

SOC 3811/5811:
BASIC SOCIAL STATISTICS

Data and Data Analysis

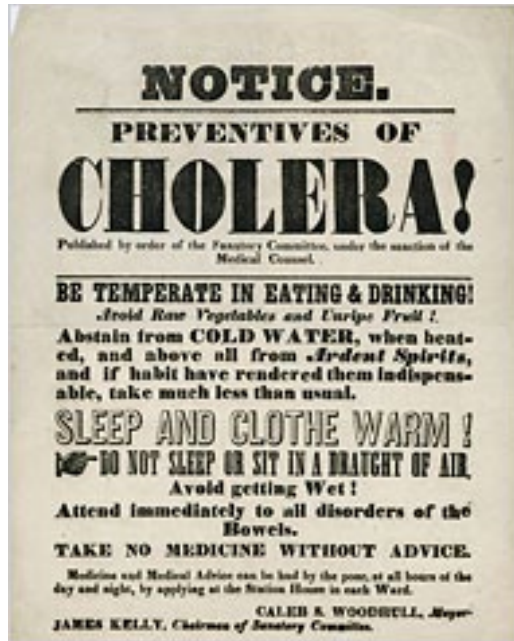
Reality → Data → Knowledge



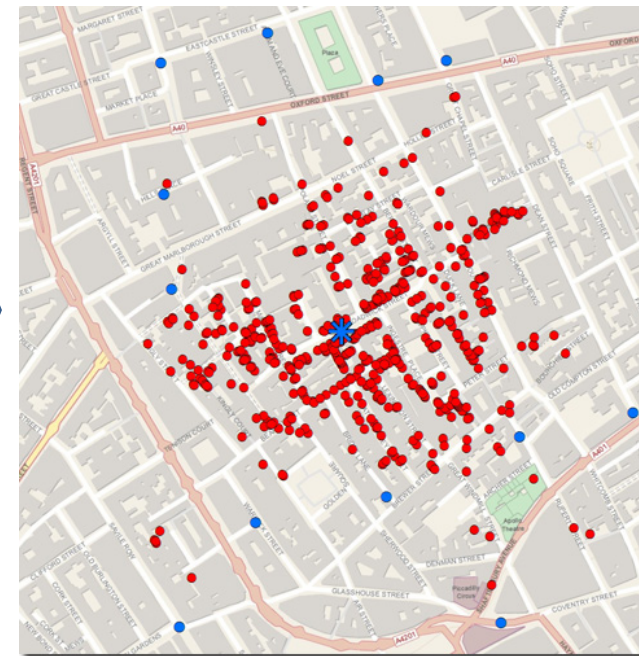
2954462187388221348
8711102682140050851
9291758426493659357
1797402968538728539
4401671369193970834
0165785879715135279
6365488407871392124
2920353540363160907
9546090581563252173
8998220343490929571
8159244365888845385
0229635205640107762
4385245611536981630
9196626803867353805
4956868092019177602
1078342587201850530
2250228297490991053
5575276026995705920
9224813120745604486
7096379166974664461
5350989184936109354



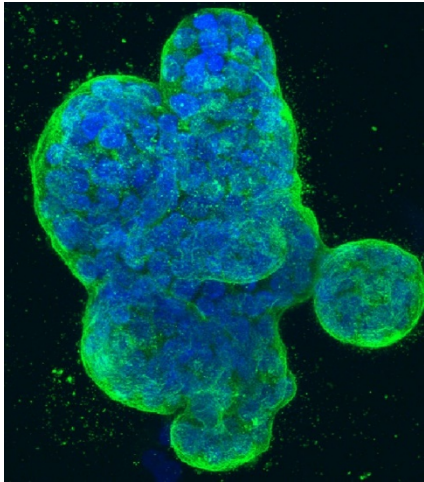
Reality → Data → Knowledge



2954462187388221348
8711102682140050851
9291758426493659357
1797402968538728539
4401671369193970834
0165785879715135279
6365488407871392124
2920353540363160907
9546090581563252173
8998220343490929571
8159244365888845385
0229635205640107762
4385245611536981630
9196626803867353805
4956868092019177602
1078342587201850530
2250228297490991053
5575276026995705920
9224813120745604486
7096379166974664461
5350989184936109354



Reality → Data → Knowledge



2954462187388221348
8711102682140050851
9291758426493659357
1797402968538728539
4401671369193970834
0165785879715135279
6365488407871392124
2920353540363160907
9546090581563252173
8998220343490929571
8159244365888845385
0229635205640107762
4385245611536981630
9196626803867353805
4956868092019177602
1078342587201850530
2250228297490991053
5575276026995705920
9224813120745604486
7096379166974664461
5350989184936109354



Reality → Data → Knowledge

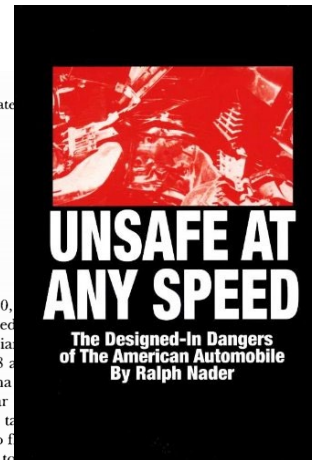


2954462187388221348
8711102682140050851
9291758426493659357
1797402968538728539
4401671369193970834
0165785879715135279
6365488407871392124
2920353540363160907
9546090581563252173
8998220343490929571
8159244365888845385
0229635205640107762
4385245611536981630
9196626803867353805
4956868092019177602
1078342587201850530
2250228297490991053
5575276026995705920
9224813120745604486
7096379166974664461
5350989184936109354



222 Corporate

Is New Again



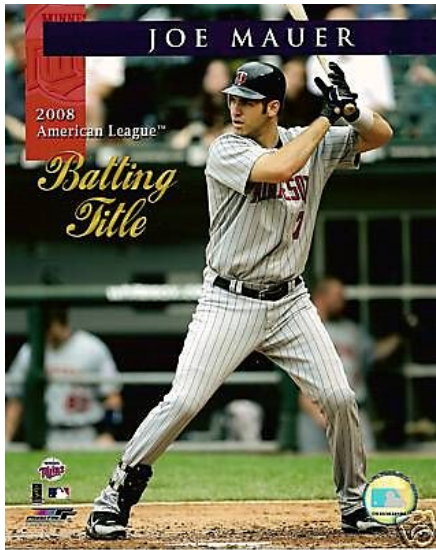
W. Michael Hoffman

On August 10, 1977, a fire occurred in Goshen, Indiana. A 17-year-old boy, Richard (ages 18 and 19), was driving a red Ford Pinto. The car burst into flames. The gas tank was located behind the rear seat. The car was burned to a crisp. Subsequently an Elkhart County grand jury returned a criminal homicide charge against Ford, the first ever against an American corporation. During the following 20-

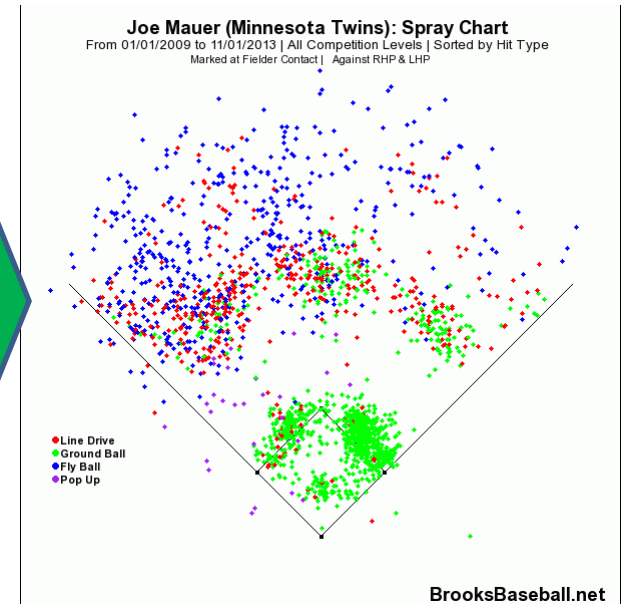
found to this trial some discussion of controversy is necessary. In 1977 the *Mother Jones* broke a story by the general manager of *Mother Jones* operations, accusing Ford of knowing on the road an unsafe car—the which hundreds of people have suffered burn deaths and even been scarred and disfigured due to burns. In his article "Pinto Madness" Dowie charges that:

more have been scarred and disfigured due to burns. In his article "Pinto Madness" Dowie charges that: Fighting strong competition from Volkswagen

Reality → Data → Knowledge



2954462187388221348
8711102682140050851
9291758426493659357
1797402968538728539
4401671369193970834
0165785879715135279
6365488407871392124
2920353540363160907
9546090581563252173
8998220343490929571
8159244365888845385
0229635205640107762
4385245611536981630
9196626803867353805
4956868092019177602
1078342587201850530
2250228297490991053
5575276026995705920
9224813120745604486
7096379166974664461
5350989184936109354



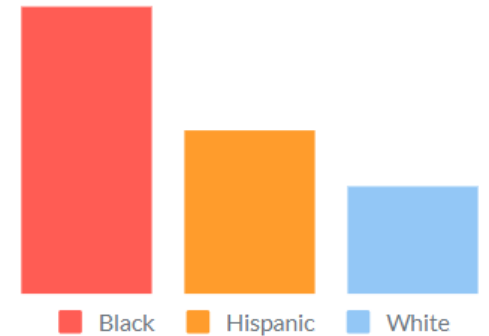
Reality → Data → Knowledge



2954462187388221348
8711102682140050851
9291758426493659357
1797402968538728539
4401671369193970834
0165785879715135279
6365488407871392124
2920353540363160907
9546090581563252173
8998220343490929571
8159244365888845385
0229635205640107762
4385245611536981630
9196626803867353805
4956868092019177602
1078342587201850530
2250228297490991053
5575276026995705920
9224813120745604486
7096379166974664461
5350989184936109354



3x more likely to be killed by police than white people.



Police Killings per 1 million population

Reality → Data → Knowledge



2954462187388221348
8711102682140050851
9291758426493659357
1797402968538728539
4401671369193970834
0165785879715135279
6365488407871392124
2920353540363160907
9546090581563252173
8998220343490929571
8159244365888845385
0229635205640107762
4385245611536981630
9196626803867353805
4956868092019177602
1078342587201850530
2250228297490991053
5575276026995705920
9224813120745604486
7096379166974664461
5350989184936109354



Food Insecurity in America

food insecurity (n) - access to adequate food is limited by a lack of money and other resources

in 2015:

12.7%
of US households were food insecure

which equals

15.8 million
households

13.1 million
children under 18 were food insecure

food insecurity in children leads to

- lower grades
- tardiness/absences
- obesity
- low mental health
- low physical health

“We do have food we don’t have the right food, so more food apartheid.”
- Karen Washington
founder, Black Urban Growers

sources:
ERS, USDA
<http://www.ers.usda.gov/publications/err-economic-research-report/err140/report-summary.aspx>
http://www.ers.usda.gov/media/2137657/err215_summary.pdf
http://www.ers.usda.gov/datafiles/Food_Access_Research_Aides/Download_the_Data/Archived_Version/archived_documentation.pdf
Be Kid Hungry
<https://www.nokidhungry.org/sites/default/files/child-economy-study.pdf#Project>
Yao, Slack and Bell
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2667853/>
Images:
The Noon Project [https://thnoonproject.com](https://th Noon Project https://thnoonproject.com)

Reality → Data → Knowledge



2954462187388221348
8711102682140050851
9291758426493659357
1797402968538728539
4401671369193970834
0165785879715135279
6365488407871392124
2920353540363160907
9546090581563252173
8998220343490929571
8159244365888845385
0229635205640107762
4385245611536981630
9196626803867353805
4956868092019177602
1078342587201850530
2250228297490991053
5575276026995705920
9224813120745604486
7096379166974664461
5350989184936109354



Reality → Data → Knowledge



Reality → Data → Knowledge

Research Methods (e.g., SOC 3801, SOC 8801)

Reality → Data → Knowledge

Vocabulary

Reality → Data → Knowledge

Unit of Observation

The type of object that we observe

(e.g., people, countries, classrooms, police encounters)

Population

All of the individuals about whom we wish to draw conclusions

Sample

A selected subset of individuals from the population

Sampling

The process we use to select a sample from the population

Variable

Any attribute that can differ (or vary) from individual to individual

Worksheet

Lecture Worksheet
Tuesday 9/8/2020

MAIN POINTS OF LECTURE

- I. Statistical reasoning is one important tool for understanding the world. In sociology, the goal is to understand the social world; statistics is one means to that end. This class is about the *logic* of statistical reasoning as much as about the *mechanics*
- II. Key Terms
 1. Unit of observation: The types of individuals we are studying (typically people in social sciences)
 2. Population: All of the individuals about whom we seek to draw conclusions
 3. Sample: The subset of the population that we actually observe in doing research
 4. Variable: An attribute or characteristic that may vary across individuals
 5. Measurement: The process of observing values of a variable for individuals in a sample
 6. Continuous variables categorize observations in terms of their numeric value of some attributes
 7. Discrete variables categorize observations in terms of qualitative, non-numeric attributes
 8. Nominal variables are discrete variables in which the categories cannot be put in order
 9. Ordinal variables are discrete variables in which the categories can be put in order
 10. Data File: A computer file containing values of variables for each individual in a sample
 11. Statistical Software: A computer program used for analyzing data files

QUESTIONS

1. We want to study how eligible voters in our congressional district will vote in the upcoming election. There are 125,000 eligible voters. We randomly select 1,000 of them and mail them a survey that asks them who they will vote for in the presidential election; 161 people return those surveys.
 - a. What is the unit of observation?
 - b. What is the population?
 - c. What is the sample?
 - d. What is the sampling procedure?
 - e. What key variable(s) did we measure?
2. We want to understand how the recent wildfires in Colorado have affected moose living in Rocky Mountain National Park. To find out, we draw blood from the first 10 moose we find inside the park, and we measure the blood's cortisol level as a way to assess how stressed the moose are.
 - a. What is the unit of observation?
 - b. What is the population?
 - c. What is the sample?
 - d. What is the sampling procedure?
 - e. What key variable(s) did we measure?
3. Are the following variables continuous or discrete? If discrete, are they nominal or ordinal?
 - a. Height
 - b. Hair color
 - c. Number of own children
 - d. Social class
 - e. Day of the week
 - f. Ever been arrested
4. What questions do you have for me about this course, about its content, or about how it will be run?
5. Please tell me three things about yourself ... things by which I can remember you.

Worksheet

We want to study how eligible voters in our congressional district will vote in the upcoming election. There are 125,000 eligible voters. We randomly select 1,000 of them and mail them a survey that asks them who they will vote for in the presidential election; 161 people return those surveys.

1. What is the unit of observation?
2. What is the population?
3. What is the sample?
4. What is the sampling procedure?
5. What key variable(s) did we measure?

Worksheet

We want to understand how the recent wildfires in Colorado have affected moose living in Rocky Mountain National Park. To find out, we draw blood from the first 10 moose we find inside the park, and we measure the blood's cortisol level as a way to assess how stressed the moose are.

1. What is the unit of observation?
2. What is the population?
3. What is the sample?
4. What is the sampling procedure?
5. What key variable(s) did we measure?

Reality → Data → Knowledge

Measurement

The process of ascertaining the value of a variable for an individual

Validity

The degree to which a measure of a concept is accurate

Reliability

The degree to which a measure of a concept is consistent

Reality → Data → Knowledge

How do we turn a measure into data?

Concept: Wealth

Measure: Ask people for the dollar value of all their savings, investments, and other assets (Reliable? Valid?)

Possible Values: \$0, \$1, \$2, \$3, ... to ∞

Data: 0 = \$0

1 = \$1

2 = \$2

etc



Values are in a logical order and the assigned values have a real numeric meaning

2954462187388221348
8711102682140050851
9291758426493659357
1797402968538728539
4401671369193970834
0165785879715135279
6365488407871392124
7353540363160907
9546090581563252173
1671369193970834
0165785879715135279
6365488407871392124
2920353540363160907
9546090581563252173
8998220343490929571
8159244365888845385
0229635205640107762
4385245611536981630
9196626803867353805
4956868092019177602
1078342587201850530
2250228297490991053
5575276026995705920
9224813120745604486

Reality → Data → Knowledge

How do we turn a measure into data?

Concept: Healthiness

Measure: Ask people whether they are in excellent, good, fair, or poor health (Reliable? Valid?)

Possible Values: excellent, good, fair, poor

Data: 1 = Excellent

2 = Good

3 = Fair

4 = Poor



Values are in a logical order but the assigned values have no numeric meaning

2954462187388221348
8711102682140050851
9291758426493659357
1797402968538728539
4401671369193970834
0165785879715135279
6365488407871392124
9546090581563252173
4401671369193970834
0165785879715135279
6365488407871392124
2920353540363160907
9546090581563252173
8998220343490929571
8159244365888845385
0229635205640107762
4385245611536981630
9196626803867353805
4956868092019177602
1078342587201850530
2250228297490991053
5575276026995705920
9224813120745604486

Reality → Data → Knowledge

How do we turn a measure into data?

Concept: Sexual orientation

Measure: Ask people whether they are gay, straight, bisexual, or something else (Reliable? Valid?)

Possible Values: gay, straight, bisexual, something else

Data: 1 = Gay

2 = Straight

3 = Bisexual

4 = Something else

← Values are not in a logical order and the assigned values have no numeric meaning

2954462187388221348
8711102682140050851
9291758426493659357
1797402968538728539
4401671369193970834
0165785879715135279
6365488407871392124
9546090581563252173
4401671369193970834
0165785879715135279
6365488407871392124
2920353540363160907
9546090581563252173
8998220343490929571
8159244365888845385
0229635205640107762
4385245611536981630
9196626803867353805
4956868092019177602
1078342587201850530
2250228297490991053
5575276026995705920
9224813120745604486

Reality → Data → Knowledge

Continuous Variables

Variables that categorize observations in terms of their numeric value of some attributes

In theory, continuous variables can take on all possible numerical values in a given interval

(e.g., IQ, wealth in dollars, height in inches)

Reality → Data → Knowledge

Discrete Variables

Variables that categorize observations in terms of qualitative, non-numeric attributes

Nominal Variables

Discrete variables in which the categories cannot be put in order
(e.g., favorite color, state of birth)

Ordinal Variables

Discrete variables in which the categories can be put in order
(e.g., highest degree, level of agreement with a statement)

Reality → Data → Knowledge

Variable Has An *Infinite* Number of Possible Values

→ It's a Continuous Variable

Variable Has a *Finite* Number of Possible Values...

→ It's a Discrete Variable

...and the categories *can* be put in order.

→ The Discrete Variable is Ordinal

...and the categories *cannot* be put in order.

→ The Discrete Variable is Nominal

There Are Gray Areas...

Worksheet

Are the following variables continuous or discrete? If discrete, are they nominal or ordinal?

1. Height
2. Hair color
3. Number of own children
4. Social class
5. Day of the week
6. Ever been arrested

Reality → Data → Knowledge

How do we use data to...

...describe a sample?

...make inferences about populations?

...test existing theory or evaluate existing policy?

...inform the development of new theory or policy?

These questions motivate Sociology 3811 (and 5811 and 8811 and parallel courses elsewhere)