


Stat 104: Quantitative Methods
Class 2: Getting Good Data



## Populations and Samples

- Consider the following research questions:

1) What is the average mercury content in sword fish in the Atlantic Ocean?
2) Over the last 5 years, what is the average time to degree for UCLA undergraduate students?
3) Does the drug sulphinpyrazone reduce the number of deaths in heart attack patients?
4) Single gender social organizations are bad.



## Old School: Anecdotal Evidence

- Let's consider the following possible responses to our three research questions:

1) A man on the news got mercury poisoning from eating swordfish, so the average mercury concentration in swordfish must be dangerously high.
2) I met two students who took more than 10 years to graduate from UCLA, so it must take longer to graduate at UCLA than at many other colleges.
3) My friend's grandfather had a heart attack and died after they gave him sulphinpyrazone. The drug must not work.






## Anecdotal Evidence=Bad Science

- These are all anecdotal evidence, based on a limited sample size that might not be representative of the population.
- Anecdotal evidence is typically composed of unusual cases that we recall based on their striking characteristics.
- What we should do, is sample appropriately from the population of interest so we can do good science.



## One of many definitions…

- 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.

- In statistics, the questions we have are about a population.

# Questions about the Population



- We seek a particular truth and want to learn something about the population that we cannot easily observe.
- Perhaps the ...
  - Unemployment rate among recent college graduates
  - Effectiveness of a flu vaccine
  - Number of eggs contaminated with Salmonella in a restaurant

# Population



- The *population* is the entire collection of objects or individuals about which information is desired.
- Ex: If we want to determine the % of men on this campus, the population is ALL the people on this campus.
- Ex: If we want to determine the percentage of men in the US, the population is ALL people in the US.

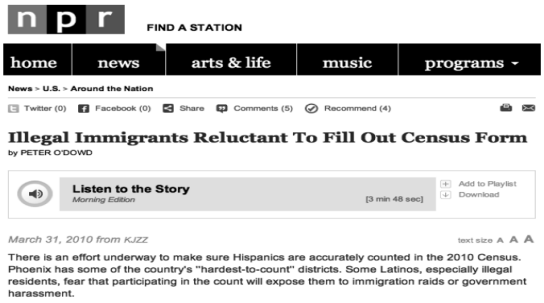
# Can we obtain the population?



- For the college example, the university probably has the exact list of students on this campus so yes, the population can be obtained.
- But, for the US, there is not a list of all the people. And what about tourists, etc.. Who exactly makes up your population?
- Examining an entire population is called taking a census, but it can be very difficult and/or expensive.

# Difficulties with Examining the Population



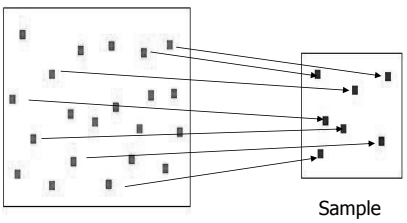



Obama's Lead Looks Stronger in Polls That Include Cellphones

# Sampling



If we can't obtain the population, we'll look to obtain a sample which is a subset of the population.




The 10% rule; in this class we will assume our sample is less than 10% of the population (and that the population is large).

# Sampling is a natural thing to do



- Think about sampling something you are cooking - you taste (examine) a small part of what you're cooking to get an idea about the dish as a whole.
- If you walk into a store that you're not familiar with, in order to decide if the store is reasonably priced, you wouldn't check the tag of every single item in the store. You would instead try to check out the price of a variety of items (a representative sample) and based on what you see decide if you think the store overall is overpriced or not.

# Again-Why Sample?



- Expense: cheaper than a census
  - Nielson ratings: based on 5000 out of an estimated 105.5 million US households with TVs
- Time: quicker than a census
  - Exit polls: gives news agencies valuable (?) information on election day in order to project election before all votes (census) are counted
- Sampled units must sometimes be destroyed (or changed) to measure characteristics
  - Reliability studies: testing lifetime of light bulbs, strength of windshields, etc.

# From sample to population



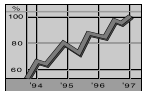
Statistics: *the science of collecting, presenting, analyzing and interpreting data to assist in making effective decisions.*

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

# Descriptive Statistics



- **Collect data**
  - e.g., Survey, Observation, Experiments
- **Present data**
  - e.g., Charts and graphs
- **Characterize (describe) data**
  - e.g., Sample mean




# Inferential Statistics

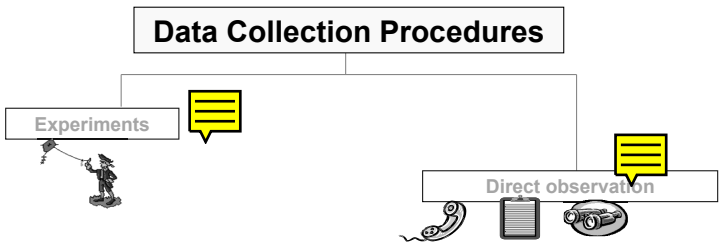


Drawing conclusions and/or making decisions concerning a population based on sample results.

- **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

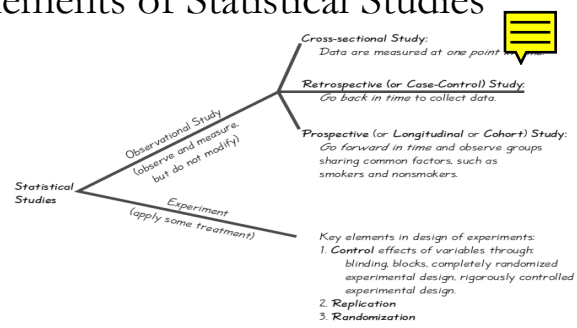
# Procedures for Collecting Data






# Elements of Statistical Studies



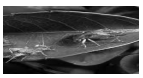


# Observational Study



- **Observational study:** Record data on individuals without attempting to influence the responses. We typically *cannot* prove cause & effect this way.
- A survey is a type of observational study.

- Example: Based on observations you make in nature, you suspect that female crickets choose their mates on the basis of their health.
  → Observe health of male crickets that mated.



# Different Types of Observational Studies



- In a cross-sectional study, data are observed, measured and collected at one point in time.
- In a retrospective study, data are collected from the past by going back in time (databases, old interviews, etc).
- In a prospective (longitudinal) study, data are collected in the future from groups sharing common factors.

# Example: Cellular Phones and Brain Tumors



- Researchers Joachim Schüz and associates wanted "to investigate cancer risk among Danish cellular phone users who were followed for up to 21 years."
- To do so, they kept track of 420,095 people whose first cellular telephone subscription was between 1982 and 1995. In 2002, they recorded the number of people out of the 420,095 people who had a brain tumor and compared the rate of brain tumors in this group to the rate of brain tumors in the general population.

# Cell Phone Study (cont)



- They found no significant difference in the rate of brain tumors between the two groups.
- The researchers concluded "cellular telephone was not associated with increased risk for brain tumors." (*Source*: Joachim Schüz et al. "Cellular Telephone Use and Cancer Risk: Update of a Nationwide Danish Cohort," *Journal of the National Cancer Institute* 98(23): 1707-1713, 2006)

# A Caveat On Observational Studies



- In an observational study, there can always be lurking (confounding) variables affecting the results.
- This means that observational studies can never show causation.
- It is easier to adjust for lurking variables in an experiment.

# Lurking Variables



A **lurking (confounding) variable** is a variable not included in the study design that does have an effect on the variables studied.

Lurking variables can *falsely suggest* a relationship.

**Relationship between boys foot size and reading ability**

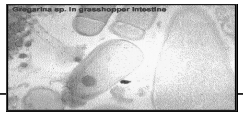
**Strong positive association between number of firefighters at a fire site and the amount of damage a fire does.**

# Experimental Study



- **Experimental study:** Deliberately impose a treatment on individuals and record their responses. Lurking variables can be controlled.
- In principle, experiments can give good evidence for causation through what we call randomized controlled comparative experiments.

- Example: Deliberately infect some males with intestinal parasites and see whether females tend to choose healthy rather than ill males.




# Basic Randomized Controlled Experiment



- Individuals are randomly assigned to groups, then the groups are randomly assigned to treatments.


# Blind or Double Blind Studies



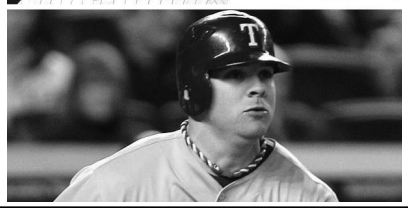
- Blind studies means the participants are not aware which group they are in
- Double Blind means participants and researchers don't know.

# Example



"Balance Necklaces"

The company explains that the titanium in its product works like small magnets to realign the body's biologic electrical fields



# The Air Force was interested



Some Major League Baseball players wear the company's nylon-coated necklaces, although it is not known whether it is for the perceived physiological benefits, because of superstition, or a combination of both.[3] However, there is no medical or scientific evidence supporting any physiological effect of Phiten's products,[4][5] and for many they are seen as mere superstition and ritual.[6]

In 2016 instructors at the United States Air Force Academy conducted a double-blind controlled test of the Phiten titanium necklace and found that there was no significant difference between a piece of clothesline and the Phiten necklace.

# Tax Dollars at work



Instructors at the United States Air Force Academy conducted a double-blind study of the Phiten claim. Initially the instructors became interested in this product because of the companies claims that it could cause relaxation, maximize strength and energy. When looking into the product before testing they flagged several concerns; Most studies supporting Phiten were also funded by Phiten, The claims sound like the definition of pseudoscience with no mechanism to show how it could work and the claims are unfalsifiable. The packaging claims that '"not all users will experience the intended benefits."' Leaving the user to believe that it will work or it will not work. Also the instructors felt that the Phiten product was similar to other pseudoscientific medical accessories like copper bracelets, crystals and magnetic bands.[8]

Forty-eight Academy cadets wore a tape-covered necklace for 48-hours, half were necklaces purchased from the Phiten website and the other half were clothesline purchased from The Home Depot. All necklaces were the same length and wrapped in masking tape. The cadets selected a necklace from a bin and they and the instructors did not know who wore a clothesline and who wore a Phiten necklace. At the end of the 48-hours each cadet was asked to fill out a survey rating from one to nine, if they felt relaxed, angry or energized? The conclusion was that there was no "statistically significant difference" between the clothesline and the Phiten necklace.[8]

## Example: Cellular Phones and Brain Tumors



- Researchers Joseph L. Roti and associates examined "whether chronic exposure to radio frequency (RF) radiation at two common cell phone signals–835.62 megahertz, a frequency used by analogue cell phones, and 847.74 megahertz, a frequency used by digital cell phones–caused brain tumors in rats.
- The rats in group 1 were exposed to the analogue cell phone frequency; the rats in group 2 were exposed to the digital frequency; the rats in group 3 served as controls and received no radiation. The exposure was done for 4 hours a day, 5 days a week for 2 years. The rats in all three groups were treated the same, except for the RF exposure.

## Study 2 Continued



- After 505 days of exposure, the researchers reported the following after analyzing the data. "We found no statistically significant increases in any tumor type, including brain, liver, lung or kidney, compared to the control group."
- (*Source*: M. La Regina, E. Moros, W. Pickard, W. Straube, J. L. Roti Roti. "The Effect of Chronic Exposure to 835.62 MHz FMCW or 847.7 MHz CDMA on the incidence of Spontaneous Tumors in Rats." Bioelectromagnetic Society Conference, June 25, 2002.)

## Disadvantages of Experiments



- They can be unethical to perform on the subjects in which you are interested.
- It can be difficult to monitor subjects to ensure that they are doing what they are told.
- They can take many years, even decades, to complete.
- Results of experiments that use animals do not generalize to humans.
- They are unnecessary when the question of interest does not involve trying to assess causality.

## We Will Focus on Observational Studies

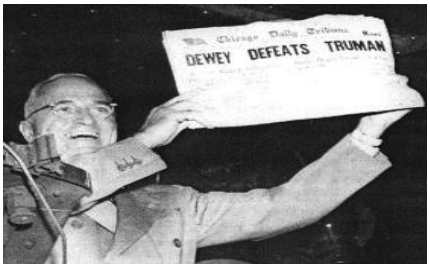


- Most of the data we will work with in this course will be from observational studies such as surveys.


## Important Warning



Surveys are not that easy to do correctly.


## What's the Matter With Polling?



The New York Times

By CLIFF ZUKIN JUNE 20, 2015


Jennifer Daniel

OVER the past two years, election polling has had some spectacular disasters. Several organizations tracking the 2014 midterm elections did not catch the Republican wave that led to strong majorities in both houses; polls in Israel badly underestimated Prime Minister Benjamin Netanyahu's strength, and pollsters in Britain predicted a close election only to see the Conservatives win easily. What's going on here? How much can we trust the polls as we head toward the 2016 elections?

## Important Warning



- Statistical studies are driven by data.
- How we collect data is at least as important as how we analyze it.
- In particular, a sample should be representative of the population, and *random sampling* (everyone in the population is equally likely to be selected), is often the best way to achieve this.

## Simple Random Sampling (SRS)



- It is the easiest most widespread form of sampling.
- Each subject has an equal chance to be in the sample.
- The sample enables us to determine how likely it is that descriptive statistics (like the sample mean) fall close to corresponding values for which we would like to make inference (like the population mean).
- Fundamental rule:

Available data (sample) can be used to make inferences about a much larger group (population) if the data can be considered ***representative*** (of the population) for ***the question(s) of interest***.

## Fundamental Rule Examples



- Do Harvard students represent all college students?
  - It depends on the question. If we're measuring heights, then maybe yes? If we're estimating the number of times you've visited New England, then no.
- Do the 100 U.S. Senators in 2015 represent all U.S. Senators (across all years)?
  - No. The first female senator was 87-year-old Rebecca Felton (D-GA) in 1922.

## Sampling Bias



- Systematic errors that result in a sample that is not representative of the overall population of interest
- We must be cautious of potential sources of bias in our sampling results

## Sometimes its not just bias




## Selection Bias



- An example of selection bias is wanting to know how all 8th graders feel about the coming basketball game, but only asking the basketball players what they think.
- Interviewing students at the Hong Kong at 1am about how much they study to understand all Harvard students.

# Voluntary Response Bias



- Advice columnist Ann Landers asked her readers, "If you had it to do over again, would you have children?"
- A few weeks later, her column was headlined: "70% OF PARENTS SAY KIDS NOT WORTH IT."
- The people who responded felt strongly enough to take the trouble to write Ann Landers. Their letters showed that many of them were angry at their children.
- These people don't fairly represent all parents.
- A statistically designed opinion poll on the same issue a few months later found that 91% of parents would have children again

http://www.stats.uwo.ca/faculty/bellhouse/stat353annlanders.pdf

# Voluntary Response Bias



- Two words-Amazon Reviews
- Results are often biased because people with strong opinions (usually negative) are more likely to respond and be included in the sample
- They also can be whack:

Dell Alienware Area-51 Gaming Machine (Intel Core i7-5820K 6-cores, 32GB DDR4 Ram, 18 TB HDD, Windows 7 Professional)
by Dell
256 customer reviews | 51 answered questions
Price: $7,939.00 & FREE Shipping

# Wording-Deliberate Bias



- "If you found a wallet with $20 in it, would you return the money?"
- "If you found a wallet with $20 in it, would you do the right thing and return the money?"

# Don't Ask Confusing Questions



- In one of the most infamous examples of flawed polling, a 1992 poll conducted by the Roper organization for the American Jewish Committee found that 1 in 5 Americans doubted that the Holocaust occurred.
- How could 20 percent of Americans report being Holocaust deniers?
- The answer became clear when the original question was re-examined.

# The Question Asked



- "Does it seem possible or does it seem impossible to you that the Nazi extermination of the Jews never happened?"
- This awkwardly-phrased question contains a confusing double-negative which led many to report the opposite of what they believed. Embarrassed Roper officials apologized, and later polls, asking clear, unambiguous questions, found that only about 2 percent of Americans doubt the Holocaust.



The New York Times **U.S.**

WORLD | U.S. | N.Y. / REGION | BUSINESS | TECHNOLOGY | SCIENCE | HEALTH | SPORTS | OPINION
POLITICS | EDUCATION | TEXAS

## Poll on Doubt Of Holocaust Is Corrected

By MICHAEL R. KAGAY
Published: July 8, 1994

Only a small segment of Americans deny that the Holocaust happened, according to a new Roper poll for the American Jewish Committee.

The new poll was conducted to correct a flawed question in an earlier survey that had suggested higher levels of Holocaust denial and, both the researchers and sponsors say, to set the polling record straight.

"Committed or consistent deniers of the Holocaust make up only a small segment of the population, about 2 percent or less," said Dr. Tom W. Smith, of the University of Chicago's National Opinion Research Center, who analyzed both the new Roper poll and all other available United States surveys on the Holocaust for the American Jewish Committee.

FACEBOOK
TWITTER
GOOGLE+
EMAIL
SHARE
PRINT
REPRINTS

# Wording-Unintentional Bias



- "Over the past few days, how many servings of fruit have you eaten?"
  - ❑ How many days are you considering?
  - ❑ What constitutes a serving?

# The Hite Report




http://www.hite-research.com/main.htm

# The 1987 Hite Report



- 84% of women are "not satisfied emotionally with their relationships"
- 70% of all women "married five or more years are having sex outside of their marriages
- 95% of women "report forms of emotional and psychological harassment from men with whom they are in love relationships"
- 84% of women report forms of condescension from the men in their love relationships

# How Was the Survey Executed



- The conclusions on the previous slide were drawn from a self-selected sample
- Hite mailed 100,000 questionnaires, only 4.5% were returned
- Questionnaires were mailed to groups of women professionals, counseling centers, church societies, and senior citizens centers
- Hite did not have a well-defined target population and non-response bias.

# More Recent Example



What's the biggest challenge facing *Fortune 500* companies?

That's the question we put to CEOs of all 500, and gave them a list to choose from: skilled labor shortages, competition from China, competition from start-ups, cyber security, shareholder activism, management diversity, employee diversity, geopolitical risk, and increased regulation.

Last year, in the final year of President Obama's eight-year term, it was increased regulation that had them most concerned, with 69% of those responding saying it was either their single biggest challenge or one of thei top three or four challenges. But this year, with a new man in Washington, that concern has subsided, with only 40% ranking it so high.

The bigger concerns this year: The rapid pace of technological change (73%, up from 64% last year) and cyber security (61%, up from 59% last year.)

FORTUNE
CEO DAILY
SUCCEED BIG IN BUSINESS

SHARE THIS SHARE THIS

JUNE 8, 2017

You can see more of the survey results here. A word on methodology, since my former colleagues at the Pew Research Center will ask. We sent an email survey directly to each of the 500 CEOs, and asked that they personally answer the questions. Seventy-two of them responded—a response rate of 14%. Given the busy lives these folks lead, that's not too shabby. If you would like to try the survey yourself, you may do so here.

# The Literary Digest Poll



- *Literary Digest* began summarizing contemporary opinion in 1890.
- *Literary Digest* correctly predicted the election of Harding in 1920, Coolidge in 1924, Hoover in 1928, and Franklin Roosevelt in 1932.

# Bring on 1936



The Literary Digest

NEW YORK — NOVEMBER 21, 1936

*Topics of the day*

**LANDON, 1,293,669; ROOSEVELT, 972,897**

Final Returns In The Digest's Poll of Ten Million Voters



Final Report "Literary Digest" 1936 Presidential Poll

They poll ten million voters and predict Landon to beat Roosevelt



The Literary Digest

NEW YORK — NOVEMBER 14, 1936

*Topics of the day*

**WHAT WENT WRONG WITH THE POLLS?**

None of Straw Votes Got Exactly the Right Answer—Why?







"LITERARY DIGEST" LEANS TALK TO THE EDITOR!

But they were wrong

Their sample frame for the poll was constructed from telephone directories, country club memberships and automobile registrations. The *Digest* predicted that Landon would win, but in reality FDR won by a landslide.

# Enter George Gallop



- George Gallup started a weekly column, “America Speaks,” in 1935.
- His sample totaled *only* 50,000 people.
- He predicted FDR would win in 1936 with 54 percent of the popular vote and 315 electoral votes.
- Gallup became “America’s oracle.”

# Gallup’s 1948 Polling Debacle



- The Major Candidates
  - Incumbent President Harry S. Truman (Democrat)
  - N.Y. Gov. Thomas E. Dewey (Republican)
  - S.C. Gov. Strom Thurmond (States’ Rights Democrat/“Dixiecrat”)
  - Former Vice President Henry A. Wallace (Progressive)

# Data Can Be Time Dependent

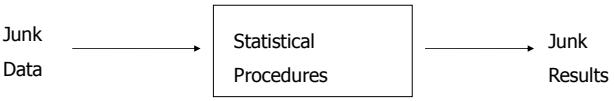


- Roughly two weeks before election day in November, Gallup confidently proclaimed Republican challenger Thomas Dewey to be the winner over the Democratic incumbent, President Harry S. Truman.
- Thinking Dewey to be far ahead and unbeatable, Gallup stopped interviewing voters. However, many voters changed their minds, switching from Dewey to Truman during that two week period. Gallup interviewers were not there to record those voting changes.
- It must be remembered that a poll only represents a "snapshot of opinion" at a particular point in time. The result was a Truman upset victory over Dewey.

# Remember : Junk in = Junk out



Ultimately, you must have good, representative data to draw the right conclusions. Otherwise, your results are worthless.



## Things you should know

- Population versus Sample
- Experimental and Observational Studies
- Sampling and Surveys Can Have Serious Issues