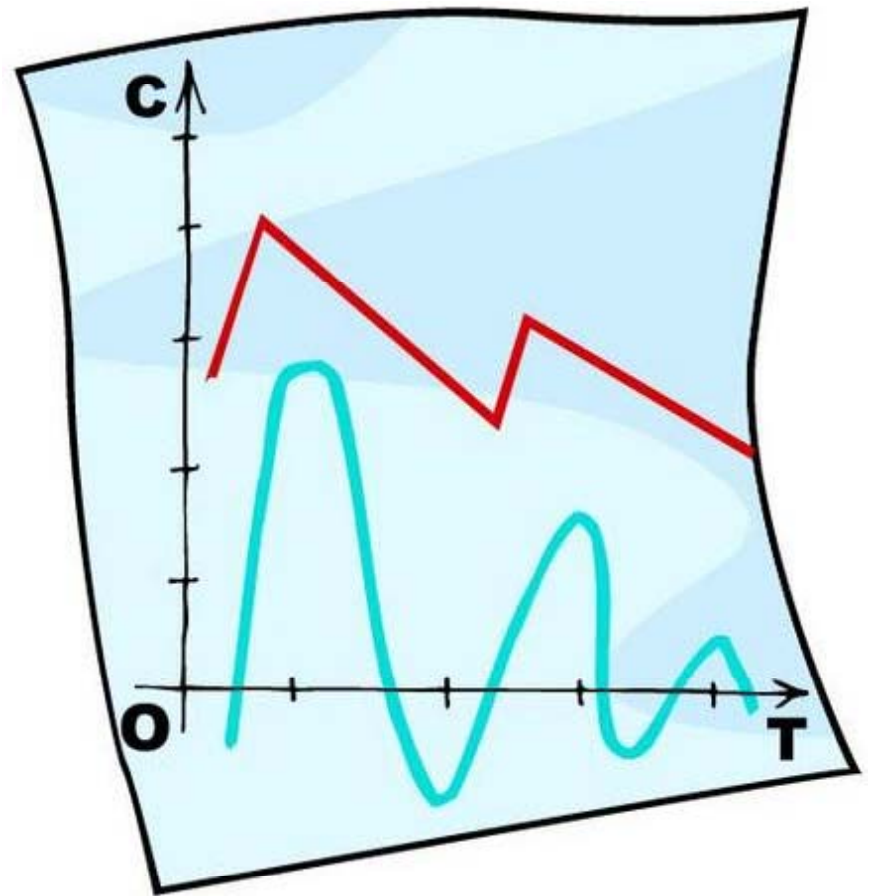
1

1.1 How Can You Investigate Using Data?

Learning Objectives

Objectives include:

1. Understand the usefulness of **data**.
2. Define **statistics**.
3. Identify **three aspects** of a **study**.



Data



beanactivist.files.wordpress.com

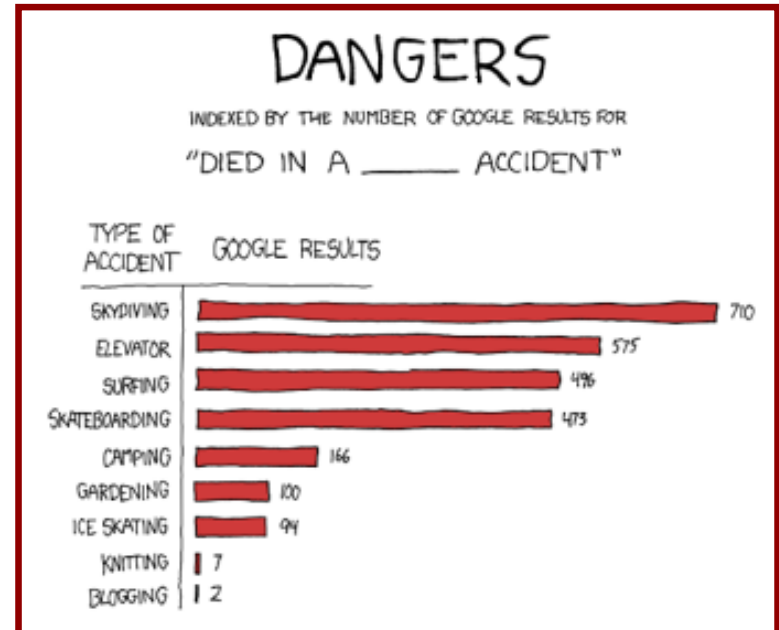
Data is **information we gather** through experiments and surveys.

1. Experiment on low carb diet
 - ▣ Data: weight of subjects before and after
2. Survey on effectiveness of a TV ad
 - ▣ Data: percentage who went to Starbucks since ad aired

Statistics

Statistics is the art and science of:

1. **Designing** studies,
2. **Analyzing** resulting data, and
3. **Translating** data into knowledge and understanding.



xkcd

Three Aspects of a Study



www.icts.uiowa.edu

1. **Design:** Planning how to obtain data
2. **Description:** Summarizing the data
3. **Inference:** Making decisions and predictions

1st Aspect of a Study: Design

How do we **conduct the experiment** or **select people** for the **survey** to insure trustworthy results?

Design Examples:

1. Planning data collection to study effects of Vitamin E on athletic strength
2. For a marketing survey, selecting people to provide proper coverage



fineartamerica.com

2nd Aspect of a Study: **Description**



www.emecogroup.org

Summarize raw data and present in useful formats (e.g., average, charts or graphs)

Description Examples:

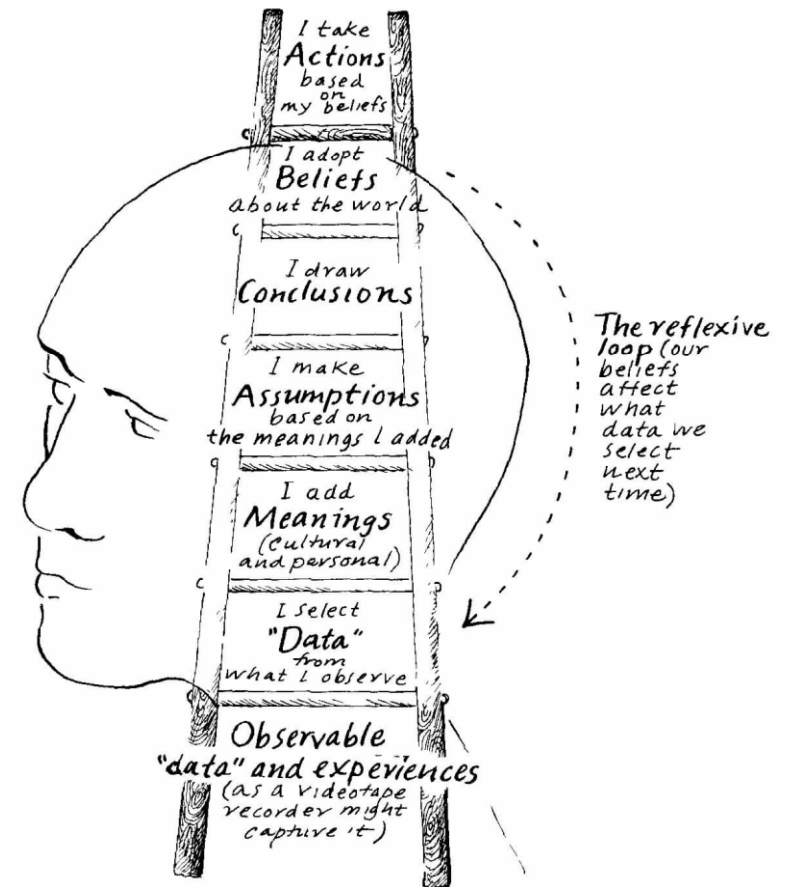
- ❑ A graph showing total precipitation in Clarksville for each month of 2005
- ❑ Average age of students in a statistics class is 25 years

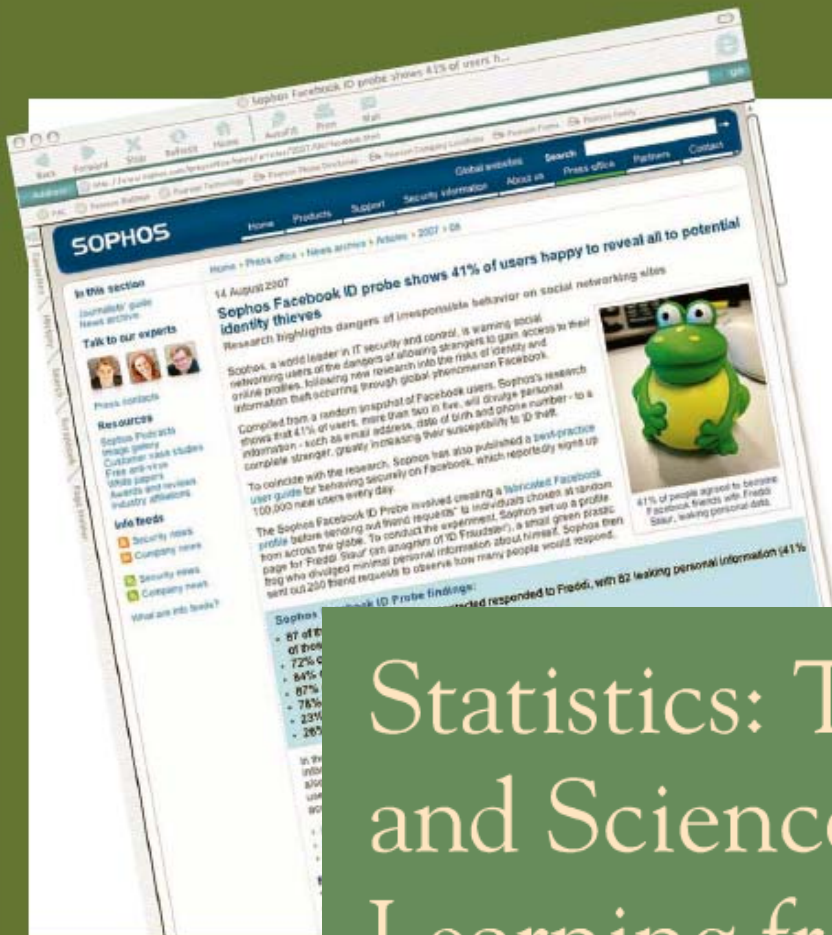
3rd Aspect of a Study: Inference

Make decisions or predictions
based on the data

Inference Examples:

- ▣ Relationship between smoking cigarettes and getting emphysema
- ▣ 47% of the registered voters in Illinois will vote in the primary





Statistics: The Art and Science of Learning from Data

1

1.2 We Learn about Populations Using Samples

Learning Objectives



www.popcom.gov.ph

Objectives include defining and differentiating:

1. Subjects
2. Population and sample
3. Descriptive statistics and inferential statistics
4. Sample Statistics and Population Parameters
5. Randomness and Variability

Subjects

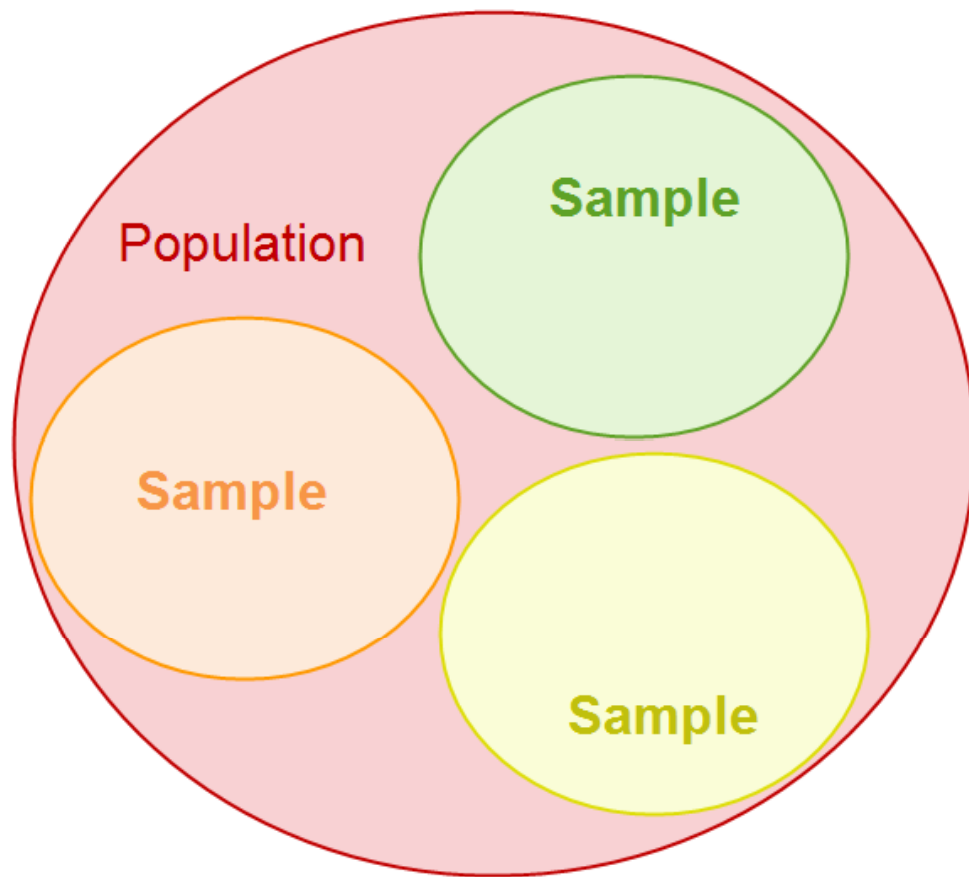
Subjects - The entities that we measure in a study

Subjects could be

1. individuals,
2. schools,
3. rats,
4. counties,
5. widgets



Population and Samples



- **Population:** All subjects of interest
- **Sample:** Subset of the population for whom we have data

Sample & Population for an Exit Poll

- In California in 2003, a special election was held to consider whether Governor Gray Davis should be recalled from office.
- An exit poll sampled 3160 of the 8 million people who voted. Define the **sample** and the **population** for this exit poll.



ltn.co.uk

Descriptive vs. Inferential Statistics



mallimages.mallfinder.com

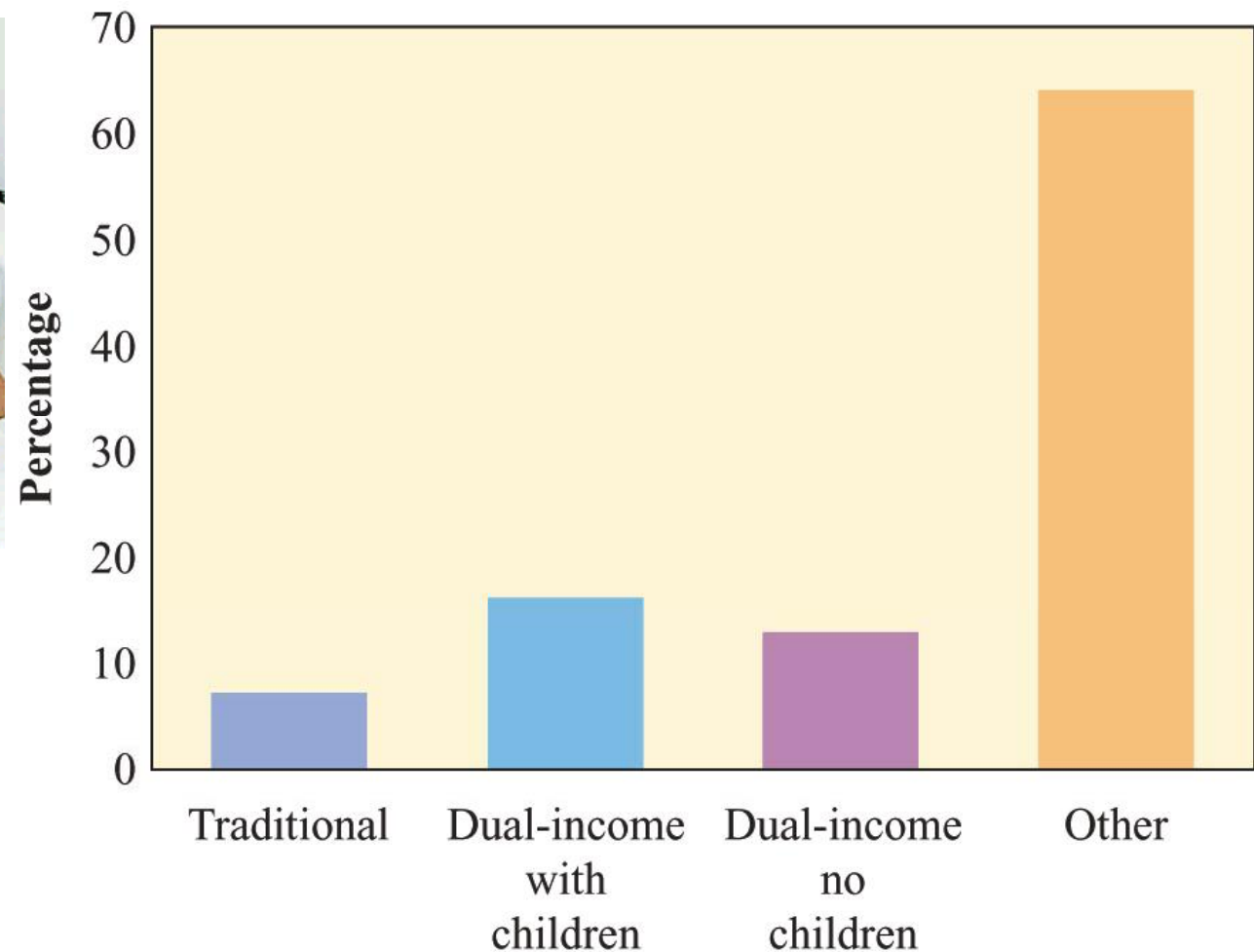
- **Descriptive** statistics summarize data – graphs and numbers such as averages and percentages
- **Inferential** statistics make decisions or predictions about a population based on data obtained from a sample of that population.

Descriptive Statistics Example



www.freedomsphoenix.com

Types of U.S. Households



Inferential Statistics Example

By surveying 1 000 likely voters, we find **39%** who approve of the job President Bush is doing.

- We are 95% confident that the population proportion of likely voters who approve of the job President Bush is doing is between **36%** and **42%**.



Sample Statistics & Population Parameters

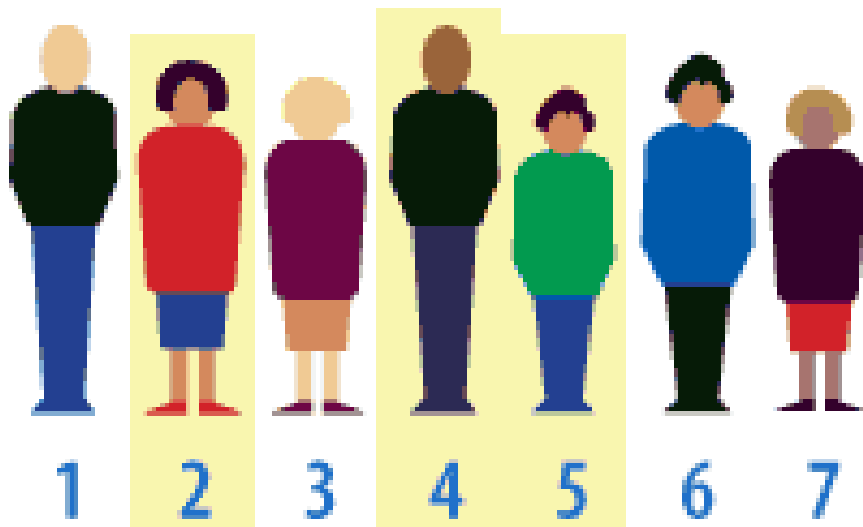
Parameter and Statistic

A **parameter** is a numerical summary of the population.

A **statistic** is a numerical summary of a sample taken from the population.



Randomness



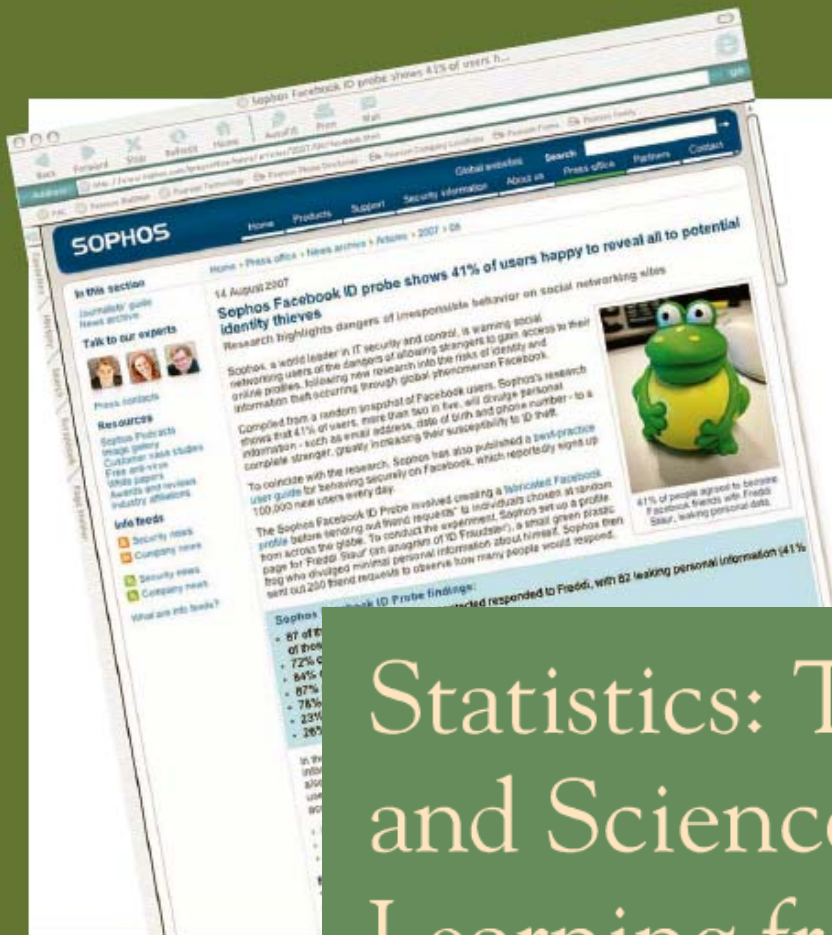
Assign Numbers,
Auto-Generate Random
Selections

- **Simple Random Sampling:**
each subject in the population has the same chance of being included in the sample
- Randomness is crucial to insuring that the sample is representative of the population so that powerful inferences can be made

Variability

- Measurements may vary from subject to subject, and
 - Measurements may vary from sample to sample.
- Predictions will therefore be more accurate for larger samples.





Statistics: The Art and Science of Learning from Data

1

1.3 What Role Do Computers Play in Statistics?

What Role Do Computers Play in Statistics?



www.masternewmedia.org

- **Data files** - Large data sets organized in a spreadsheet format known as a data file
 - ▣ Each row contains measurements for a particular subject and column for a particular characteristic
- **Databases** – An existing archive collection of data files

Sources should always be checked for reliability.

What Role Do Computers Play in Statistics?

- **Applets** – A short application program for performing a specific task
 - ▣ Useful for performing activities that illustrate the ideas of statistics

People still select valid analyses.



www.atelier-us.com

Image Sources

Statistics: The Art and Science of Learning from Data, 2nd Edition, Agresti and Franklin

<http://www-dse.ec.unipi.it/persona/ricercatori/Guerrazzi/Pictures/Data.jpg>

http://beanactivist.files.wordpress.com/2008/07/starbucks_escher.jpg

<http://philippe.gambette.free.fr/Blog/Xkcd/Xkcd369.png>

<http://www.icts.uiowa.edu/images/events/stats.jpg>

<http://fineartamerica.com/images-medium/geometric-design-number-seven-jason-nelson.jpg>

http://www.emecogroup.org/media/17356/graphs_charts.jpg

http://www.reply-mc.com/UserFiles/Image/ladder_of_inference%281%29.jpg

<http://www.popcom.gov.ph/regions/03/popchart.jpg>

<http://kiriko-moth.com/blog/wp-content/uploads/2009/09/ages.jpg>

<http://itn.co.uk/storyf2fbd7b06ca0af000a279e06c853465a.jpg>

http://mallimages.mallfinder.com/images/customheader///Guess/guess_store_ATS.jpg

<http://www.freedomsphoenix.com/Uploads/Graphics/101-0824211022-happy-family.gif>

<http://bigjournalism.com/files/2010/01/george-bush-frowning.jpg>

<http://www.nedarc.org/nedarc/collectingData/participantSelection/probabilitySampling/images/simpleSample.gif>

<http://www.pinguicula.org/images/divers/variation3%28HR%29.jpg>

http://www.masternewmedia.org/images/technology_stockxpertcom_id1218361_size1.jpg

http://www.atelier-us.com/upload/gallery2/2008/news/online_student.jpg