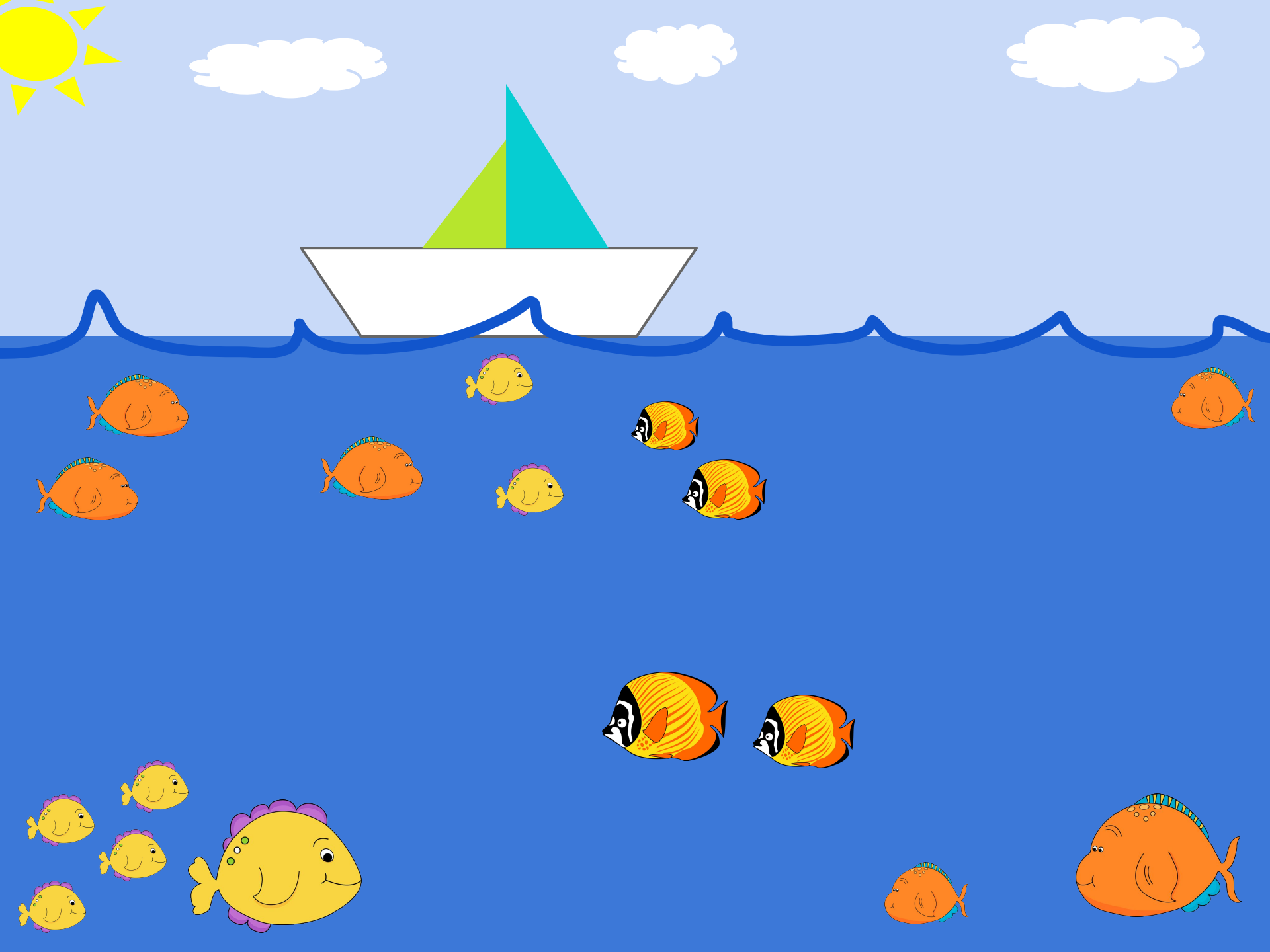
Sample Surveys

Gaston Sanchez

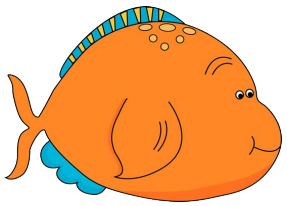
Sampling



Often we want to know something about the population

Population

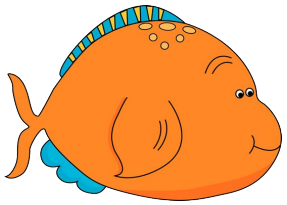
Gender of Fish



Proportion of female fish

Proportion of male fish

Length of Fish

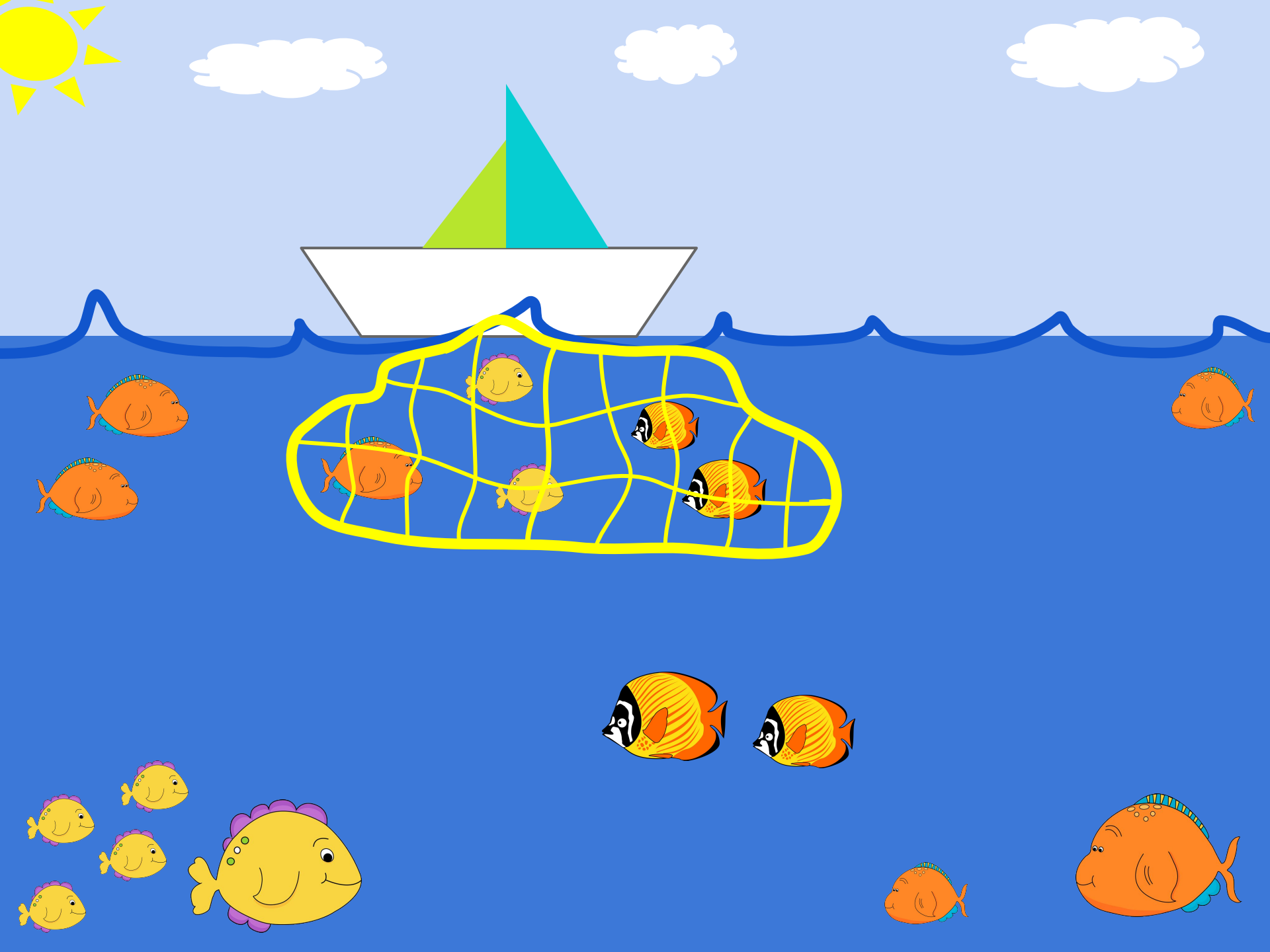


P(Avg length < 2 feet)

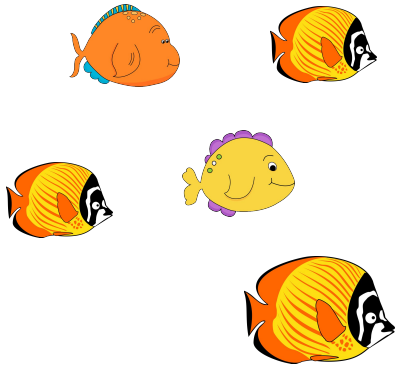
P(1ft < avg length < 2 ft)

Often we want to know something about the population

But we don’t have access to all of its individuals



Sample

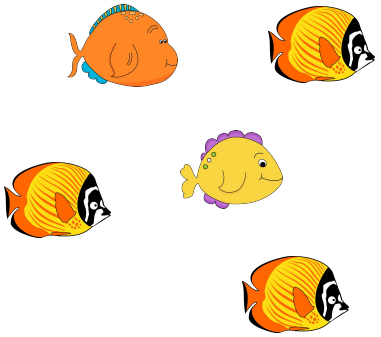


Population

Avg Weight

parameter

Sample



Avg Weight

statistic

To help you remember

Population

***P**arameter*

describes an aspect of the population

Sample

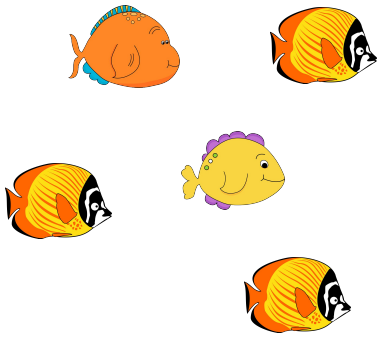
***S**tatistic*

describes an aspect of the sample

Estimation Idea

Samples statistics to estimate params

Sample



Avg Weight

statistic

Calculate statistics to estimate parameters

Population Avg Weight?

parameter

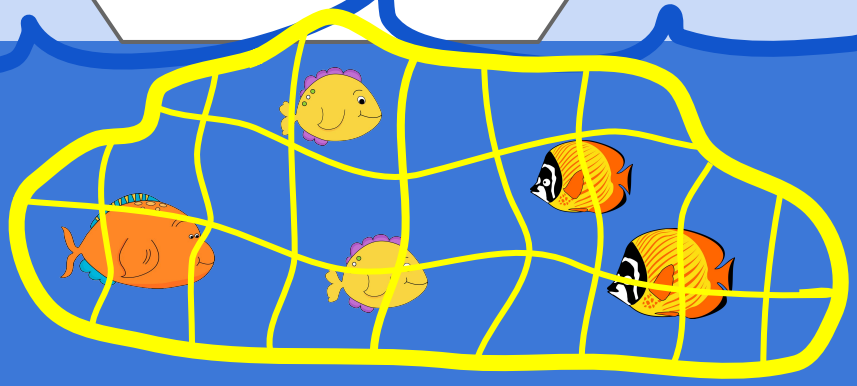

population difficult to observe

Typical limitations

- Hard to observe population
- Not enough time
- Not enough money
- Not enough resources

But...

We can get **samples** ...

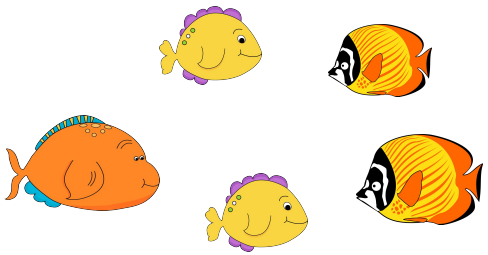


to make confident statements about populations

Instead we draw **samples** and expect to make generalizations about the population

Toward statistical inferences

What can these fish...



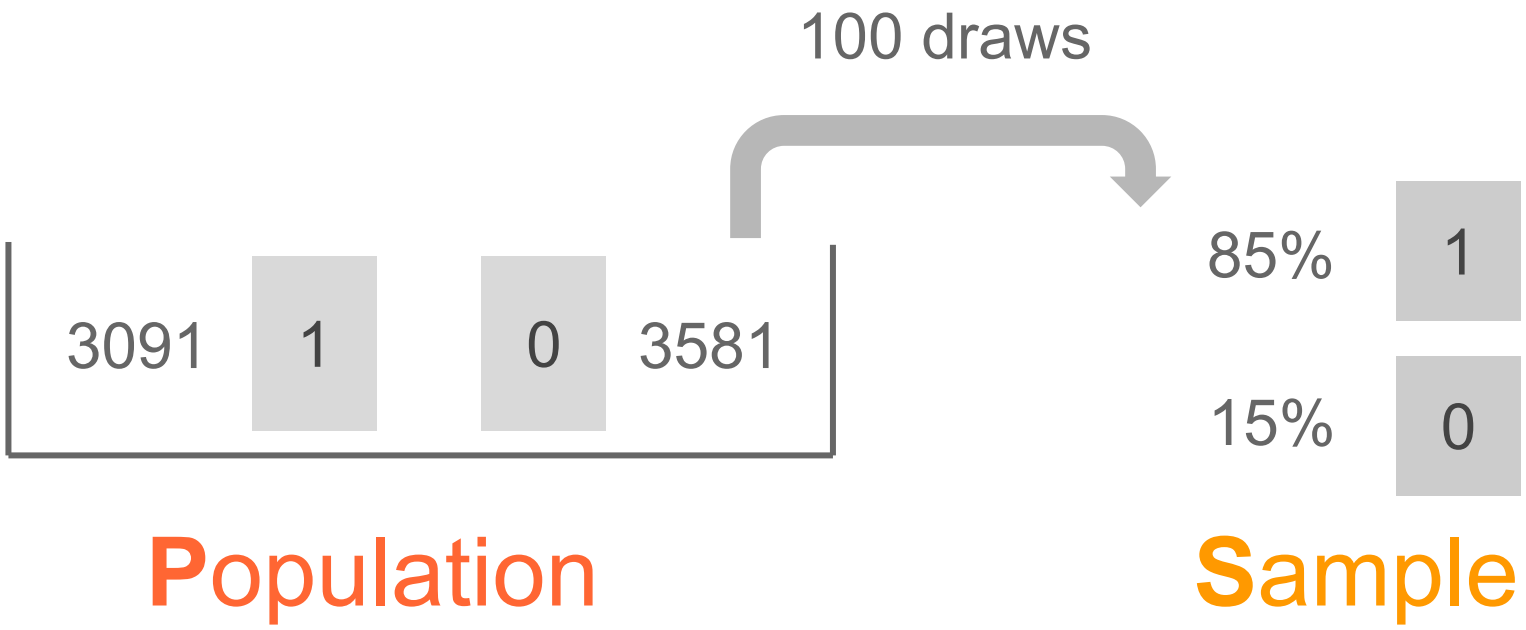
sample

tell us about these fish?

population

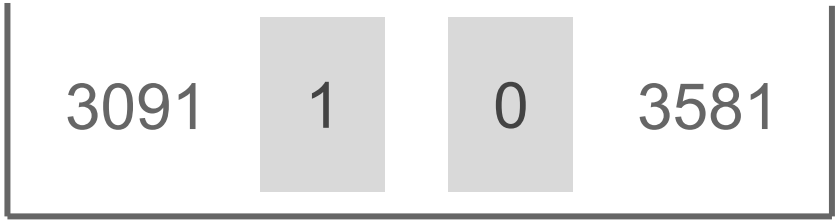
Estimation with box models

Population and Sample



Box Models so far

In our discussion so far of drawing from a box, we have known the contents of the box



100 draws

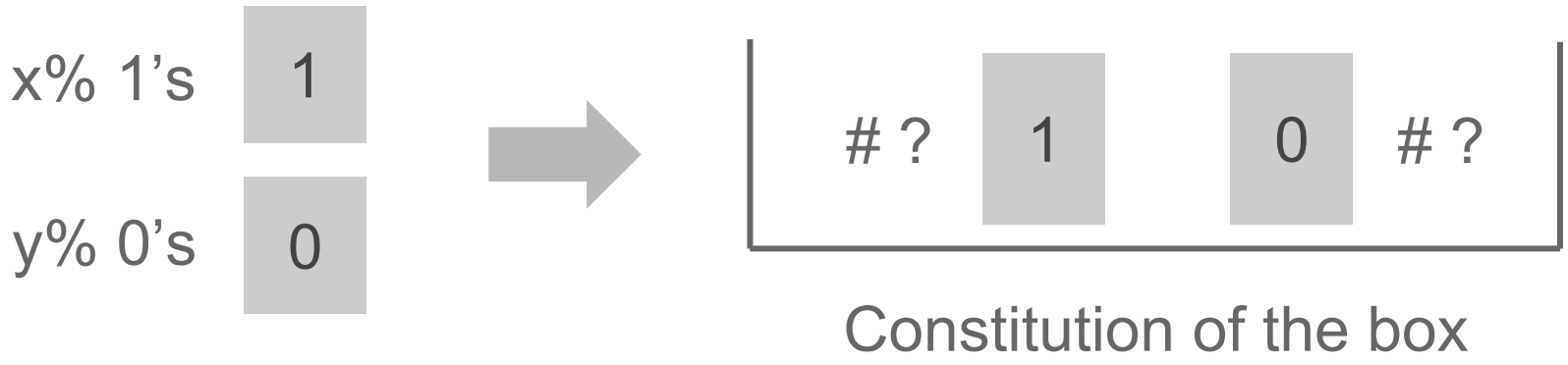
EV(sum of draws)

EV(avg of draws)

Turning to statistical estimation and inference

Now we work in the other direction: starting from a sample, we try to draw conclusions about the contents of the box (i.e. the population)

Sample of *b* draws



Chance Error and Bias

Chance processes

$$\text{Observed Value} = \text{Expected Value} + \text{Chance Error}$$

Sampling (in theory)

$$\text{Statistic} = \text{Parameter} + \text{Chance Error}$$

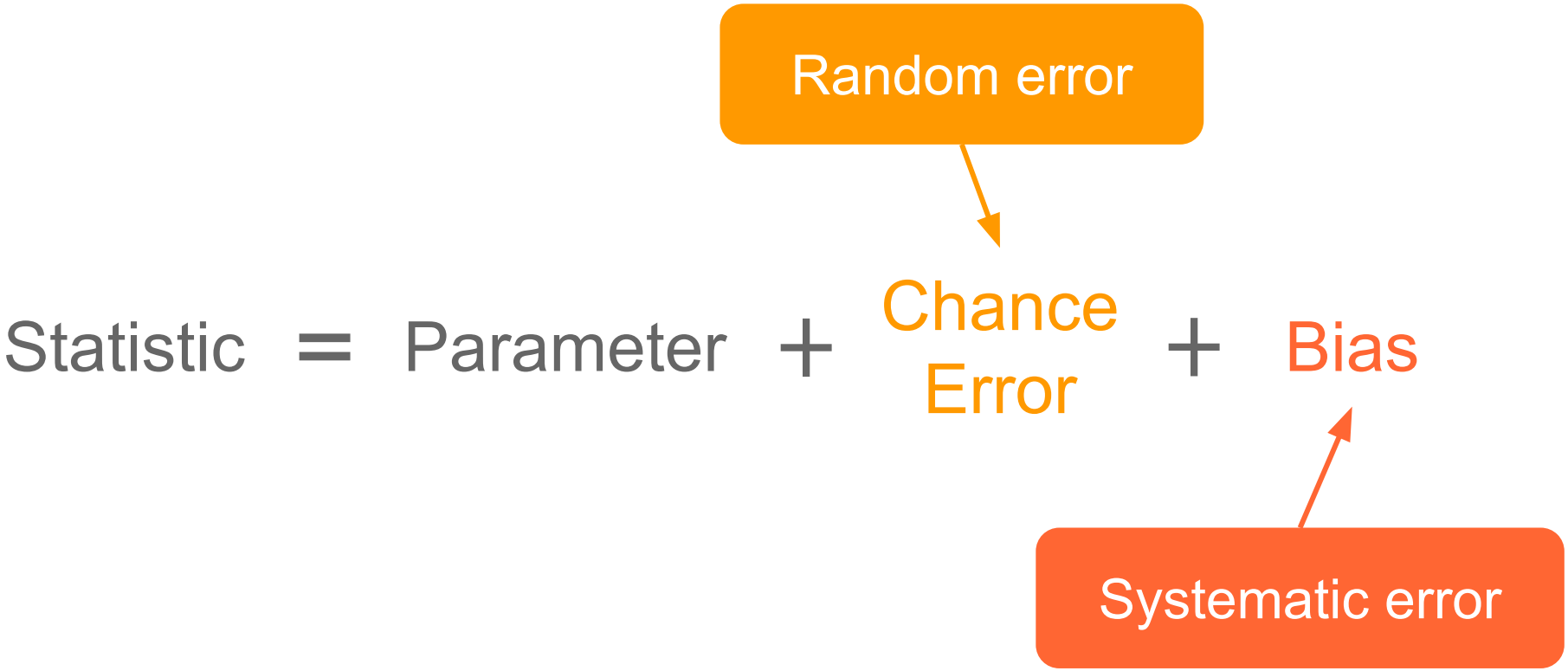
Chance processes

$$\text{Observed Value} = \text{Expected Value} + \text{Chance Error}$$

Sampling (in practice)

$$\text{Statistic} = \text{Parameter} + \text{Chance Error} + \text{Bias}$$

Errors in sampling



Why Sampling?

Some advantages of sampling

Sampling reduces the number of measurements that need be made:

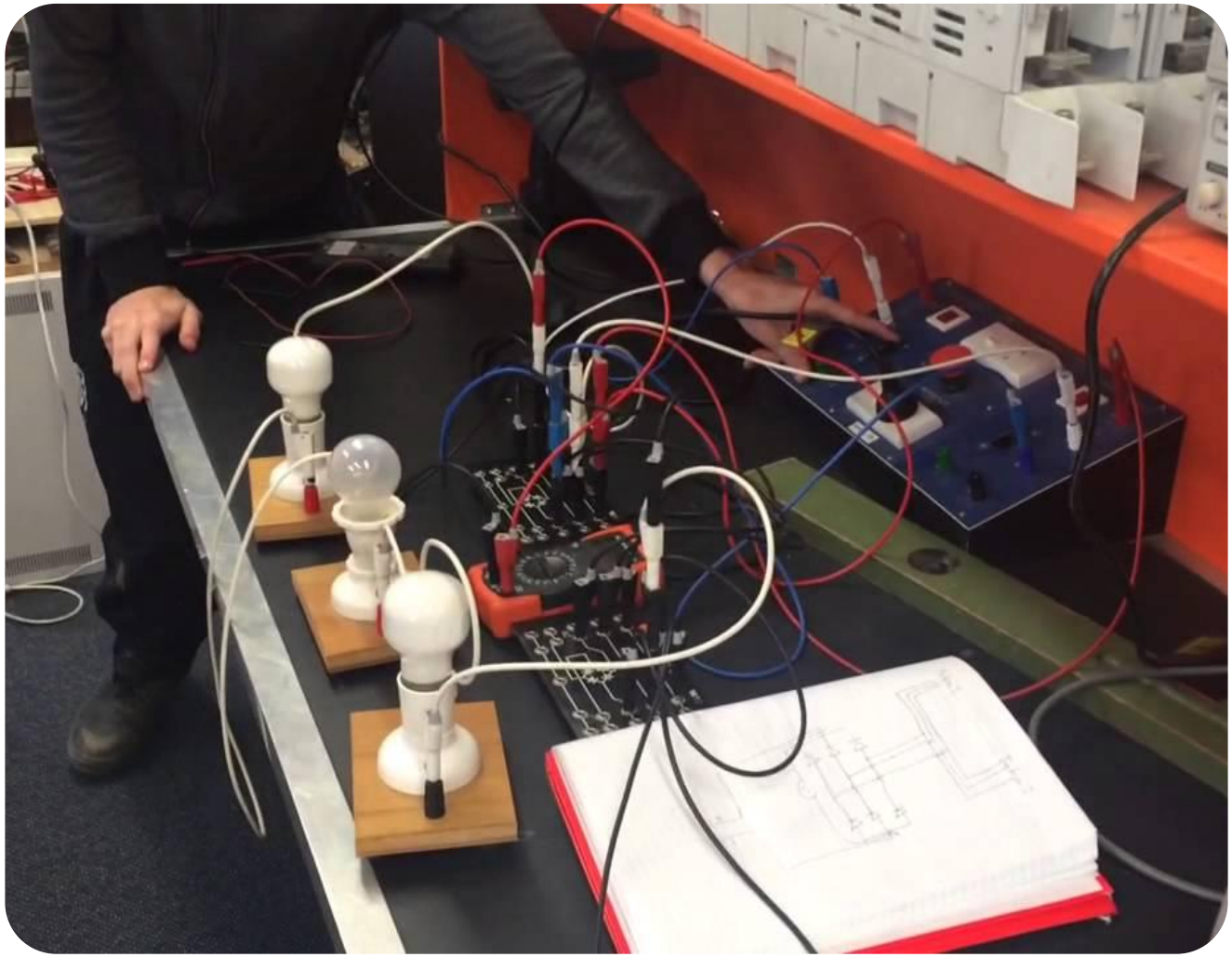
Allow more expensive measurement apparatus

Save money, time, and resources

Make timely estimates possible

Be essential in destructive testing

Lightbulb testing

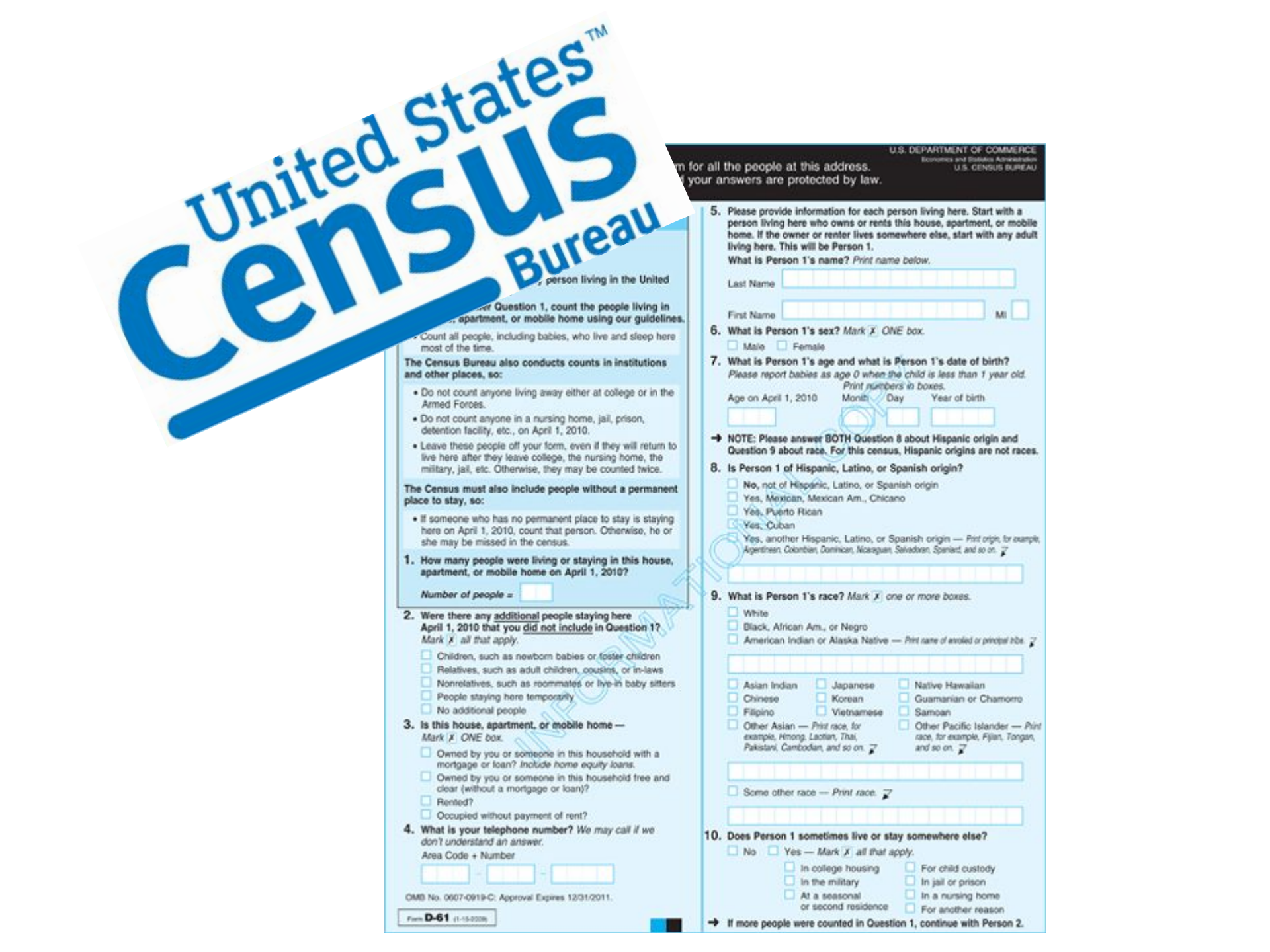


Car crash testing

Sample Design

Whether results from a sample can be extrapolated to the population from which the sample was drawn depends on the **DESIGN of the sample** (how the sample was drawn)

Sample Surveys


United States™
Census
Bureau
U.S. DEPARTMENT OF COMMERCE
Economics and Statistics Administration
U.S. CENSUS BUREAU
rn for all the people at this address.
your answers are protected by law.
person living in the United
er Question 1, count the people living in
apartment, or mobile home using our guidelines.
Count all people, including babies, who live and sleep here most of the time.
The Census Bureau also conducts counts in institutions and other places, so:
• Do not count anyone living away either at college or in the Armed Forces.
• Do not count anyone in a nursing home, jail, prison, detention facility, etc., on April 1, 2010.
• Leave these people off your form, even if they will return to live here after they leave college, the nursing home, the military, jail, etc. Otherwise, they may be counted twice.
The Census must also include people without a permanent place to stay, so:
• If someone who has no permanent place to stay is staying here on April 1, 2010, count that person. Otherwise, he or she may be missed in the census.
1. How many people were living or staying in this house, apartment, or mobile home on April 1, 2010?
Number of people =
2. Were there any additional people staying here April 1, 2010 that you did not include in Question 1? Mark ✗ all that apply.
Children, such as newborn babies or foster children
Relatives, such as adult children, cousins, or in-laws
Nonrelatives, such as roommates or live-in baby sitters
People staying here temporarily
No additional people
3. Is this house, apartment, or mobile home — Mark ✗ ONE box.
Owned by you or someone in this household with a mortgage or loan? Include home equity loans.
Owned by you or someone in this household free and clear (without a mortgage or loan)?
Rented?
Occupied without payment of rent?
4. What is your telephone number? We may call if we don't understand an answer.
Area Code + Number
OMB No. 0607-0919-C: Approval Expires 12/31/2011.
Form D-61 (1-15-2009)
5. Please provide information for each person living here. Start with a person living here who owns or rents this house, apartment, or mobile home. If the owner or renter lives somewhere else, start with any adult living here. This will be Person 1.
What is Person 1's name? Print name below.
Last Name
First Name
MI
6. What is Person 1's sex? Mark ✗ ONE box.
Male
Female
7. What is Person 1's age and what is Person 1's date of birth?
Please report babies as age 0 when the child is less than 1 year old.
Print numbers in boxes.
Age on April 1, 2010
Month
Day
Year of birth
➜ NOTE: Please answer BOTH Question 8 about Hispanic origin and Question 9 about race. For this census, Hispanic origins are not races.
8. Is Person 1 of Hispanic, Latino, or Spanish origin?
No, not of Hispanic, Latino, or Spanish origin
Yes, Mexican, Mexican Am., Chicano
Yes, Puerto Rican
Yes, Cuban
Yes, another Hispanic, Latino, or Spanish origin — Print origin, for example, Argentinean, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, and so on.
9. What is Person 1's race? Mark ✗ one or more boxes.
White
Black, African Am., or Negro
American Indian or Alaska Native — Print name of enrolled or principal tribe.
Asian Indian
Japanese
Native Hawaiian
Chinese
Korean
Guamanian or Chamorro
Filipino
Vietnamese
Samoan
Other Asian — Print race, for example, Hmong, Laotian, Thai, Pakistani, Cambodian, and so on.
Other Pacific Islander — Print race, for example, Fijian, Tongan, and so on.
Some other race — Print race.
10. Does Person 1 sometimes live or stay somewhere else?
No
Yes — Mark ✗ all that apply.
In college housing
In the military
At a seasonal or second residence
For child custody
In jail or prison
In a nursing home
For another reason
➜ If more people were counted in Question 1, continue with Person 2.


SURVEY
VERY SATISFIED
ATISFIED
9
8
7
6
5
4
3

Sample Surveys

Sample surveys attempt to determine the opinions, beliefs, behavior, or other parameters of a population of people from the responses of a sample (via a questionnaire or interview)

Ideally we want a sample that "matches" the population as close as possible

Typical Problem:

There is NO list having everybody in the population

Frame Bias:

What we have is a frame that tries to match the population as close as possible

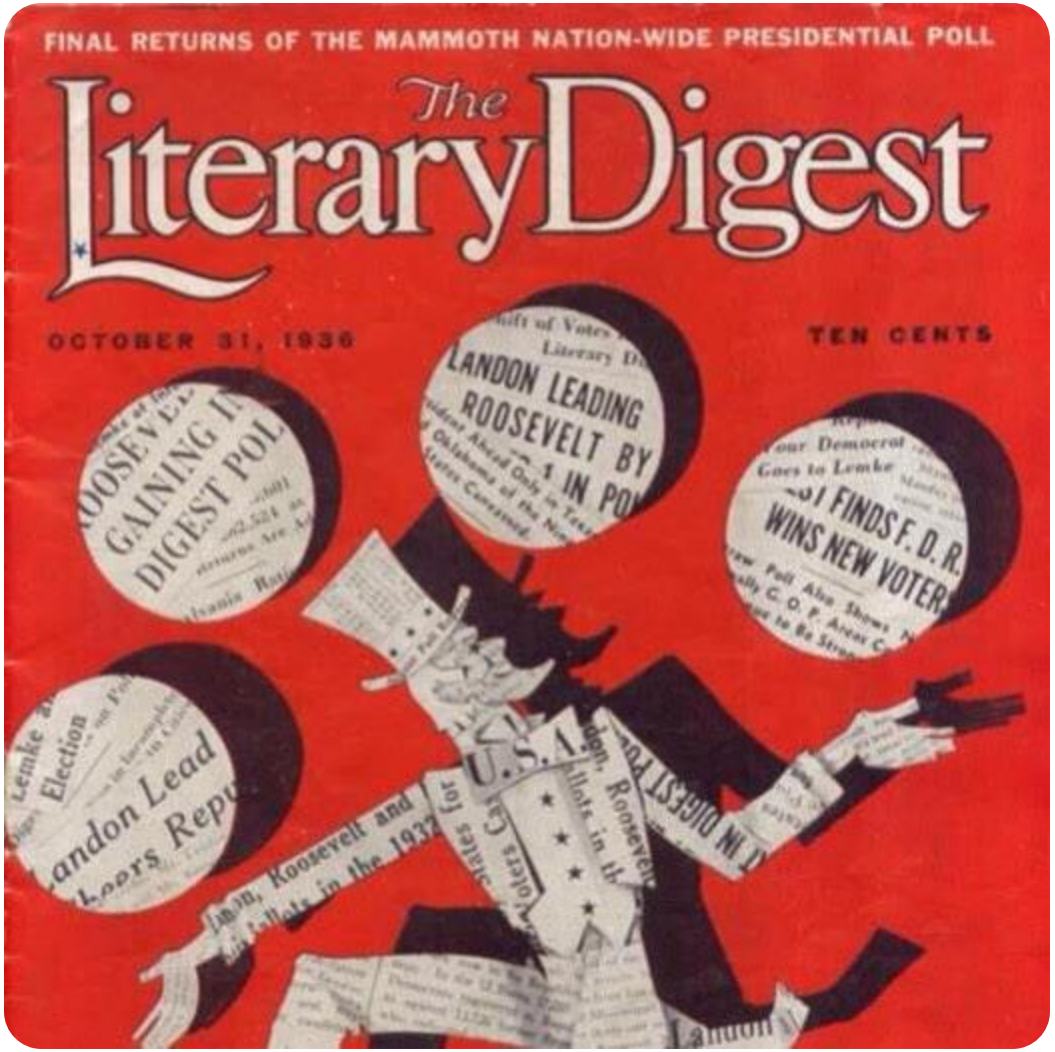
Literary Digest Example

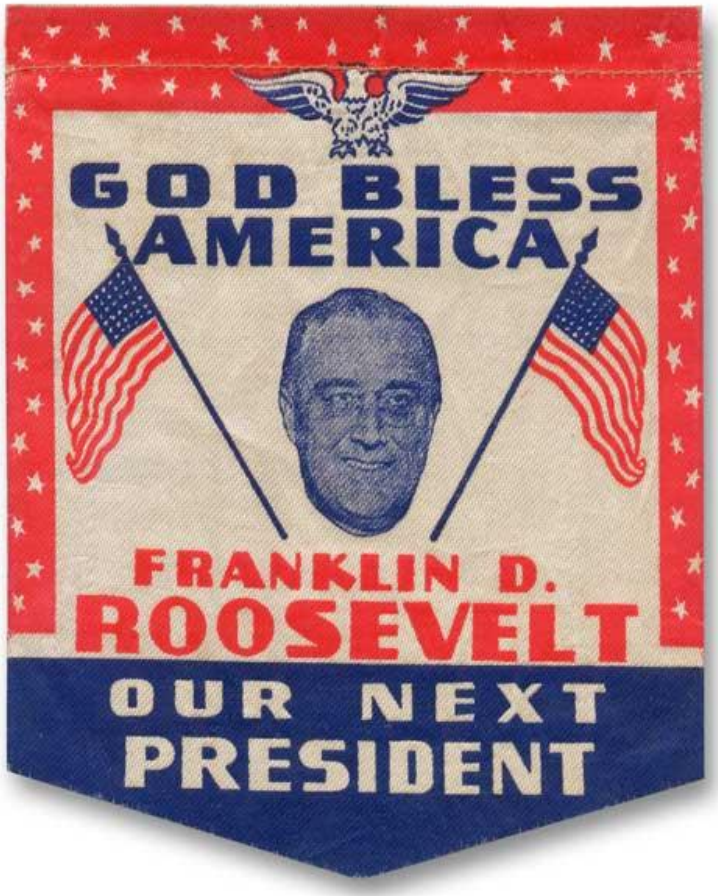
FINAL RETURNS OF THE MAMMOTH NATION-WIDE PRESIDENTIAL POLL

The Literary Digest

OCTOBER 31, 1936

TEN CENTS




GOD BLESS AMERICA
FRANKLIN D.
ROOSEVELT
OUR NEXT PRESIDENT


LANDON
FOR PRESIDENT

1936 Literary Digest Poll

1936 Presidential Elections

Literary Digest Poll

Took a list of 10 million people

Answers from about 2 million (20% response)

More than half responses favored Landon

What went wrong in the poll?

Not really a sample from the population

List made from phone directories and suscribers

Sample from that list (not representative of population)

Oversample of Republicans

Selection Bias

Questions about sensitive topics

How often do you use illegal drugs?

Do you pay your taxes?

How often do you cheat in tests?

Have you had extra-marital affairs?

People do not always answer with the truth

Difference in responses to given questions

Question wording bias

Should a woman have control over her own body, including her reproductive system?

Should a doctor be allowed to murder unborn children who can't defend themselves?

These questions frame things differently

Questions about sensitive topics

Sensitive questions

How often do you bathe?

Have you used illegal drugs?

Do you pay your taxes?

etc

People will not always answer with the truth