

Audience

40%: School teachers: Current or previous
30%: College faculty: Current or previous
20%: Education, non-profit
10%: Industry, commercial

Statistical Literacy 2017

Milo Schield, Augsburg College

Elected Member: International Statistical Institute US Rep: International Statistical Literacy Project VP. National Numeracy Network

CME Presentation in Toronto
Fields Institute
April 29, 2016

www.StatLit.org/pdf/2017-Schield-CME-Slides.pdf

Statistical Literacy 2017: Overview

"We teach the wrong stuff; We teach it the wrong way; We teach it in the wrong order." Richard de Veaux

Statistical Literacy 2017:

- 1. What is it in general?
- 2. Who needs it?
- 3. What is it in particular?
- 4. Who can implement it?

la ^{VIA} 281 CASE

What are Statistics?

- a. Data; numerical data, classifications of data, or numerical summaries of data [Ambiguous]
- b. Outcomes from a random process;randomly-selected or randomly-assigned groups[Technical distinction]
- c. Numbers in context where the **context matters**: Quantitative summaries of real things: things that have natures, connections & causes

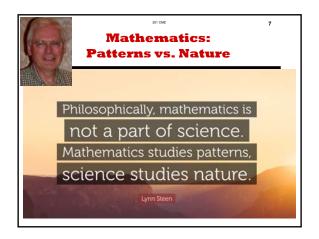
la ^{VIA}
Statistics is Different
from Mathematics

Math ignores the context.

- a. Math deals with form (ignores the matter)
- b. Math deals with variables and values (no natures)
- c. Math deals with associations and co-variates
- d. Math has no operator for "causes"

Statistics depends on the context

- a. Statistics deals with the matter: its nature
- b. Statistics deals with subjects and characteristics
- c. Statistics deals with "confounders"
- d. Statistics deals with "causes"



"Statistics Come From Data" s is like saying "Babies Come from Hospitals" T

HOSPITAL L

It's true but it leaves out the interesting details

What is statistical literacy?
In general terms

Statistical literacy is needed by citizens and social decision makers to enable them to understand and evaluate the statistics they encounter everyday.

Everyday statistics are used as evidence in arguments. Legal:

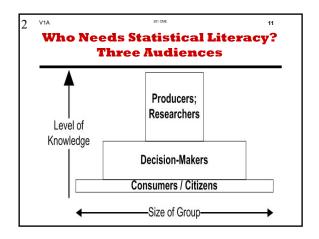
- Describe: 90% of a restaurant's staff speaks Spanish
- · Compare: Most Mexican restaurant staff speak Spanish
- · Evaluate: Mexican restaurants discriminate in hiring .

What is statistical literacy?
Examples:

Statistics are answers to questions or interests.

Medical:

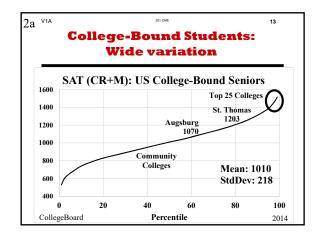
- · Describe: Japanese, who live long, eat low-fat diet
- · Compare: People with high-fat diets die sooner
- Evaluate: High-fat diet causes shorter lifespan. Social:
- Describe: Average school class size is 24
- Compare: Best performing classes are smaller
- · Evaluate: Smaller classes will improve outcomes

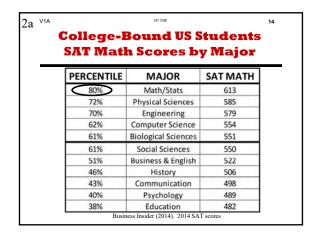


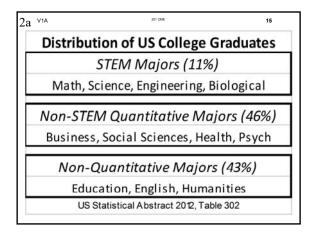
Three Audiences:
More detail

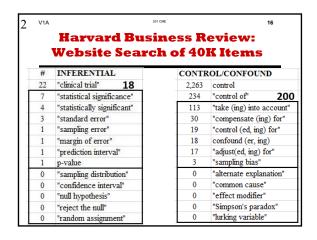
- 1. STEM majors and those who conduct surveys, studies and research.
- 2. Social decision-makers:
 Politicians, bureaucrats, business leaders, doctors
 - Those who inform citizens and decision makers:
 journalists, analysts, lawyers, economists, consultants,
 sociologists, political scientists, policy advocates,
 psychologists and educators.
- 3. Citizen in a modern republic or democracy.

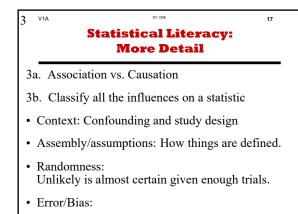
Statistical Literacy: 2017

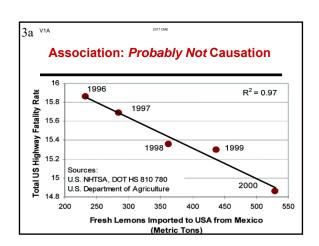




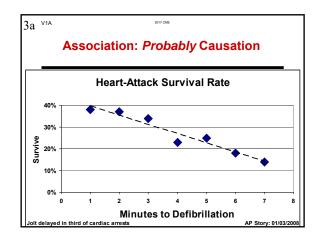


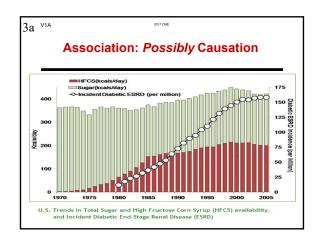






Statistical Literacy: 2017





Distinguish Causation
from Association

Causation (8%): cause, effects, results, prevents
Association (2%): associate, relate, correlate,
Between (67%):
 Action verbs: ups, cuts, raises, boosts, increases
 Other: due to, because of, attributed to
Inappropriate use of "causes":

Obesity causes later onset of puberty in boys

Junk food causes a third of heart attacks.

Schield and Raymond (2009) study 2,000 newspaper headlines involving quantity



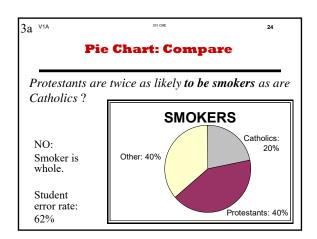
Association-Causation

Baseball players whose names begin with the letter "D" are more likely to die young

Drinking a full pot of coffee every morning will add years to your life, but one cup a day increases the risk of pancreatic cancer.

Asian-Americans are most susceptible to heart attacks on the fourth day of the month

Source: Standard Deviations: Flawed Assumptions, Tortured Data, and Other Ways to Lie with Statistics by Gary Smith (2015).

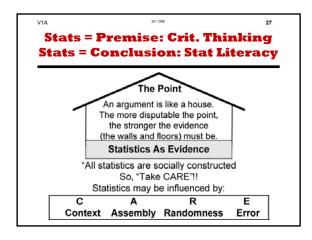


V1a 4/29/2017



Association vs. Causation 11 Headlines, Same Story

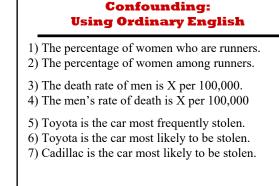
- 1. Study: 45,000 Uninsured *Die* a Year (CBS News)
- 2. 45,000 deaths attributable to uninsurance
- 3. 45,000 US deaths associated with lack of insurance
- 4. No health coverage tied to 45,000 deaths a year
- 5. Lack of insurance linked to 45,000 deaths
- 6. Study: 45,000 U.S. Deaths From Lack of Insurance
- 7. One death every 12 minutes due to no health insurance
- 8. 45,000 ... die because of lack of health insurance
- 9. Lack of Health Insurance Kills 45,000 a Year
- 10. Lack of Health Insurance cause 44,789 deaths
- 11. Lack of insurance to blame for almost 45,000 deaths



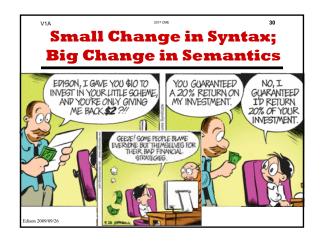
3b VIA 2010E 2 Statistical Literacy in detail: "Take CARE"

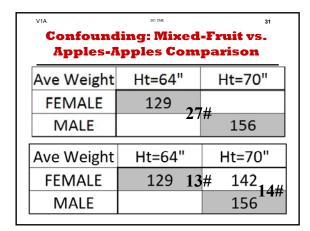
Statistical literacy studies all influences on statistic:

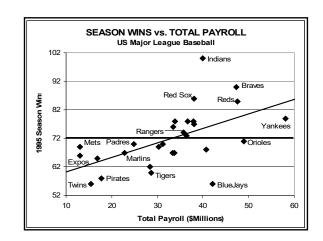
- · Confounding:
 - what was and was not controlled for
 - what kind of study was involved
- Assembly/Assumptions:
 - · how statistics are collected, defined and grouped
 - how statistics are summarized, compared & presented
- Randomness: small samples and big data
- Error/bias



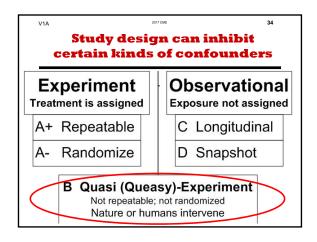
3b V1A

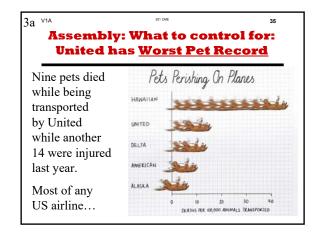


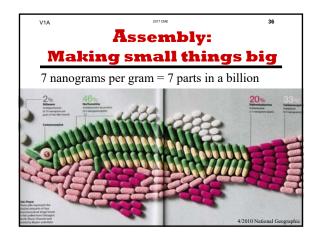




VIA	T-17E	2017 CME	L SCO	DFC	33
				RES	
Average SAT-V	1981	2002	Change	1981	2002
All Test-Takers	504	504	0	100%	100%
White	519	527	8	85%	65%
Black	412	431	19	9%	11%
Asian	474	501	27	3%	10%
Mexican	438	446	8	2%	4%
Puerto Rican	437	455	18	1%	3%
American Indian	471	479	8	0%	1%

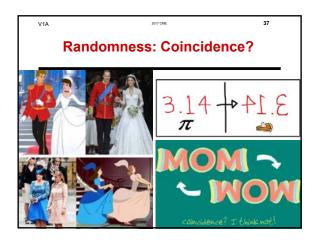






V1a 4/29/2017

Statistical Literacy: 2017



Error/Bias

A recent survey shows that most Republicans surveyed prefer Obama as President.

Question: Who would you prefer as President?

- · Barack Obama
- The captain of the Italian linear that crashed
- Charlie Sheehan
- · Lady Gaga

4a VIA 2010E 39

What is Impeding

Statistical Literacy

Math is the most privileged discipline in academia.

Math and statistics have successfully resisted all attempts to support statistical literacy.

This resistance is not a commission: a statement denying the need for statistical literacy.

This resistance is an omission: a total silence on whether math is responsible for deciding what various groups of students need.

4b V1A 201 CME

The Challenge



"Quantitative Literacy (QL), the ability to use numbers and data analysis in everyday life, is everybody's orphan.

Despite every person's need for QL, in the discipline-dominated K-16 education system in the United States, there is neither an academic home nor an administrative promoter for this critical competency." *Quantitative Literacy: Why Numeracy Matters.* p. 153 Bernard Madison

4b V1A 221 CME 41

Statistical Literacy Support
by NCTM Past President

"Statistical literacy has risen to the top of my advocacy list, right alongside numeracy, and perhaps even ahead of "algebra for all." By statistical literacy, I mean ... developing the ability to reason in the presence of, or under conditions of uncertainty. ... the facility to read and interpret statistical information and make informed inferences...." J. Michael Shaughnessy

www.statlit.org/pdf/2010Shaughnessy-StatisticsForAll-NCTM.pdf

Tension: Statistics vs. Stat Literacy



what most statisticians actually practice is typically more than the average person needs to be an informed citizen, intelligent consumer or skilled worker.

What everyone needs is typically called statistical thinking or statistical literacy, a crucial component of quantitative literacy."

Lynn Steen (2004). Achieving Quantitative Literacy p. 43



What Needs to be done? Support!

Mathematics Canada has a unique opportunity to become a world leader in supporting statistical literacy in grades 10-18.

The need is obvious, the tools are available. There is support from the American Statistical Association for multivariate thinking.

Lynn Steen (MAA past president) and J. Michael Shaughnessy (NCTM past president) support it.

Mathematics is a highly privileged discipline

Mathematics controls all of the quantitative courses taken in K-12.

Mathematics decides whether to offer algebra in 8th grade or 9th grade.

Mathematics decides what courses should be taken by students in non-quantitative majors.

No discipline has as much power as Mathematics.

V1A

Mathematics has great responsibility

With great power comes great responsibility!

Mathematics often polls other disciplines to see what they want for their students.

Problem: Most other disciplines don't know what mathematics their students should

Mathematics must take the lead. Mathematics must identify what students in all disciplines need.

V1A

Mathematics opportunities

Review the literature to see what students need to know about statistics.

Identify the math needed by all college graduates

Join with American statisticians (ASA) in supporting a multivariate focus on observational studies with a strong emphasis on confounding.

Support the National Numeracy Network.

References

Business Insider (2014). http://www.businessinsider.com/heres-the-averagesat-score-for-every-college-major-2014-10

47

De Veaux, D. (2015). Introductory Statistics in the 21st Century. USCOTS slides

Schield, M. (2015). Statistical Inference for Managers. ASA www.statlit.org/pdf/2015-Schield-ASA.pdf

Schield, M. (2014). Two Big Ideas for Teaching Big Data: ECOTS. www.statlit.org/pdf/2014-Schield-ECOTS.pdf

Schield, M. (2013). Reinventing Business Statistics. MBAA. www.StatLit.org/pdf/2013-Schield-MBAA.pdf

Tintle, Chance, Cobb, Rossman, Roy, Swanson & VanderStoep (2014)
Challenging the state of the art in post-introductory statistics.
http://2013.isiproceedings.org/Files/IPS032-P1-S.pdf



Audience

40%: School teachers: Current or previous

30%: College faculty: Current or previous

20%: Education, non-profit

10%: Industry, commercial

Statistical Literacy 2017

Milo Schield, Augsburg College

Elected Member: International Statistical Institute
US Rep: International Statistical Literacy Project
VP. National Numeracy Network

CME Presentation in Toronto
Fields Institute
April 29, 2016

www.StatLit.org/pdf/2017-Schield-CME-Slides.pdf

Statistical Literacy 2017: Overview

"We teach the wrong stuff; We teach it the wrong way; We teach it in the wrong order." Richard de Veaux

Statistical Literacy 2017:

- 1. What is it in general?
- 2. Who needs it?
- 3. What is it in particular?
- 4. Who can implement it?

What are Statistics?

- a. Data; numerical data, classifications of data, or numerical summaries of data [Ambiguous]
- b. Outcomes from a random process;randomly-selected or randomly-assigned groups[Technical distinction]
- c. Numbers in context where the **context matters**: Quantitative summaries of real things: things that have natures, connections & causes

Statistics is Different from Mathematics

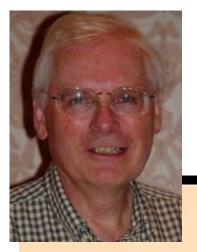
Math ignores the context.

- a. Math deals with form (ignores the matter)
- b. Math deals with variables and values (no natures)
- c. Math deals with associations and co-variates
- d. Math has no operator for "causes"

Statistics depends on the context

- a. Statistics deals with the matter: its nature
- b. Statistics deals with subjects and characteristics
- c. Statistics deals with "confounders"
- d. Statistics deals with "causes"

201 CME **7**



Mathematics: Patterns vs. Nature

Philosophically, mathematics is

not a part of science.

Mathematics studies patterns,

science studies nature.

Lynn Steen

1a

Saying "Statistics Come From Data" is like saying "Babies Come from Hospitals"





It's true but it leaves out the interesting details

Statistics are answers to questions or interests.

What is statistical literacy? In general terms

Statistical literacy is needed by citizens and social decision makers to enable them to understand and evaluate the statistics they encounter everyday.

Everyday statistics are used as evidence in arguments. Legal:

- Describe: 90% of a restaurant's staff speaks Spanish
- Compare: Most Mexican restaurant staff speak Spanish
- Evaluate: Mexican restaurants discriminate in hiring.

What is statistical literacy? Examples:

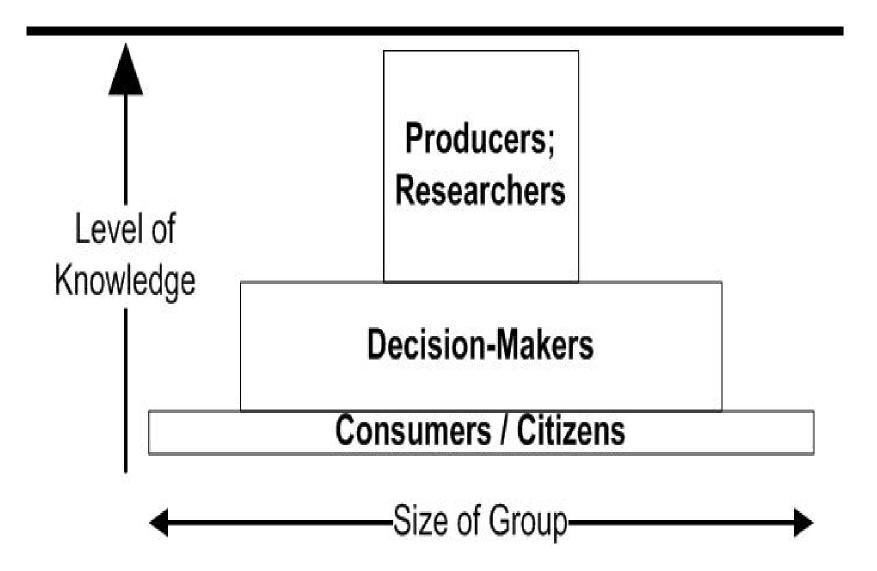
Medical:

- Describe: Japanese, who live long, eat low-fat diet
- Compare: People with high-fat diets die sooner
- Evaluate: High-fat diet causes shorter lifespan.

Social:

- Describe: Average school class size is 24
- Compare: Best performing classes are smaller
- Evaluate: Smaller classes will improve outcomes

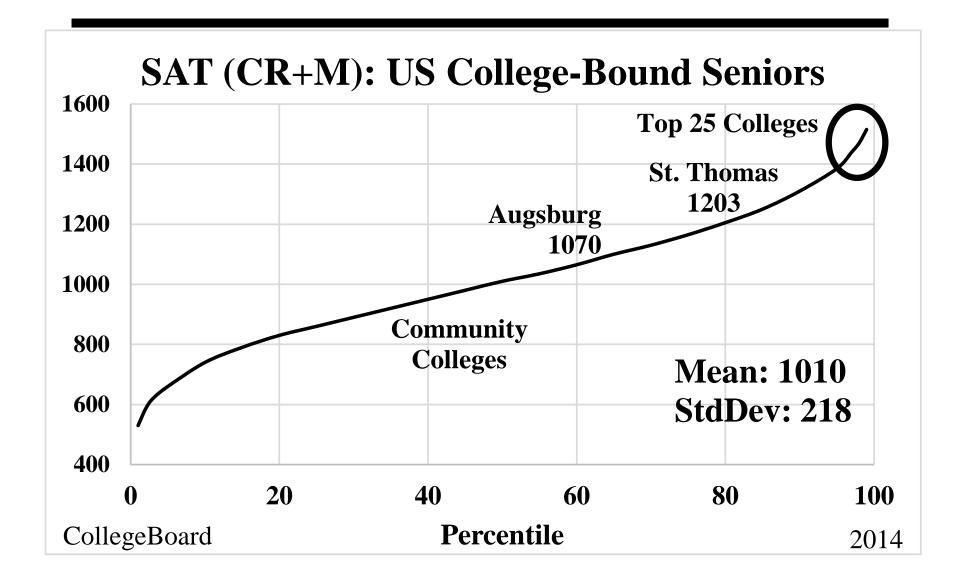
Who Needs Statistical Literacy? Three Audiences



Three Audiences: More detail

- 1. STEM majors and those who conduct surveys, studies and research.
- 2. Social decision-makers:
 Politicians, bureaucrats, business leaders, doctors
 - Those who inform citizens and decision makers: journalists, analysts, lawyers, economists, consultants, sociologists, political scientists, policy advocates, psychologists and educators.
- 3. Citizen in a modern republic or democracy.

College-Bound Students: Wide variation



College-Bound US Students SAT Math Scores by Major

PERCENTILE	MAJOR	SAT MATH
80%	Math/Stats	613
72%	Physical Sciences	585
70%	Engineering	579
62%	Computer Science	554
61%	Biological Sciences	551
61%	Social Sciences	550
51%	Business & English	522
46%	History	506
43%	Communication	498
40%	Psychology	489
38%	Education	482

Business Insider (2014). 2014 SAT scores

Distribution of US College Graduates

STEM Majors (11%)

Math, Science, Engineering, Biological

Non-STEM Quantitative Majors (46%)

Business, Social Sciences, Health, Psych

Non-Quantitative Majors (43%)

Education, English, Humanities

US Statistical Abstract 2012, Table 302

Harvard Business Review: Website Search of 40K Items

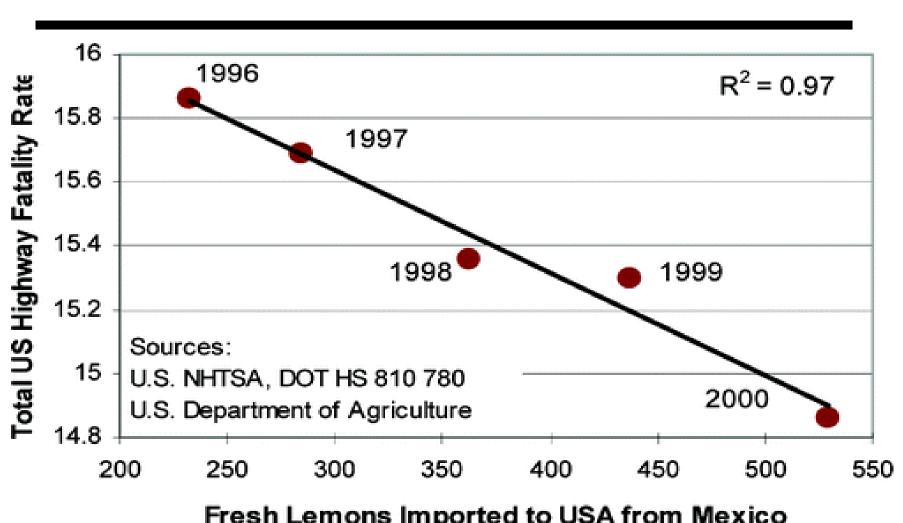
#	INFERENTIAL	1
22	"clinical trial"	18
7	"statistical significa	ince"
4	"statistically significant	cant"
3	"standard error"	
1	"sampling error"	
1	"margin of error"	
1	"prediction interval	
1	p-value	
0	"sampling distributi	on"
0	"confidence interva	al"
0	"null hypothesis"	
0	"reject the null"	
0	"random assignmen	nt"

CONTROL/CONFOUND				
2,263	control			
234	"control of" 200			
113	"take (ing) into account"			
30	"compensate (ing) for"			
19	"control (ed, ing) for"			
18	confound (er, ing)			
17	"adjust(ed, ing) for"			
3	"sampling bias"			
0	"alternate explanation"			
0	"common cause"			
0	"effect modifier"			
0	"Simpson's paradox"			
0	"lurking variable"			

Statistical Literacy: More Detail

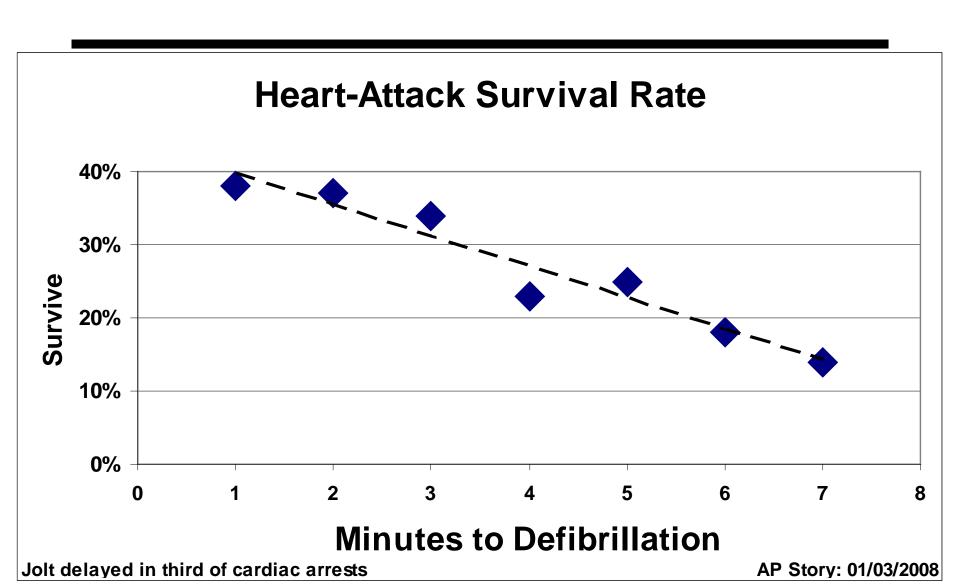
- 3a. Association vs. Causation
- 3b. Classify all the influences on a statistic
- Context: Confounding and study design
- Assembly/assumptions: How things are defined.
- Randomness: Unlikely is almost certain given enough trials.
- Error/Bias:

Association: Probably Not Causation

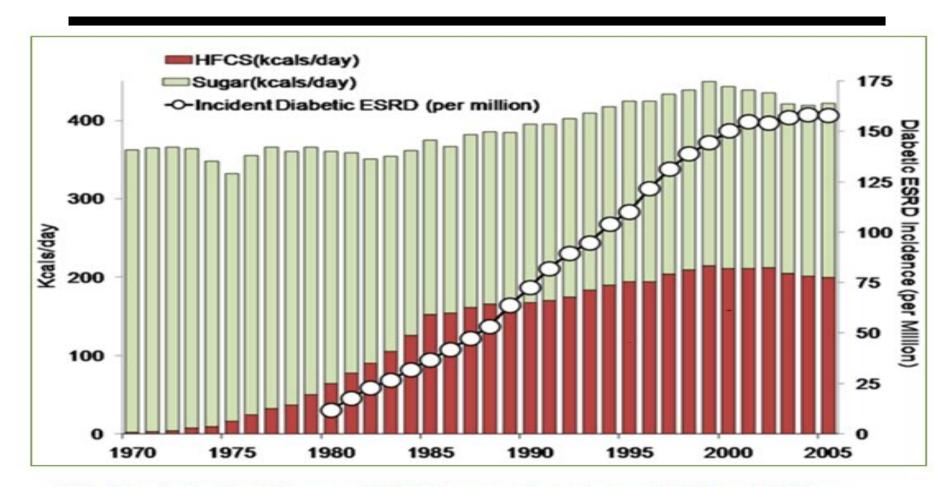


Fresh Lemons Imported to USA from Mexico (Metric Tons)

Association: Probably Causation



Association: Possibly Causation



U.S. Trends in Total Sugar and High Fructose Corn Syrup (HFCS) availability, and Incident Diabetic End-Stage Renal Disease (ESRD)

Distinguish Causation from Association

Causation (8%): cause, effects, results, prevents

Association (2%): associate, relate, correlate,

Between (67%):

Action verbs: ups, cuts, raises, boosts, increases Other: due to, because of, attributed to

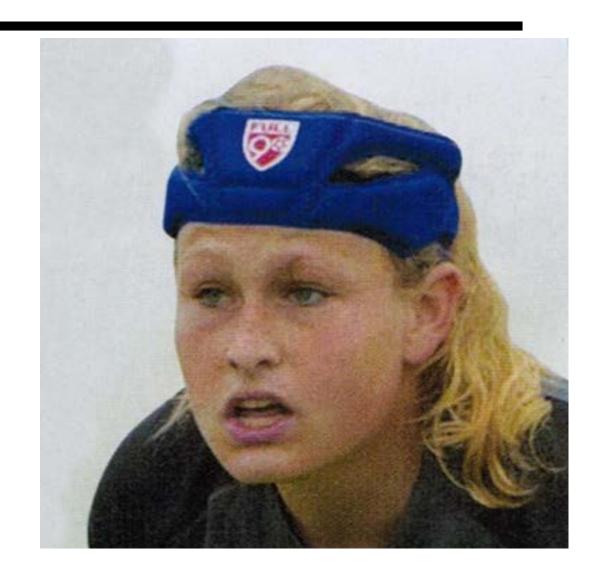
Inappropriate use of "causes":

- Obesity causes later onset of puberty in boys
- Junk food causes a third of heart attacks.

Action-Verb Association

"Research shows that the headgear reduces the concussion rate by more than 50 percent."

8/2011 P. 41



Association-Causation

Baseball players whose names begin with the letter "D" are *more likely to die young*

Drinking a full pot of coffee every morning will add years to your life, but one cup a day increases the risk of pancreatic cancer.

Asian-Americans are *most susceptible* to heart attacks on the fourth day of the month

Source: Standard Deviations: Flawed Assumptions, Tortured Data, and Other Ways to Lie with Statistics by Gary Smith (2015).

Pie Chart: Compare

Protestants are twice as likely to be smokers as are

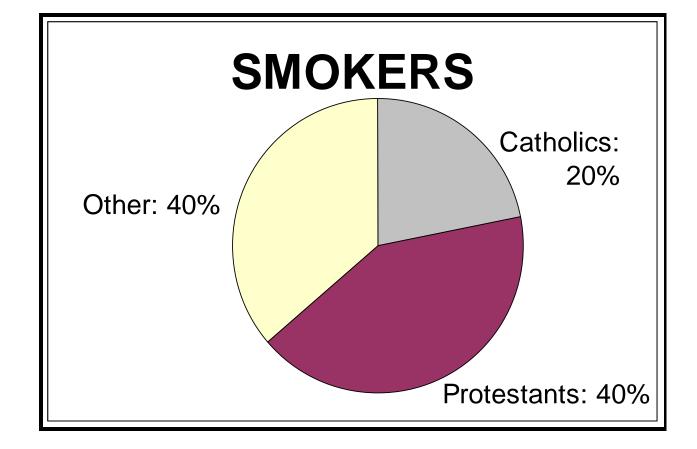
Catholics?

NO:

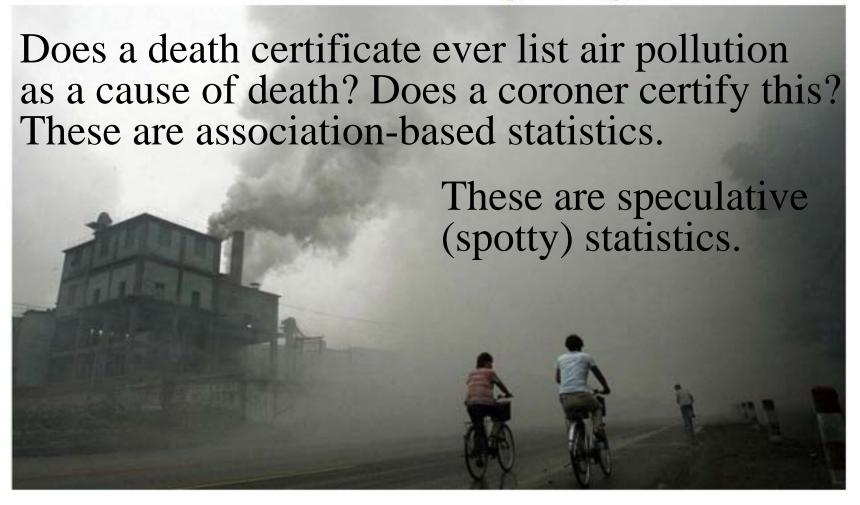
Smoker is whole.

Student error rate:

62%



Air Pollution Linked to 6.5 Million Deaths a Year, Study Says



Melissa Chan Time June 27, 2016

Association vs. Causation 11 Headlines, Same Story

- 1. Study: 45,000 Uninsured *Die* a Year (CBS News)
- 2. 45,000 deaths attributable to uninsurance
- 3. 45,000 US deaths associated with lack of insurance
- 4. No health coverage *tied to* 45,000 deaths a year
- 5. Lack of insurance *linked to* 45,000 deaths
- 6. Study: 45,000 U.S. Deaths *From* Lack of Insurance
- 7. One death every 12 minutes *due to* no health insurance
- 8. 45,000 ... die because of lack of health insurance
- 9. Lack of Health Insurance Kills 45,000 a Year
- 10. Lack of Health Insurance cause 44,789 deaths
- 11. Lack of insurance to blame for almost 45,000 deaths

Stats = Premise: Crit. Thinking Stats = Conclusion: Stat Literacy

The Point

An argument is like a house.
The more disputable the point,
the stronger the evidence
(the walls and floors) must be.

Statistics As Evidence

"All statistics are socially constructed So, "Take CARE"!! Statistics may be influenced by:

C A R E
Context Assembly Randomness Error

Statistical Literacy in detail: "Take CARE"

Statistical literacy studies all influences on statistic:

- Confounding:
 - what was and was not controlled for
 - what kind of study was involved
- Assembly/Assumptions:
 - how statistics are collected, defined and grouped
 - how statistics are summarized, compared & presented
- Randomness: small samples and big data
- Error/bias

Confounding: Using Ordinary English

- 1) The percentage of women who are runners.
- 2) The percentage of women among runners.
- 3) The death rate of men is X per 100,000.
- 4) The men's rate of death is X per 100,000
- 5) Toyota is the car most frequently stolen.
- 6) Toyota is the car most likely to be stolen.
- 7) Cadillac is the car most likely to be stolen.

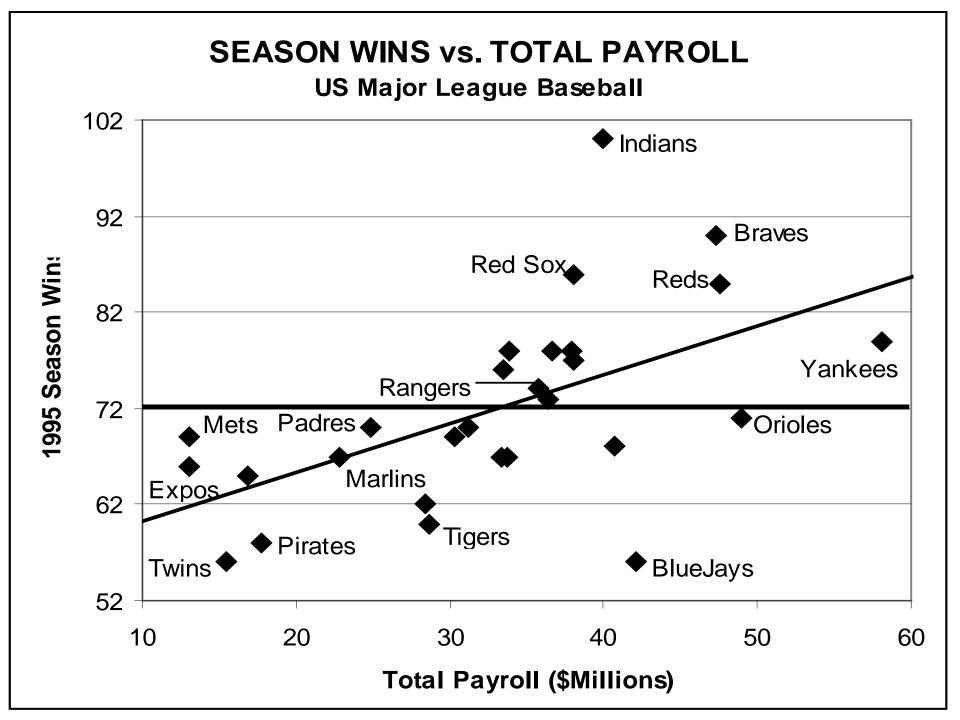
Small Change in Syntax; Big Change in Semantics



Confounding: Mixed-Fruit vs. Apples-Apples Comparison

Ave Weight	Ht=64"	Ht=70"
FEMALE	129 27	7-44
MALE	41	[#] 156

Ave Weight	Ht=64"	Ht=70"		
FEMALE	129 13	# 142 ₁₄		
MALE		156		



US SAT-VERBAL SCORES

Average SAT-V	1981	2002	Change	1981	2002
All Test-Takers	504	504	0	100%	100%
White	519	527	8	85%	65%
Black	412	431	19	9%	11%
Asian	474	501	27	3%	10%
Mexican	438	446	8	2%	4%
Puerto Rican	437	455	18	1%	3%
American Indian	471	479	8	0%	1%

Study design can inhibit certain kinds of confounders

Experiment

Treatment is assigned

A+ Repeatable

A- Randomize

Observational

34

Exposure not assigned

C Longitudinal

D Snapshot

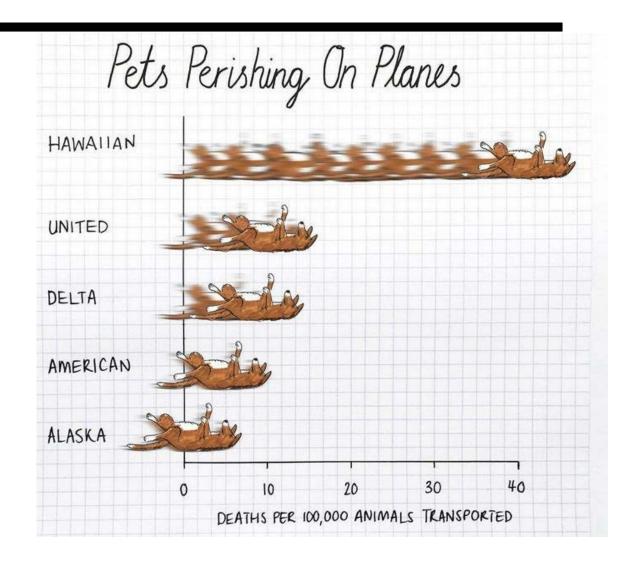
B Quasi (Queasy)-Experiment

Not repeatable; not randomized Nature or humans intervene

Assembly: What to control for: United has <u>Worst Pet Record</u>

Nine pets died while being transported by United while another 14 were injured last year.

Most of any US airline...



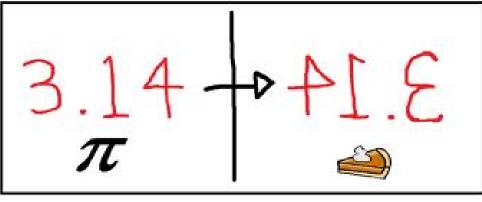
Assembly: Making small things big

7 nanograms per gram = 7 parts in a billion

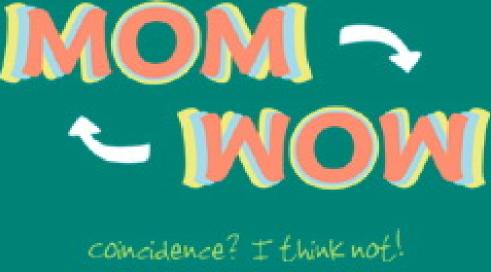


Randomness: Coincidence?









Error/Bias

A recent survey shows that most Republicans surveyed prefer Obama as President.

Question: Who would you prefer as President?

- Barack Obama
- The captain of the Italian linear that crashed
- Charlie Sheehan
- Lady Gaga

What is Impeding Statistical Literacy

Math is the most privileged discipline in academia.

Math and statistics have successfully resisted all attempts to support statistical literacy.

This resistance is not a commission: a statement denying the need for statistical literacy.

This resistance is an omission: a total silence on whether math is responsible for deciding what various groups of students need. **1h** V1A 201 CME





"Quantitative Literacy (QL), the ability to use numbers and data analysis in everyday life, is everybody's orphan.

Despite every person's need for QL, in the discipline-dominated K-16 education system in the United States, there is neither an academic home nor an administrative promoter for this critical competency." *Quantitative Literacy: Why Numeracy Matters.* p. 153 Bernard Madison

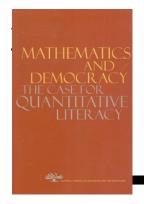
Statistical Literacy Support by NCTM Past President

"Statistical literacy has risen to the top of my advocacy list, right alongside numeracy, and perhaps even ahead of "algebra for all."

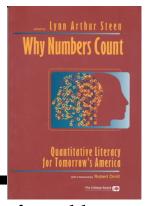


By statistical literacy, I mean ... developing the ability to reason in the presence of, or under conditions of uncertainty. ... the facility to read and interpret statistical information and make informed inferences...." J. Michael Shaughnessy

www.statlit.org/pdf/2010Shaughnessy-StatisticsForAll-NCTM.pdf



Tension: Statistics vs. Stat Literacy



what most statisticians actually practice is typically more than the average person needs to be an informed citizen, intelligent consumer or skilled worker.

What everyone needs is typically called statistical thinking or statistical literacy, a crucial component of quantitative literacy."

Lynn Steen (2004). Achieving Quantitative Literacy p. 43

What Needs to be done? Support!

Mathematics Canada has a unique opportunity to become a world leader in supporting statistical literacy in grades 10-18.

The need is obvious, the tools are available. There is support from the American Statistical Association for multivariate thinking.

Lynn Steen (MAA past president) and J. Michael Shaughnessy (NCTM past president) support it.

Mathematics is a highly privileged discipline

Mathematics controls all of the quantitative courses taken in K-12.

Mathematics decides whether to offer algebra in 8th grade or 9th grade.

Mathematics decides what courses should be taken by students in non-quantitative majors.

No discipline has as much power as Mathematics.

Mathematics has great responsibility

With great power comes great responsibility!

Mathematics often polls other disciplines to see what they want for their students.

Problem: Most other disciplines don't know what mathematics their students should

Mathematics must take the lead. Mathematics must identify what students in all disciplines need.

Mathematics opportunities

Review the literature to see what students need to know about statistics.

Identify the math needed by all college graduates

Join with American statisticians (ASA) in supporting a multivariate focus on observational studies with a strong emphasis on confounding.

Support the National Numeracy Network.

References

- Business Insider (2014). http://www.businessinsider.com/heres-the-average-sat-score-for-every-college-major-2014-10
- De Veaux, D. (2015). Introductory Statistics in the 21st Century. USCOTS slides
- Schield, M. (2015). Statistical Inference for Managers. ASA www.statlit.org/pdf/2015-Schield-ASA.pdf
- Schield, M. (2014). Two Big Ideas for Teaching Big Data: ECOTS. www.statlit.org/pdf/2014-Schield-ECOTS.pdf
- Schield, M. (2013). Reinventing Business Statistics. MBAA. www.StatLit.org/pdf/2013-Schield-MBAA.pdf
- Tintle, Chance, Cobb, Rossman, Roy, Swanson & VanderStoep (2014) Challenging the state of the art in post-introductory statistics. http://2013.isiproceedings.org/Files/IPS032-P1-S.pdf