Statistical Literacy

Milo Schield

University of New Mexico Statistical Literacy Coordinator Fellow, American Statistical Association US Coordinator, International Statistical Literacy Project

Sept 2, 2022 Slides and paper:

www.StatLit.org/pdf/

2022-Schield-Metro-State-Slides.pdf 2022-Schield-Metro-State.pdf APPROACH

- 1. Review ASA 2005 and 2016 GAISE guidelines
- 2. Investigate multivariate thinking and confounding
- 3. Are statistical educators adverse to confounding?
- 4. UNM offers Math 1300 Statistical Literacy
- 5. Statistical Literacy and confounding
- 6. Work problems involving confounding
- 7. Course outcomes and components
- 8. Student comments
- 9. Conclusions

1 vi

ASA 2005 GAISE Guidelines: Six Recommendations

- 1. Emphasize <u>statistical literacy</u> and develop statistical thinking
- 2. Use real data
- 3. Stress conceptual understanding, rather than mere knowledge of procedures
- 4. Foster active learning in the classroom
- 5. *Use technology* for developing conceptual understanding and analyzing data
- 6. Use assessments to improve & evaluate student learning

#1

ASA 2005 GAISE Guidelines: Definitions

<u>Statistical literacy</u>: understanding the basic language of statistics (e.g., knowing what statistical terms and symbols mean and being able to read statistical graphs) and fundamental ideas of statistics.

Statistical thinking: the type of thinking that statisticians use when approaching or solving statistical problems. *The practical operation of statistics is to collect and analyze data to answer questions.*

#1 vı

2022 Schield StatLit Metro State

ASA 2005 GAISE Guidelines: Learning Outcomes

Students should believe and understand why:

- 1. Data beat anecdotes
- 2. Variability is natural, predictable, and quantifiable
- 3. Random sampling allows results of surveys and experiments to be extended to the population...
- 4. Random assignment in comparative experiments allows cause-and-effect conclusions to be drawn
- 5. Association is not causation

#1b vı

2022 Schield StatLit Metro State

2016 GAISE Guidelines: Recommendations

- 1. "Teach statistical thinking.
 - Teach statistics as an investigative process of problem-solving and decision-making.
 - Give students experience with multivariable thinking
- 2. Focus on conceptual understanding.
- 3. Integrate real data with a context and purpose.
- 4. Foster active learning.
- 5. Use technology to explore concepts and analyze data.
- 6. Use assessments to improve & evaluate student learning"

#1b v1

2016 GAISE Guidelines: Clarification

"There is no single introductory statistics course.

- Some introductory courses address <u>statistical</u> literacy, while others focus on statistical methods.
- This distinction is sometimes referred to as courses for <u>consumers</u> versus those for <u>producers</u> of analyses..."

#1b v1

2016 GAISE Guidelines: Goals of Introductory Statistics

"desired result of all introductory statistics courses is to produce <u>statistically educated</u> students, which means that students should develop the ability to <u>think statistically</u>.

 Students should become <u>critical consumers</u> of statistically-based results <u>reported in</u> <u>popular media</u>, recognizing whether reported results reasonably follow from the study and analysis conducted."

#2 vi

2016 GAISE guidelines: Multivariable Thinking

- "Multivariable models are necessary when we want to model many aspects of the world more realistically. The real world is complex and can't be described well by one or two variables.
- If students do not have exposure to <u>simple tools for</u> <u>disentangling complex relationships</u>, they may dismiss statistics as an old-school discipline only suitable for small sample inference of randomized studies.
- ... when we don't demonstrate <u>realistic models</u> students are left with the impression that statistics is trivial and not really useful."

#2 vı

2016 GAISE Guidelines: Multivariable & Confounding

Why should we teach multivariable thinking?

- Much of today's data is multivariate.
- "Perhaps the best place to start is to consider how a third variable can change our understanding of the relationship between two variables."
- students should understand the "principles of statistical design and tools to assess and account for the possible <u>impact of other measured and</u> unmeasured confounding variables." ASA 2014

#2 v1

2016 GAISE guidelines: Multivariable & Confounding

Introduce multivariable thinking in stages:

- 1. identify observational studies,
- 2. explain why randomized assignment to treatment improves the situation,
- 3. learn to be wary of cause-and-effect conclusions from observational studies...
- 4. consider potential confounding factors, and
- 5. use simple approaches (such as stratification) to address confounding

#3 vi SERJ "Statistical Literacy":
Statistical educators adverse?

Of the 17 papers, just 4 mention 'statistical literacy'.

- Future of Statistical Literacy is the Future of Statistics: Ridgway and Nicholson
- Statistical Literacy as the Earth Moves: Wild
- Statistical Literacy for Active Citizenship: Engel
- GAISE 2016 Promotes Statistical Literacy: Schield

https://iase-web.org/ojs/SERJ/issue/view/12



"The 2016 revision eliminated statistical literacy as a stated goal.

Although this looks like a rejection, this paper argues that by including multivariate thinking and – more importantly – confounding as recommended topics in introductory statistics, statistical literacy has in fact been accepted if not promoted.

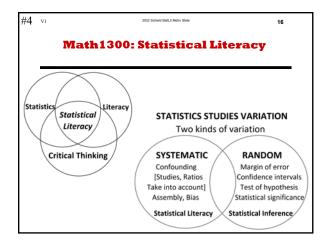
The adoption of the new guidelines will greatly advance students' statistical literacy: the ability to read and interpret statistics relevant to consumers and decision makers."

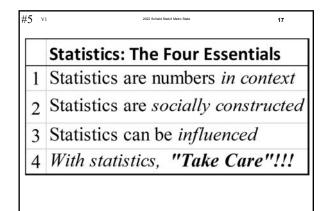


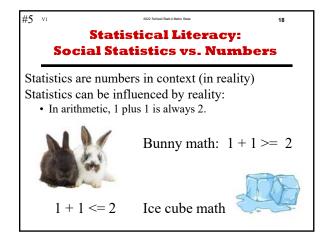
#4 vi 2021-22: UNM Offers Math1300 Statistical Literacy

This confounder-based course is different:

- designed for students in non-quantitative majors
- designed as an alternative to Stat 101.
- satisfies math requirement in GenEd curriculum.
- does not use any computer software.
- has less than a 30% overlap with Stat 101.







#5 vi 202 Statistics are Socially Constructed"



Joel Best, author of "Lies,

Damned Lies and Statistics" identified this fact as the most important, the most fundamental, aspect of all reality-based statistics.

Statistics, just like words, are created by people: people with motives, values and goals.

#5 VI 2022 Schall Statistics can Be Influenced



Q. Best advice when dealing with statistics?

A. "Take CARE". Statistics can be influenced.

All influences are grouped into four categories:

C: Confounding: Confused by related factors

A: Assembly: how things are defined, counted, etc.

R: Randomness

E: Error (including bias)

Admonition: "Take CARE"



Students like "CARE". It gives them a structure.

When asked to rank what idea they considered the most valuable, students chose "Take CARE".

#5 vi ZOZ BANKE DRAIL MAD DRAIL Today's student need to study Statistics

Disparities in

- Education, suspensions and graduation
- Policing, crime, sentencing and prison
- Wages, income, assets, loans and wealth
- · Health, health care, homicides and deaths

Disparities by

gender, race, ethnicity, religion, politics, age, etc.

All of these rely on statistics: social statistics.

#6 vi 2022 Schuld Stall More Stale Confounding

Association: People who read home and fashion magazines are more likely to get pregnant than people who read car and sport magazines.

We know that pregnancy isn't caused by magazines. We know that only women can get pregnant.

We recognize that women are more likely to read home and fashion magazines than men.

QED. This association is confounded by gender.

#6 VI 2022 SOURCE SEALS MADE SEAL

Covid vaccinated cases are MORE likely to die than unvaccinated cases.

Covid Death Ra	ates Per Cas	ie		
	Crude Rate	# Cases		
Unvaccinated	0.17%	151,052		
Vaccinated	0.41%	117,114		
41/17 = 2.4		UK NHS 2021		

This result is unexpected. Doesn't make sense.



Who are LESS likely to die?

Under 50: Vaccinated. 50+? Vaccinated

Confounder: "It's the mix!"

Elderly: 23% of vaccinated, 2% of unvaccinated

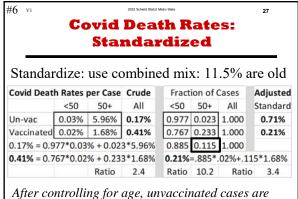
#6 vi **Covid Death Rates** Simpson's Paradox

Paradox: Two group comparison has opposite direction from that in each of the sub-groups. Impossible with counts. Possible with ratios.

Simpson's paradox: the elephant in the room for the social sciences.

What causes it? It's the "mix"!

A **crude comparison**: a "mixed fruit" comparison!



3.4 times as likely to die as are vaccinated cases.

Confounding is the elephant in observational statistics

Teachers know this.

Not in intro. statistics or research methods.

It should be taught in an introductory course.

There isn't time in traditional statistics.



#7 vi **StatLit Course Outcomes:** Students should be able to ...

- read and interpret the statistics encountered in everyday life: in statements, tables and graphs.
- think hypothetically about how the statistics could have been influenced.
- evaluate the sensitivity of the statistics to being influenced.
- communicate their findings and conclusions using ordinary English.

#6 vi Students should understand "Taking into Account"

Taking into account the influence of a related factor converts a mixed-fruit comparison (apples and oranges) into a same fruit comparison (apples and apples).

Taking something into account can reverse the direction of a comparison: Simpson's paradox.

#7 vı **Math1300: Statistical Literacy** Four Parts (% of grade)

25% Odyssey writing forum: 1-2 cases / week.

20% Exercises: 15-20 multi-choice (2 try) / week. Write 2-4 one-line statements: Manual grade. Compare counts; describe ratios (part/whole)

5%: Attendance:

50%: Quizzes (two 12%@) and Final (26%)

Textbook: Schield 2021 and 2021B.

1. Odyssey Writing Forum

Each case is a challenge. No research required.

Odyssey forum is different:

#7 vi

- 1. Everyone is anonymous (no avatar)
- 2. No free-riders. Can't see others till after you post
- 3. Everyone grades everyone else.
- 4. Odyssey computes each student's power.

More info: www.statlit.org/pdf/2014-Schield-ICOTS.pdf

2. Exercises

Multiple choice: field tested (1,000+ students). Immediate feedback on right-wrong. Helpful hints/comments on why it is wrong. Closely aligned to the textbook.

One-line essay:

Compare two counts: difference, ratio, % more. Describe part-whole ratios from tables & graphs Use Percent and Percentage grammar.

3. Textbook: Eight Chapters

- 1. The story behind the statistics
- 2. Comparisons and CARE solutions
- 3. Understanding Measurements
- 4. Describing ratios: Percent & Percentage
- 5. Describing ratios: Rate, Chance and Ratio
- 6. Comparing ratios: Named ratios and Likely
- 7. Interpreting confusing ratios
- 8. Randomness

#8 vı 4. Student Evaluations www.statlit.org/UNM.htm

"I've despised every other math class I've ever taken but this one is actually enjoyable. It also seems more useful than any math course I've ever taken."

"I enjoyed critical thinking and the news stories. Both provide beneficial knowledge I will take with me into my everyday life."

#8 vı 4. Student Evaluations www.statlit.org/UNM.htm

"This course is an answer to my prayers, I am a music major and horrible at math so fulfilling my math requirement has been hard.

This is the first math class I actually liked. I loved the format; the material is about things I can apply to everyday life.

The textbook is fantastic and helped me a lot...

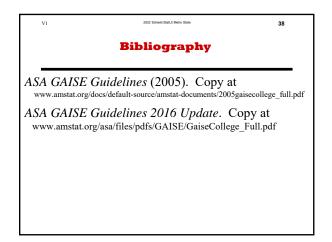
I would recommend this class for anyone."



Statistical Literacy involves statistics, writing and critical thinking. This course is very different from a typical mathematics course.

Statistical Literacy must be tailored to the students involved. First year-students are very different from Seniors or Honors students.

Students see value in taking Math1300. To help their students think critically about statistics, other colleges should offer Statistical Literacy.



V1 2022 Schield StatLit Metro State

Statistical Literacy

Milo Schield

University of New Mexico
Statistical Literacy Coordinator
Fellow, American Statistical Association
US Coordinator, International Statistical Literacy Project

Sept 2, 2022 Slides and paper:

www.StatLit.org/pdf/

2022-Schield-Metro-State-Slides.pdf 2022-Schield-Metro-State.pdf V1 2022 Schield StatLit Metro State 2

APPROACH

- 1. Review ASA 2005 and 2016 GAISE guidelines
- 2. Investigate multivariate thinking and confounding
- 3. Are statistical educators adverse to confounding?
- 4. UNM offers Math 1300 Statistical Literacy
- 5. Statistical Literacy and confounding
- 6. Work problems involving confounding
- 7. Course outcomes and components
- 8. Student comments
- 9. Conclusions

ASA 2005 GAISE Guidelines: Six Recommendations

- 1. Emphasize <u>statistical literacy</u> and develop statistical thinking
- 2. Use real data
- 3. Stress conceptual understanding, rather than mere knowledge of procedures
- 4. Foster active learning in the classroom
- 5. Use technology for developing conceptual understanding and analyzing data
- 6. Use assessments to improve & evaluate student learning

ASA 2005 GAISE Guidelines: Definitions

Statistical literacy: understanding the basic language of statistics (e.g., knowing what statistical terms and symbols mean and being able to read statistical graphs) and fundamental ideas of statistics.

Statistical thinking: the type of thinking that statisticians use when approaching or solving statistical problems. *The practical operation of statistics is to collect and analyze data to answer questions.*

ASA 2005 GAISE Guidelines: Learning Outcomes

5

Students should believe and understand why:

- 1. Data beat anecdotes
- 2. <u>Variability</u> is natural, predictable, and quantifiable
- 3. Random sampling allows results of surveys and experiments to be extended to the population...
- 4. Random assignment in comparative experiments allows cause-and-effect conclusions to be drawn
- 5. Association is not causation

2016 GAISE Guidelines: Recommendations

- 1. "Teach statistical thinking.
 - Teach statistics as an investigative process of problem-solving and decision-making.
 - Give students experience with multivariable thinking
- 2. Focus on conceptual understanding.
- 3. Integrate real data with a context and purpose.
- 4. Foster active learning.
- 5. Use technology to explore concepts and analyze data.
- 6. Use assessments to improve & evaluate student learning"

2016 GAISE Guidelines: Clarification

"There is no single introductory statistics course.

- Some introductory courses address <u>statistical</u> <u>literacy</u>, while others focus on <u>statistical methods</u>.
- This distinction is sometimes referred to as courses for <u>consumers</u> versus those for <u>producers</u> of analyses..."

2016 GAISE Guidelines: Goals of Introductory Statistics

"desired result of all introductory statistics courses is to produce <u>statistically educated</u> students, which means that students should develop the ability to <u>think statistically</u>.

1. Students should become <u>critical consumers</u> of statistically-based results <u>reported in popular media</u>, recognizing whether reported results reasonably follow from the study and analysis conducted."

2016 GAISE guidelines: Multivariable Thinking

- "Multivariable models are necessary when we want to model many aspects of the world more realistically. The real world is complex and can't be described well by one or two variables.
- If students do not have exposure to <u>simple tools for</u> <u>disentangling complex relationships</u>, they may dismiss statistics as an old-school discipline only suitable for small sample inference of randomized studies.
- ... when we don't demonstrate <u>realistic models</u> students are left with the impression that statistics is trivial and not really useful."

2016 GAISE Guidelines: Multivariable & Confounding

Why should we teach multivariable thinking?

- Much of today's data is multivariate.
- "Perhaps the best place to start is to consider how a third variable can change our understanding of the relationship between two variables."
- students should understand the "principles of statistical design and tools to assess and account for the possible impact of other measured and unmeasured confounding variables." ASA 2014

2016 GAISE guidelines: Multivariable & Confounding

Introduce multivariable thinking in stages:

- 1. identify observational studies,
- 2. explain why randomized assignment to treatment improves the situation,
- 3. learn to be wary of cause-and-effect conclusions from observational studies...
- 4. consider potential confounding factors, and
- 5. use simple approaches (such as stratification) to address confounding

SERJ "Statistical Literacy": Statistical educators adverse?

Of the 17 papers, just 4 mention 'statistical literacy'.

- Future of Statistical Literacy is the Future of Statistics: Ridgway and Nicholson
- Statistical Literacy as the Earth Moves: Wild
- Statistical Literacy for Active Citizenship: Engel
- GAISE 2016 Promotes Statistical Literacy: Schield

https://iase-web.org/ojs/SERJ/issue/view/12

GAISE 2016 Promotes Statistical Literacy. Statistical Educators Adverse?

"The 2016 revision eliminated statistical literacy as a stated goal.

Although this looks like a rejection, this paper argues that by including multivariate thinking and – more importantly – confounding as recommended topics in introductory statistics, statistical literacy has in fact been accepted if not promoted.

The adoption of the new guidelines will greatly advance students' statistical literacy: the ability to read and interpret statistics relevant to consumers and decision makers."

University of New Mexico is offering a new course!

Taught 7 sections (200 students) in 2021-22



Statistical Literacy



MATH 1300 (3)

Participants will study the social statistics encountered by consumers. Investigate the story behind the statistics. Study the influences on social statistics. Study the techniques used to control these influences. Strong focus on confounding.

Meets New Mexico General Education Curriculum Area 2: Mathematics and Statistics.

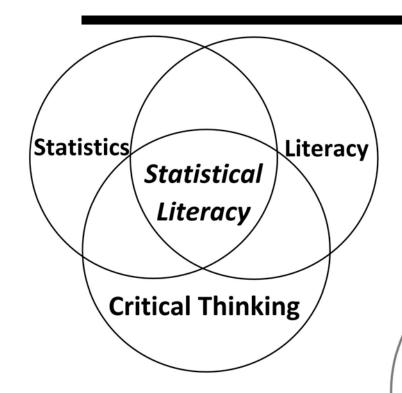
2021-22: UNM Offers Math1300 Statistical Literacy

15

This confounder-based course is different:

- designed for students in non-quantitative majors
- designed as an alternative to Stat 101.
- satisfies math requirement in GenEd curriculum.
- does not use any computer software.
- has less than a 30% overlap with Stat 101.

Math1300: Statistical Literacy



STATISTICS STUDIES VARIATION

Two kinds of variation

SYSTEMATIC

Confounding [Studies, Ratios Take into account] Assembly, Bias

Statistical Literacy

RANDOM

Margin of error
Confidence intervals
Test of hypothesis
Statistical significance

Statistical Inference

Statistics: The Four Essentials Statistics are numbers in context Statistics are socially constructed 3 Statistics can be *influenced* 4 With statistics, "Take Care"!!!

Statistical Literacy: Social Statistics vs. Numbers

Statistics are numbers in context (in reality)
Statistics can be influenced by reality:

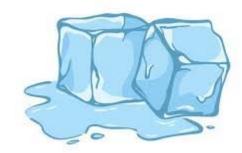
• In arithmetic, 1 plus 1 is always 2.



Bunny math: 1 + 1 >= 2

$$1 + 1 <= 2$$

Ice cube math



#5 v1

"Statistics are Socially Constructed"

Joel Best, author of "Lies, Damned Lies and Statistics" identified this fact as the most important, the most fundamental, aspect of all reality-based statistics.

Statistics, just like words, are created by people: people with motives, values and goals.

#5 v1

Statistics can Be Influenced



- Q. Best advice when dealing with statistics?
- A. "Take CARE". Statistics can be influenced.

All influences are grouped into four categories:

C: Confounding: Confused by related factors

A: Assembly: how things are defined, counted, etc.

R: Randomness

E: Error (including bias)

#5 v1 2022

Admonition: "Take CARE"



Students like "CARE". It gives them a structure.

When asked to rank what idea they considered the most valuable, students chose "Take CARE".

Today's student need to study Statistics

Disparities in

- Education, suspensions and graduation
- Policing, crime, sentencing and prison
- Wages, income, assets, loans and wealth
- Health, health care, homicides and deaths

Disparities by

gender, race, ethnicity, religion, politics, age, etc.

All of these rely on statistics: social statistics.

Confounding

Association: People who read home and fashion magazines are more likely to get pregnant than people who read car and sport magazines.

We know that pregnancy isn't caused by magazines. We know that only women can get pregnant.

We recognize that women are more likely to read home and fashion magazines than men.

QED. This association is confounded by gender.

Where is the Math? Confounding: Covid Deaths

Covid vaccinated cases are MORE likely to die than unvaccinated cases.

Covid Death Rates Per Case						
	Crude Rate	# Cases				
Unvaccinated	0.17%	151,052				
Vaccinated	0.41%	117,114				
41/17 = 2.4		UK NHS 2021				

This result is unexpected. Doesn't make sense.

Covid Death Rates by Age

Covid Death Rates Per Case				Death Rat	es by Age	
	# Cases		<50	50+		
Unvaccinated	0.17%	151,052		0.03%	5.96%	
Vaccinated	0.41%	117,114		0.02%	1.68%	
41/17 = 2.4		UK NHS 2021				

Who are LESS likely to die?

Under 50: Vaccinated. 50+? Vaccinated

Confounder: "It's the mix!"

Elderly: 23% of vaccinated, 2% of unvaccinated

Covid Death Rates Simpson's Paradox

Paradox: Two group comparison has opposite direction from that in each of the sub-groups.

Impossible with counts. Possible with ratios.

Simpson's paradox: the elephant in the room for the social sciences.

What causes it? It's the "mix"!

A crude comparison: a "mixed fruit" comparison!

Covid Death Rates: Standardized

Standardize: use combined mix: 11.5% are old

Covid Death Rates per Case		Crude		Fraction of Cases				Adjusted	
	<50	50+	All		<50	50+	All		Standard
Un-vac	0.03%	5.96%	0.17%		0.977	0.023	1.000		0.71%
Vaccinated	0.02%	1.68%	0.41%		0.767	0.233	1.000		0.21%
0.17% = 0.977*0.03% + 0.023*5.96%				0.885	0.115	1.000			
0.41% = 0.767*0.02% + 0.233*1.68% 0.21% = .885*.02% + .115*1.68%						5*1.68%			
		Ratio	2.4		Ratio	10.2	Rati	0	3.4

After controlling for age, unvaccinated cases are 3.4 times as likely to die as are vaccinated cases.

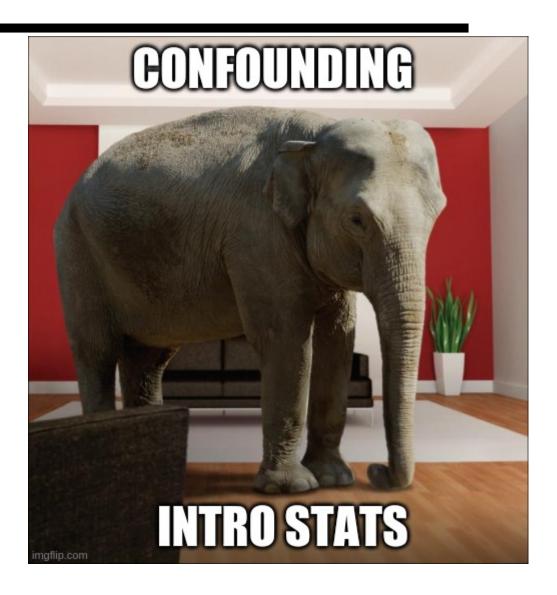
Confounding is the elephant in observational statistics

Teachers know this.

Not in intro. statistics or research methods.

It should be taught in an introductory course.

There isn't time in traditional statistics.



StatLit Course Outcomes: Students should be able to ...

- read and interpret the statistics encountered in everyday life: in statements, tables and graphs.
- think hypothetically about how the statistics could have been influenced.
- *evaluate* the sensitivity of the statistics to being influenced.
- communicate their findings and conclusions using ordinary English.

Students should understand "Taking into Account"

Taking into account the influence of a related factor converts a mixed-fruit comparison (apples and oranges) into a same fruit comparison (apples and apples).

Taking something into account can reverse the direction of a comparison: Simpson's paradox.

Math 1300: Statistical Literacy Four Parts (% of grade)

25% Odyssey writing forum: 1-2 cases / week.

20% Exercises: 15-20 multi-choice (2 try) / week. Write 2-4 one-line statements: Manual grade. Compare counts; describe ratios (part/whole)

5%: Attendance:

50%: Quizzes (two 12%@) and Final (26%)

Textbook: Schield 2021 and 2021B.

1. Odyssey Writing Forum

Each case is a challenge. No research required.

Odyssey forum is different:

- 1. Everyone is anonymous (no avatar)
- 2. No free-riders. Can't see others till after you post
- 3. Everyone grades everyone else.
- 4. Odyssey computes each student's power.

More info: www.statlit.org/pdf/2014-Schield-ICOTS.pdf

2. Exercises

Multiple choice: field tested (1,000+ students). Immediate feedback on right-wrong. Helpful hints/comments on why it is wrong. Closely aligned to the textbook.

One-line essay:

Compare two counts: difference, ratio, % more. Describe part-whole ratios from tables & graphs Use Percent and Percentage grammar.

3. Textbook: Eight Chapters

- 1. The story behind the statistics
- 2. Comparisons and CARE solutions
- 3. Understanding Measurements
- 4. Describing ratios: Percent & Percentage
- 5. Describing ratios: Rate, Chance and Ratio
- 6. Comparing ratios: Named ratios and Likely
- 7. Interpreting confusing ratios
- 8. Randomness

4. Student Evaluations www.statlit.org/UNM.htm

"I've despised every other math class I've ever taken but this one is actually enjoyable. It also seems more useful than any math course I've ever taken."

"I enjoyed critical thinking and the news stories. Both provide beneficial knowledge I will take with me into my everyday life."

4. Student Evaluations www.statlit.org/UNM.htm

"This course is an answer to my prayers, I am a music major and horrible at math so fulfilling my math requirement has been hard.

This is the first math class I actually liked. I loved the format; the material is about things I can apply to everyday life.

The textbook is fantastic and helped me a lot...

I would recommend this class for anyone."

Conclusion

37

Statistical Literacy involves statistics, writing and critical thinking. This course is very different from a typical mathematics course.

Statistical Literacy must be tailored to the students involved. First year-students are very different from Seniors or Honors students.

Students see value in taking Math1300. To help their students think critically about statistics, other colleges should offer Statistical Literacy.

V1 2022 Schield StatLit Metro State 38

Bibliography

ASA GAISE Guidelines (2005). Copy at

www.amstat.org/docs/default-source/amstat-documents/2005gaisecollege_full.pdf

ASA GAISE Guidelines 2016 Update. Copy at

www.amstat.org/asa/files/pdfs/GAISE/GaiseCollege Full.pdf