Welcome to Today's Webinar!
The Growing Role of Statistics and Data Science in Grades 6-12 Education: Implications for Mathematics Teacher Educators

Presenters:

Stephanie Casey, Eastern Michigan University
Rick Hudson, University of Southern Indiana
Hollylynne Lee, North Carolina State University
Gemma Mojica, North Carolina State University

Hosted by:

Enrique Galindo, AMTE President
Shari Stockero, AMTE Executive Director
Megan Burton, AMTE Past-President

Announcements:

- **Virtual Institute:** *Considerations for Designing, Teaching, and Assessing Methods of Teaching Mathematics Courses* on Thursday, November 9

- **STaR 2024 Cohort Applications** are due November 15

- **AMTE Conference Regular Registration** deadline is November 30

The growing role of statistics and data science in grades 6-12 education:
Implications for mathematics teacher educators

Webinar co-sponsored by:

AMTE Statistics Teacher Education Community Circle

ESTEEM II Project

October 12, 2023

Project PIs

Hollylynne Lee, NC State University

Gemma Mojica, NC State University

Stephanie Casey, Eastern Michigan University

Rick Hudson, University of Southern Indiana

Bill Finzer, Concord Consortium

The Data Revolution is Here

What is needed to *revolutionize* students' access to learning with and from data about their world?

The last decade has seen increased attention to statistics and data science in K-12 education

International, national, and state levels

Sample of International Efforts

[International Data Science in Schools Project (IDSSP)](): a cross-disciplinary project involving an international team of computer scientists and statisticians from the leading professional organizations for both disciplines. The purpose of the project is to promote and support the teaching of Introductory Data Science, particularly in the final years of K-12 schooling.

Curriculum Framework for Introductory Data Science
Released in Sept. 2019.

International Association for Statistics Education
[2021 Satellite Conference]():
Statistics Education *in the Era of Data Science*

These frameworks are endorsed by:

ACEMS

American Statistical Association

BCS The Chartered Institute for IT

International Statistical Institute

New Zealand Statistical Association

Royal Statistical Society

Statistical Society of Australia

Teaching Statistics Trust

What is Happening Across the U.S.?

INCREASED EMPHASIS ON STATISTICS AND DATA

ASA-NCTM Joint Committee Efforts in K-12
Professional Learning for Teaching Statistics
- "Meeting within a Meeting" for K-12 Teachers held with Joint Statistics Meeting, Typical attendance **15-40** teachers
- Virtual since 2020, with **80-100** teachers attending across 2 days

NSF-funded Mobilize project created an *Intro to Data Science* course in high schools since ~2013

Increase in enrollment in high school statistics classes (AP and non-AP)

Growth of AP Statistics Exam Participation

● # of Students Taking AP Statistics Exam   ▲ # of Students Taking AP Calculus AB Exam
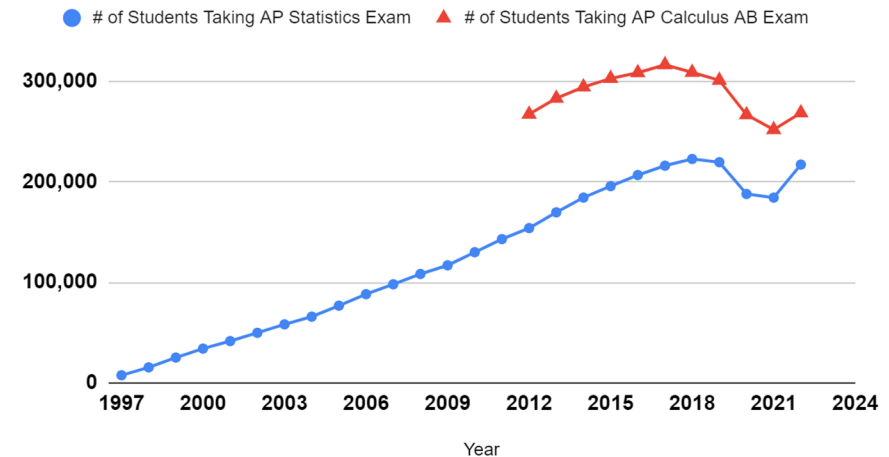
What is Happening Across the U.S.?

Launch Years Initiative: *Bringing more applicable and modern content into HS math curriculum;* 22 states

Data Science Course Framework, released 2021

NCER and IES Technical Working Group, 2021
- *Catalyzing a New Field: Data Science Education in K-12* Report and Recommendations

National Academies of Sciences
- *Foundations for Data Science in K-12*, September 2022. Summary report available

Data Science 4 Everyone, since 2019. Organizational commitments, a resource hub, and webinars

YouCubed K-12 Data Science since 2020; data talks (like number talks), curriculum, PD

Joint Position Statement

[Preparing PreK - 12 Teachers of Statistics](#), September 2022

Students, teachers, administrators, employers, and others increasingly recognize statistics and data science as important scientific fields of study. Mathematics content standards emphasize that the development of statistical thinking begins in the early grades and extends into and beyond high school. **To successfully develop students' statistical thinking, teachers must have deep knowledge and understanding of statistics and the way that students learn statistics. Teachers of secondary-level data science must have knowledge of developing students' statistical and computational thinking in the presence of data.** Consequently, the need is critical for high-quality preservice and in-service preparation and professional development that supports PK–12 teachers of mathematics and statistics, new and experienced, in developing their own statistical proficiency as well as their students' understanding of and skill in working with statistics and data science. **The preparation of PK–12 teachers of statistics and data science must also include explicit attention to equitable use of statistics and data and critical data literacy.**

NATIONAL COUNCIL OF TEACHERS OF MATHEMATICS

ASA AMERICAN STATISTICAL ASSOCIATION Promoting the Practice and Profession of Statistics®

EsTEEM

New Joint Position Statement

Data Science in K-12 Education -- to be released late 2023-early 2024

- Data science is contextual and multidisciplinary
- Data science is an investigative process
- Data science understandings and experiences are for everyone
- Data science educators must develop and practice ethical uses of data

Data Science Efforts at the State Level

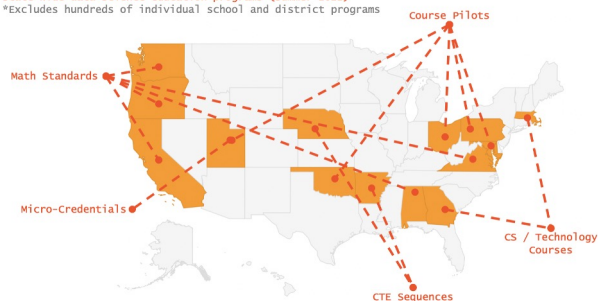*Estimated active K-12 Data Science programs:*

1,600+ schools / districts

2,000+ teachers

180,000+ students

14 statewide programs

State-wide data science education programs (Summer 2022)
*Excludes hundreds of individual school and district programs

Electives

| Algebra 1 | → | Geometry | → | Algebra 2 | | Data Science |

Math Pathways

Algebra 1 → Geometry → Data Science 1 → Data Science 2
Geometry → Algebra 2 → Pre-Calc

Modernized Courses

Algebra 1 → Geometry → Algebra 2 → Data Science 2
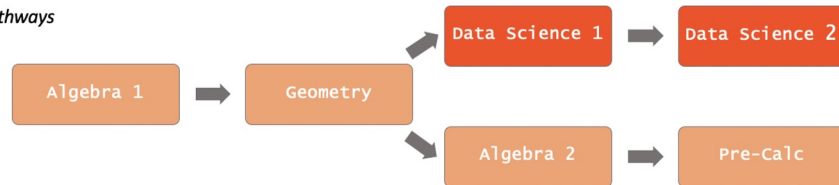
Data and images from Zarek Drozda, *Hearing from Practice: What is Happening in and Out of Schools?* September 14, 2022 at National Academies. Slide deck and paper available at https://www.nationalacademies.org/event/09-13-2022/foundations-of-data-science-for-students-in-grades-k-12-a-workshop-days-1-and-2

Examples of State Initiatives

Virginia
- Updated [Mathematics standards](#) include a course in Data Science in high school, 2023

California
- New [K-12 Mathematics Framework](#) includes strong focus on data science, 2023

Oregon
- Added data reasoning throughout [K-12 mathematics curriculum standards](#)
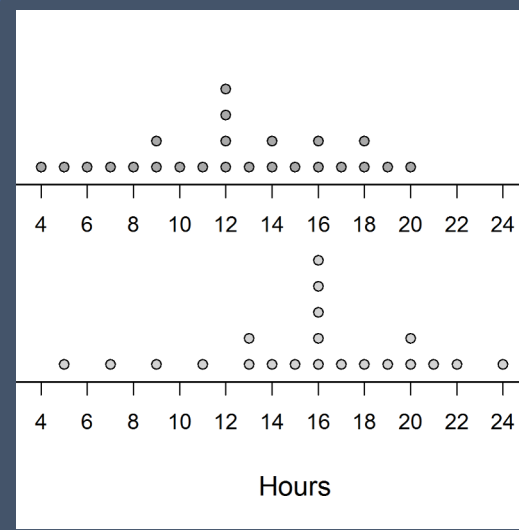- Data Science course in high school as an alternative to Algebra II

Ohio
- [5 different math course options](#) after Algebra 1 & Geometry: 1) *Statistics and Probability*, 2) *Data Science Foundations*, 3) Discrete Math & Computer Sci, 4) Quant Reasoning, or 5) Algebra 2
- Data Science Foundations course taught using *Introduction to Data Science* curriculum

North Carolina
- [Math 4,](#) since 2020 is about 1/3 statistics and probability content
- Introduction to Data Science class, started 2021 as part of Career & Technical Education

What has changed?
Statistics and Data Science

What are the typical experiences students have in learning statistics in school?
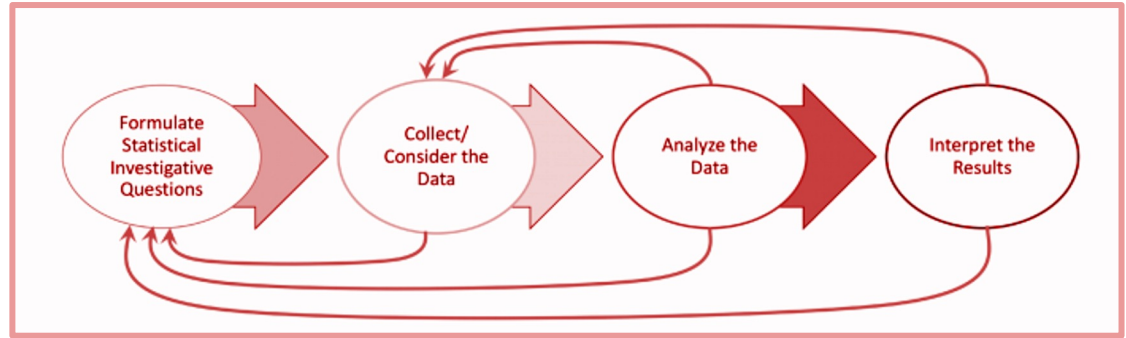
Statistics as a problem solving process

PCAI Model (Friel et al., 2006; Graham, 1987; Lee & Tran, 2015)
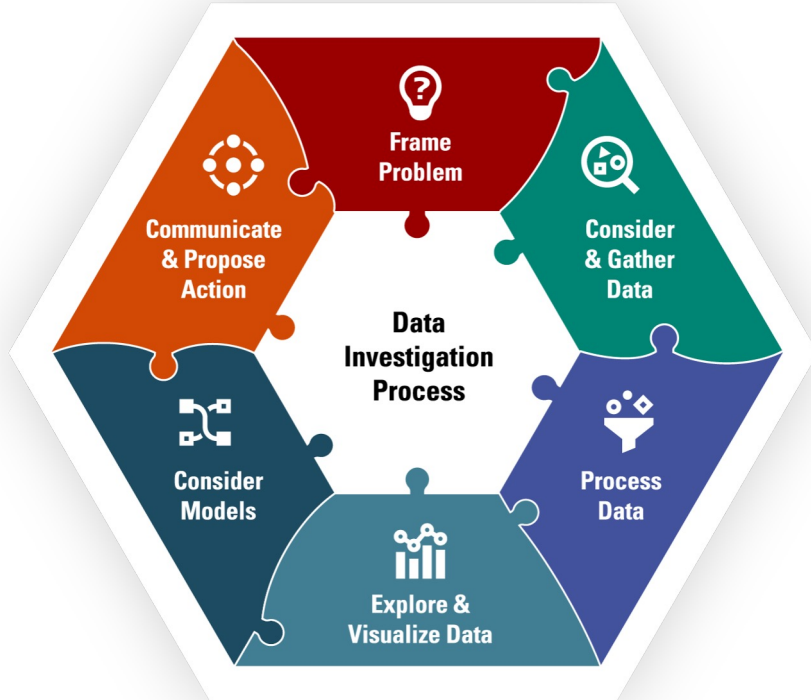
Data Cycle (Gould et al., 2016)

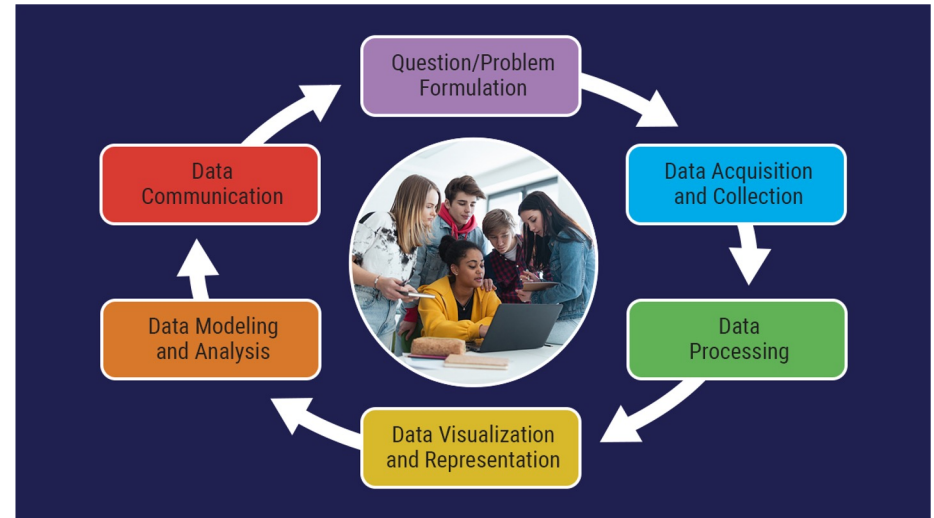Statistical Problem Solving as an Investigation Process (GAISE II, 2020)

Data Science Processes

Data Investigation Process (Lee et al, 2022)

Virginia's Data Science Cycle

How Student Expectations and Experiences Have Changed

Recall the equation of the least squares regression line is

$$\hat{y} = a + bx$$

Where the slope coefficient *b* and intercept coefficient *a* are determined from the sample data, specifically the means and standard deviations for each variable and the correlation coefficient between them:

$$b = r \frac{s_y}{s_x} \qquad a = \bar{y} - b\bar{x}$$

$\bar{x}$ is the mean of the *x* variable, s_x is the standard deviation of the *x* variable.
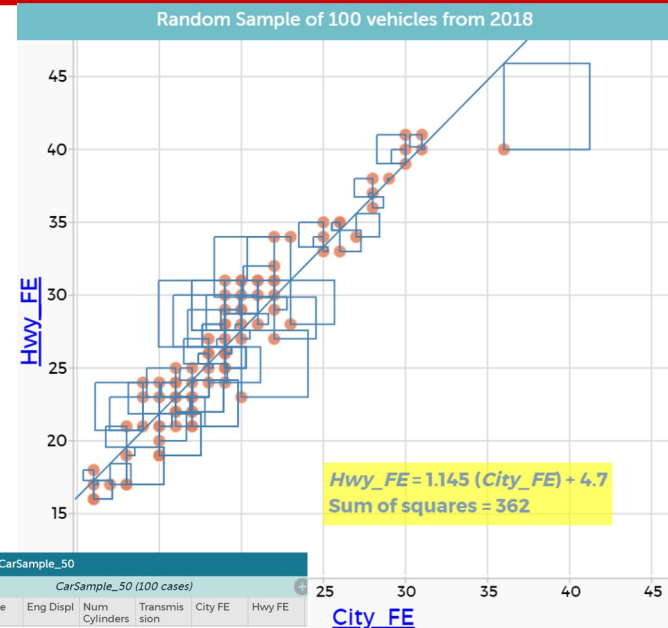
$\bar{y}$ is the mean of the *y* variable, s_y is the standard deviation of the *y* variable.

r is the correlation coefficient

The means and standard deviations of sports cars' weight and fuel efficiency and the correlation between them are reported in the table below:

	Mean	Standard Deviation	Correlation
Weight	2997	357.6	-0.816
MPG	20.867	3.044	

a) Use this information to determine (by hand) the coefficients of the least squares line for predicting a car's miles per gallon rating from its weight. Report the equation of this line.

b) Use the regression line to predict the city MPG rating for the Audi TT, whose weight is 2655 pounds.
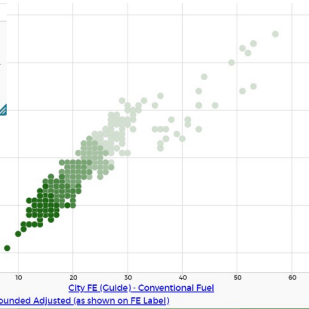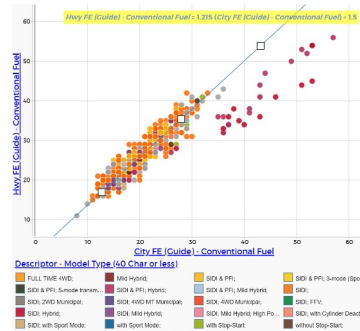
Random Sample of 100 vehicles from 2018

Hwy_FE

City_FE

$Hwy\_FE = 1.145\ (City\_FE) + 4.7$
Sum of squares = 362

CarSample_50

Divisions (34 cases)		CarSample_50 (100 cases)					
index	Division	index	Carline	Eng Displ	Num Cylinders	Transmission	City FE	Hwy FE
1	Porsche	1	Paname...	4.8	8	Auto	16	24
2	HYUND...	2	Macan T...	3.6	6	Auto	17	23
3	LEXUS	3	Cayenn...	3	6	Auto	20	29
4	NISSAN	4	911 Targ...	3.4	6	Auto	19	26
5	KIA MO...	5	Boxster ...	3.4	6	Auto	22	31
6	INFINITI	6	Boxster S	3.4	6	Manual	20	28
7	TOYOTA	1	Sonata ...	2	4	Auto	21	31
8	MAZDA	2	Santa F...	3.3	6	Auto	17	22
9	BMW	3	Santa F...	2.4	4	Auto	19	25
10	SCION	1	ES 350	3.5	6	Auto	21	31
11	Ford	2	GX 460	4.6	8	Auto	15	20
12	Honda	3	RC 350	3.5	6	Auto	19	28

How Student Expectations and Experiences Have Changed

I wonder about the fuel efficiency and CO2 emissions for vehicles in 2024

Where We Began in 2016

Enhancing Statistics Teacher Education Through E-Modules
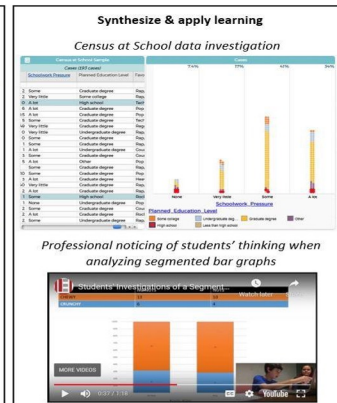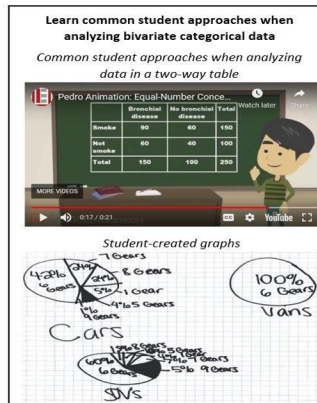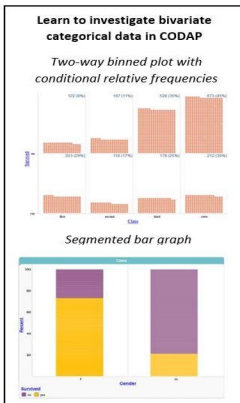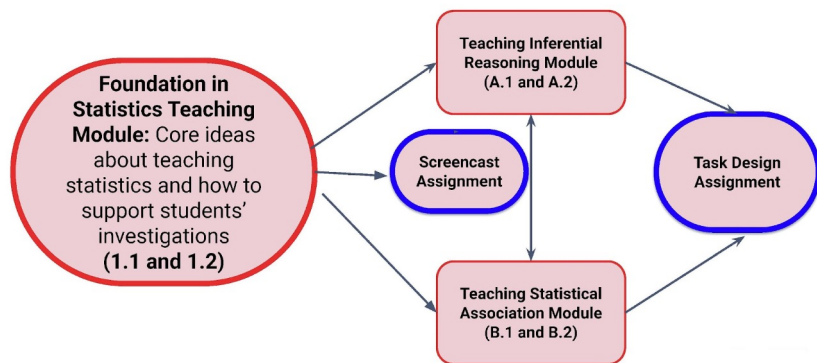
40+ hrs of multimedia material

150+ faculty as ESTEEM users

Materials are FREE and can be imported directly into LMSs

BROAD GOAL FOR ESTEEM II

Transform *undergraduate teacher preparation* to support developing teachers who are knowledgeable and confident in teaching Data Science & Statistics (DS&S)

Primary focus:

Preparation of middle and high school mathematics teachers

Secondary focus:

Preparation of other content teachers where data and statistics is utilized for learning (e.g., science, social studies, computer science, engineering)

In 2022, ESTEEM II launched…

Goal A: Investigate the current systems in undergraduate teacher preparation for teaching data science and statistics [DS&S].

Goal B: Build and sustain a DS&S teacher education networked improvement community.

Goal C: Reach a broad, large, and diverse teacher education audience through developing, curating and disseminating high quality DS&S teacher education curriculum materials.

ESTEEM II Goals

Goal A: Investigate the current systems in undergraduate teacher preparation for teaching data science and statistics [DS&S].

Goal B: Build and sustain a DS&S teacher education networked improvement community.

Goal C: Reach a broad, large, and diverse teacher education audience through developing, curating and disseminating high quality DS&S teacher education curriculum materials.

How prepared are novice mathematics teachers to teach statistics and data science? Preliminary results

Early Career Mathematics Teacher (ECMT) Study

- **Collected Data in Spring-Summer 2023**
- **Participants: 70+ grades 6-12 teachers within their first three years of teaching career**
- **Study included examining:**
  - **Practices for Teaching Statistics**
  - **Content Knowledge (LOCUS)**
  - **Pedagogical Content Knowledge**
  - **Attitudes towards Statistics (I-SOMAS)**
  - **Self-efficacy for Teaching Statistics (SETS)**

Participant Demographics

81 ECMTs shared demographic data
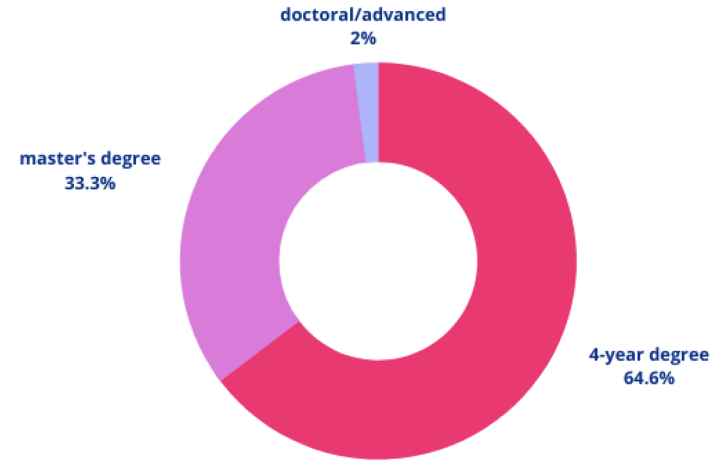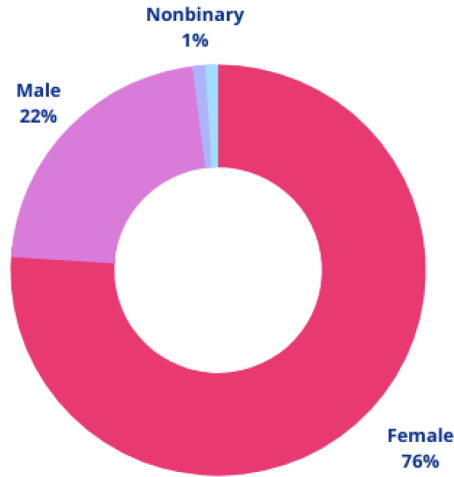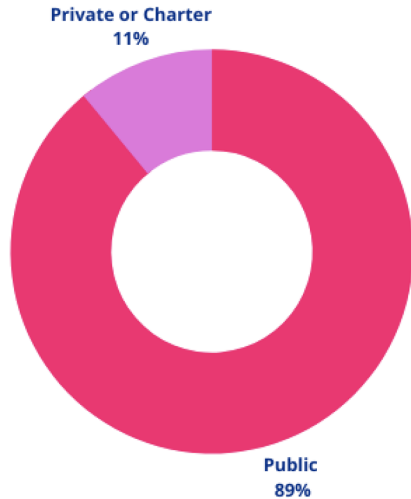
Where do they teach?

Participants represent 22 states in the U.S.
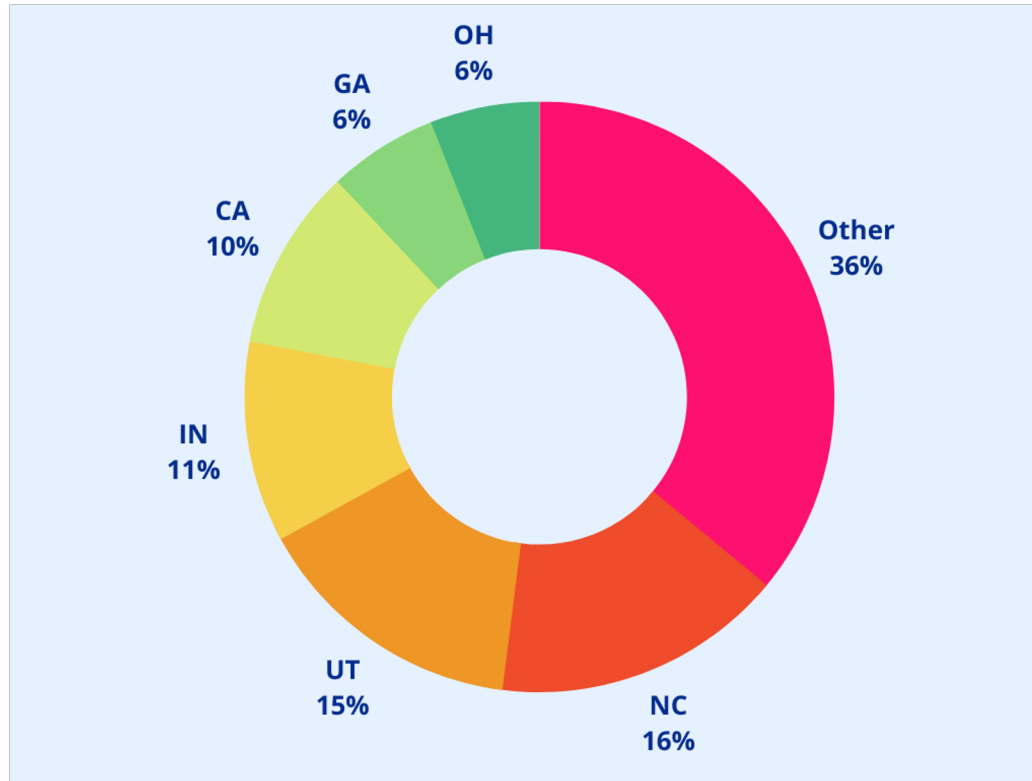
What are their ethnicities and racial identities?

Racial Identity

- White
- Asian
- Black
- Asian/White
- Did not answer
- Black/White
- American Indian/Native American

80%
10%
2.5%
2.5%
2.5%
1.25%
1.25%
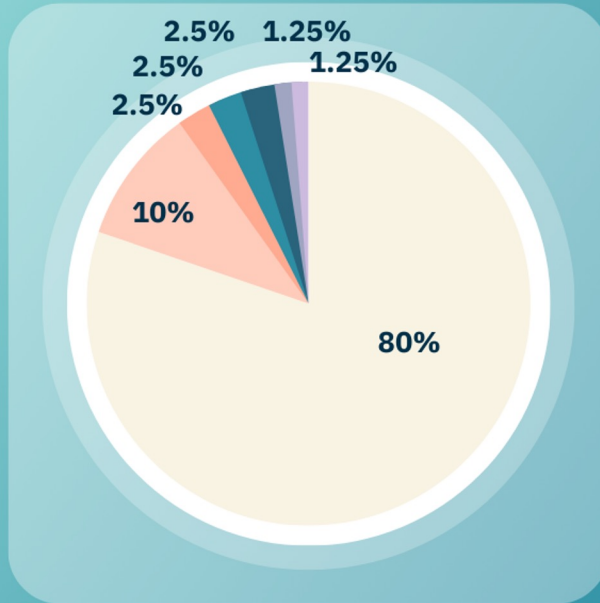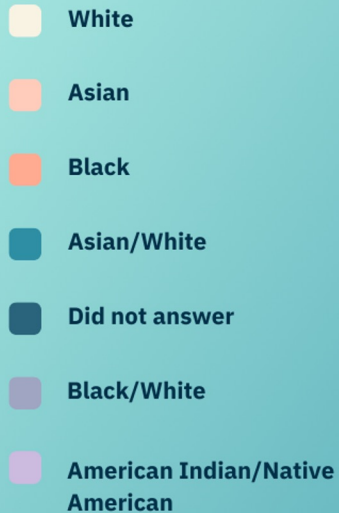
- **10% identify as Hispanic/Latino/Spanish Origin**
- **90% identify as Non-Hispanic/Latino/Spanish Origin**

How many years of teaching experience do they have and what grade levels do they teach?

2-3 yrs 26%
0-1 yr 43%
1-2 yrs 31%
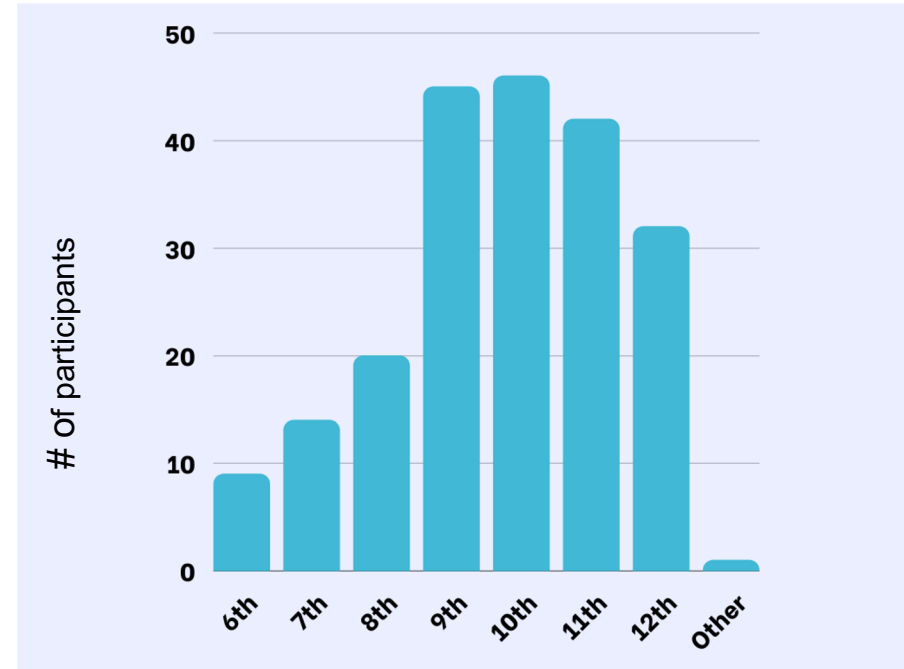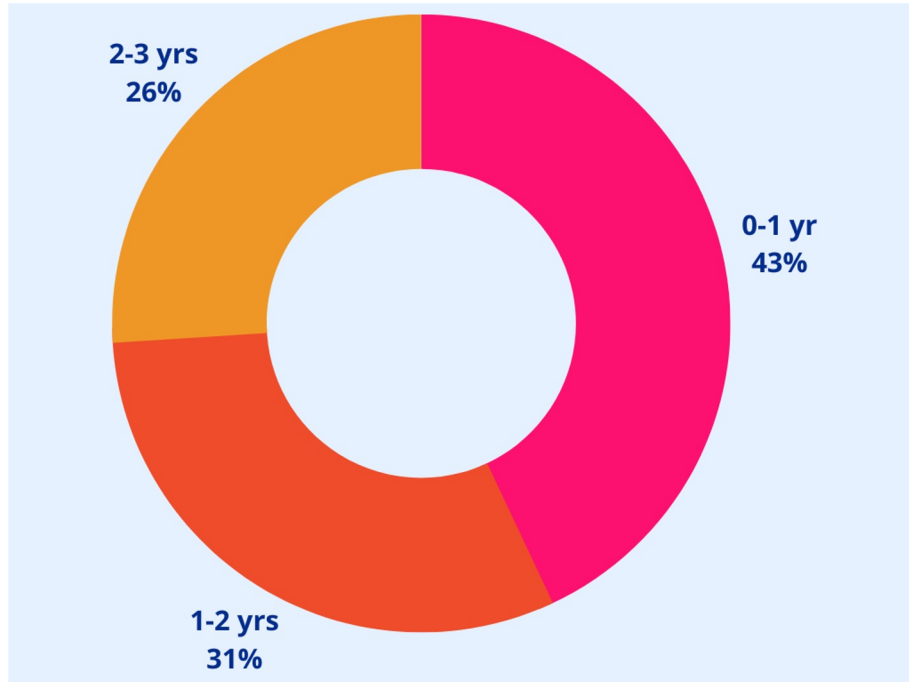
What type of teacher preparation program did they complete?
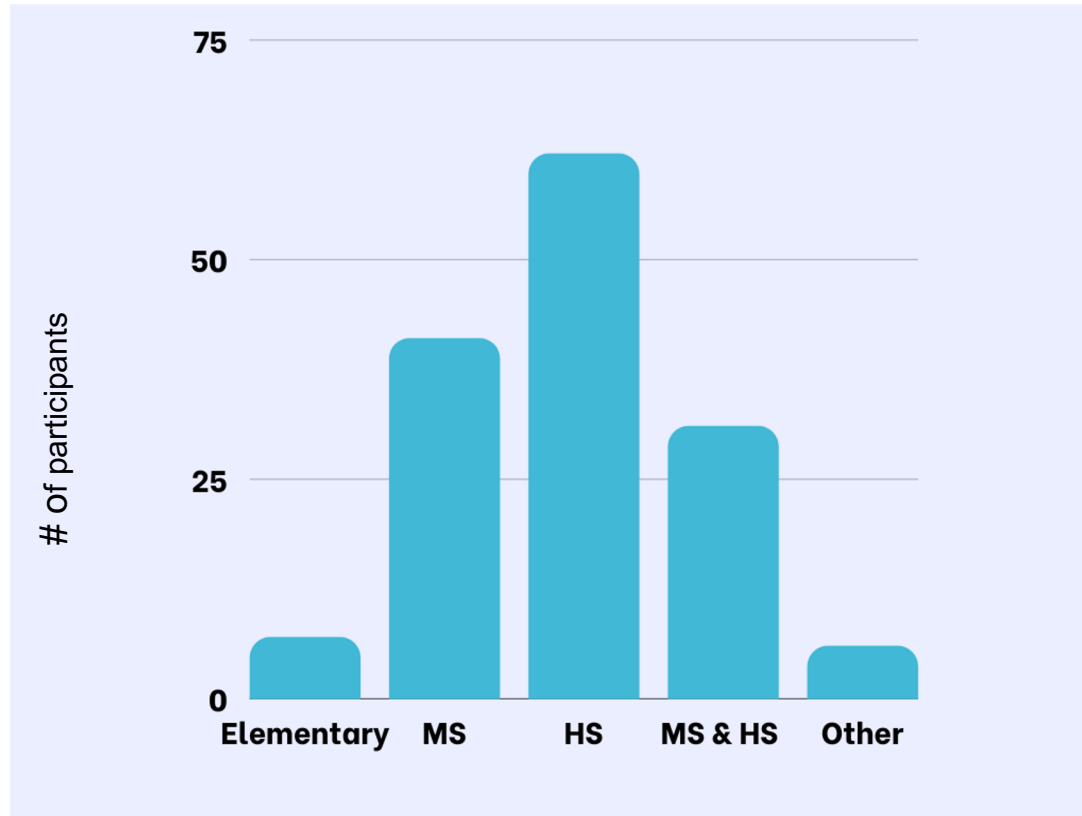
Practices

What type of technology do ECMTs use?

Technology Tool	ECMT uses for deomonstration	Students use
Spreadsheets	38%	27%
Web apps (e.g., Rossman & Chance)	15%	11%
Data visualization/analysis tools (e.g., CODAP)	15%	10%
Statistical software (e.g., StatCrunch, JMP)	6%	1%
Open source programming tools (e.g., R, Python)	6%	1%

81% indicated their students use technology tools, **other than the GC**, to explore, visualize or analyze data

EsTEEM

Practices

Do ECMTs provide opportunities for students to engage with "large data"?
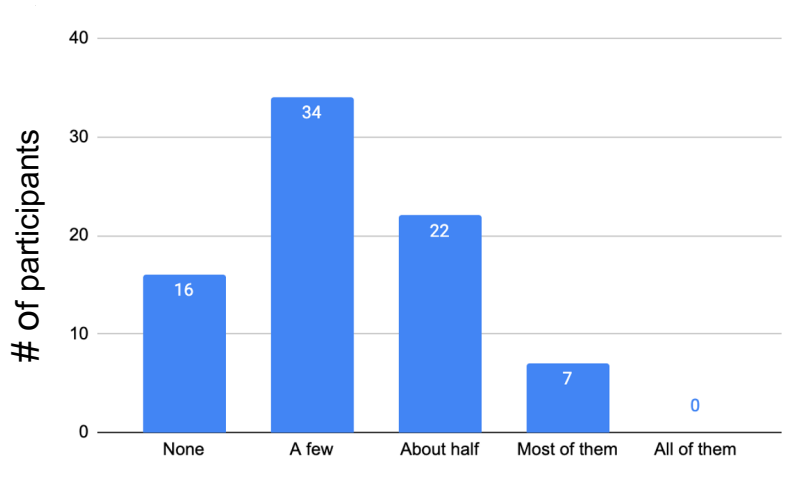
Data contain 26-100 cases
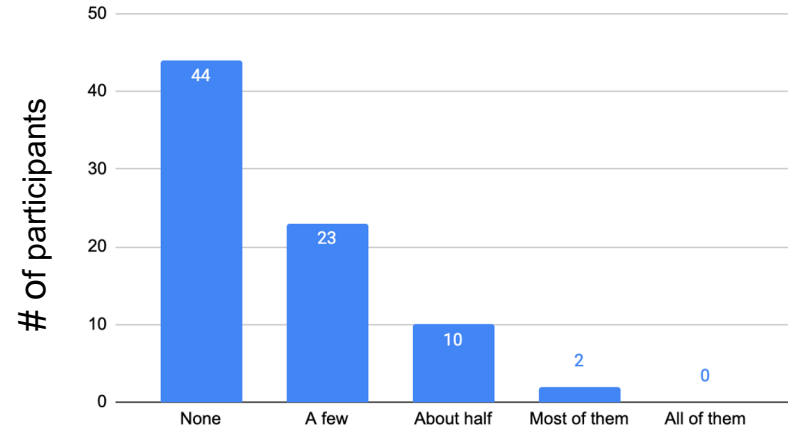
Data contain 101-1000 cases

Practices

Do ECMTs provide opportunities for students to engage with multivariate data?
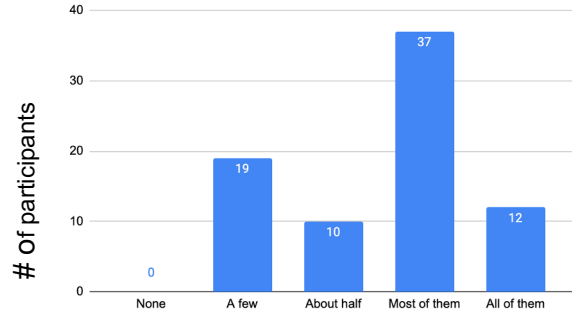
Data contain 1 or 2 attributes
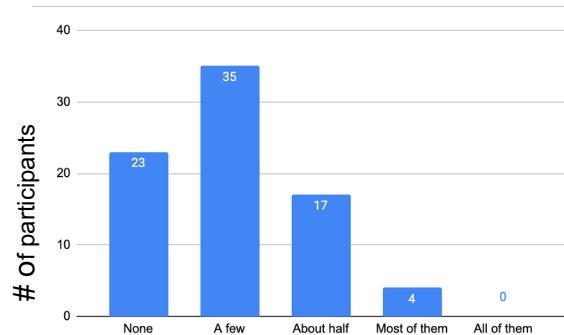
Data contain 3-10 attributes
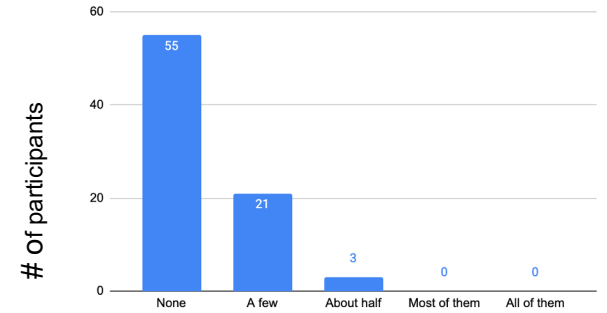
Data contain more than 10 attributes

Content Knowledge (*n* = 72)
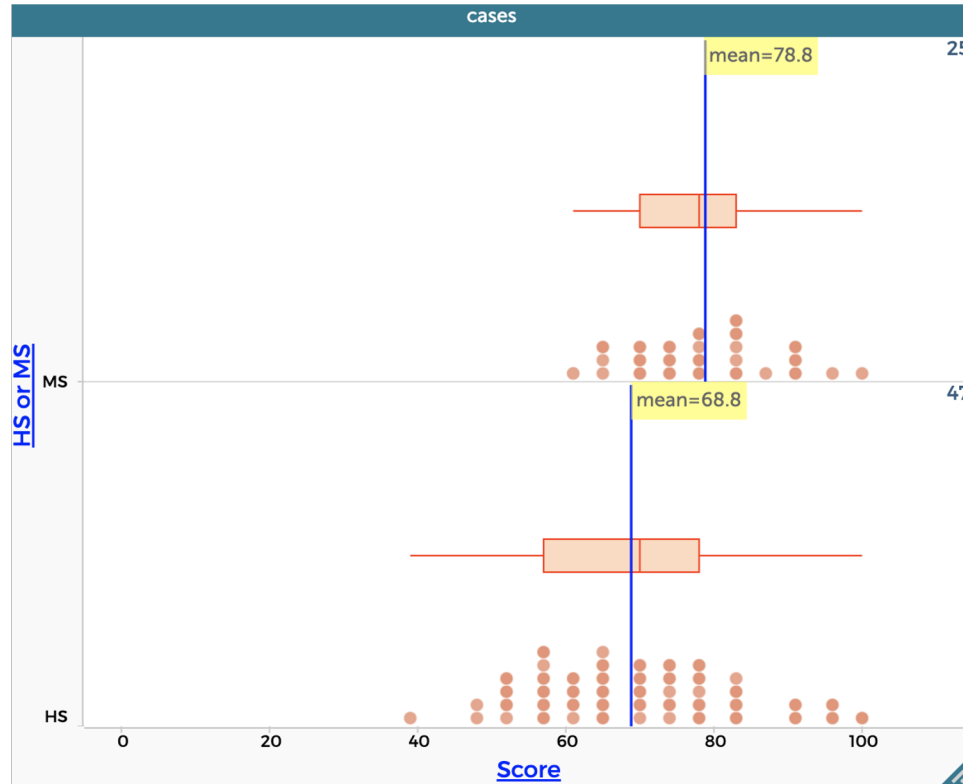
Pedagogical Content Knowledge (*n* = 79)

ECMTs' analyzed a side-by-side bar graph created by a student to analyze the association of two categorical variables with two weaknesses

- No label on vertical axis
- Comparison of raw frequencies rather than conditional relative frequencies

Weaknesses Identified

2 weaknesses
5.1%

0 weaknesses
41.8%

1 weakness
53.2%

Task from Casey, Ross, & Strayer (2022)

ESTEEM
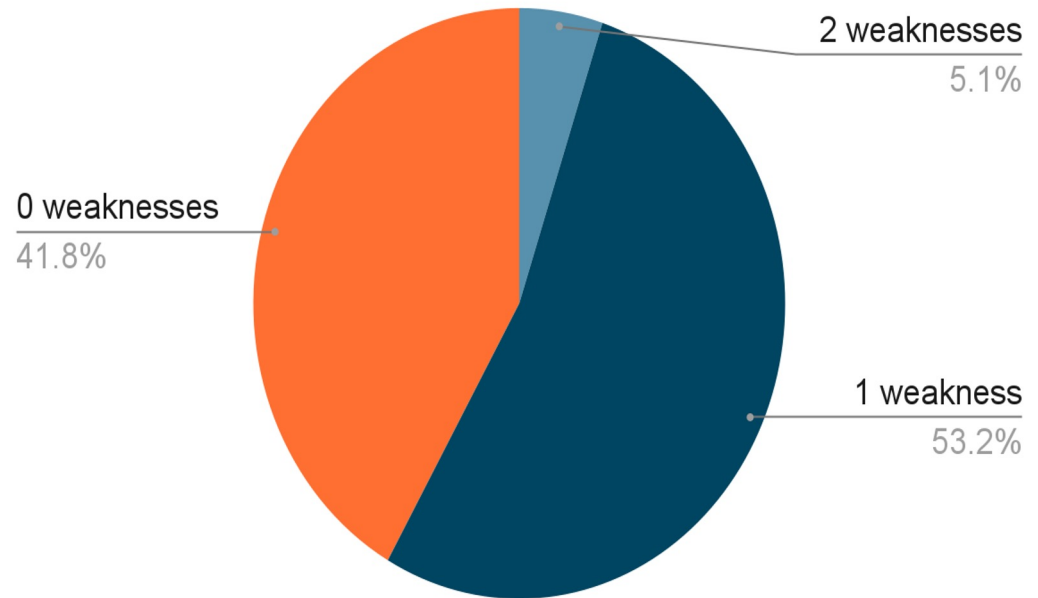
Attitudes (*n* = 78)

I am confident in my ability to teach statistics.

73.1% agreed, 18.0% disagreed.

It is important to me to be a good statistics teacher.

79.5% agreed, 9.0% disagreed.

Teaching statistics is easier than teaching other courses.

16.7% agreed, **61.5% disagreed.**

I would rather teach statistics than other courses.

37.2% agreed, **50.0% disagreed.**

Teaching statistics is part of who I am.

20.5% agreed, **48.7% disagreed.**

Self-Efficacy for Teaching Stats (n = 77)

More Confident

- Using summary values (e.g., mean) to compare groups
- Identifying trends in scatterplots
- Recognizing if a sample is representative
- Creating graphical visualizations

Less Confident

- **Building and using simulations**
- **Interpreting two-way frequency tables**
- Comparing two treatments from a randomized experiment using numerical and graphical summaries
- Assessing the fit of a model using residuals

Future Research Plans

You can help!

Study of Mathematics Teacher Education Programs

- Currently collecting data about the status of mathematics teacher preparation programs and how they are preparing teachers for data science and statistics (instep

Second Study of ECMTs

- Follow-up study involving classroom visits and data from practice

What in the current
Data Science & Statistics Secondary Teacher Preparation System
produces the current outcomes re: the
teaching & learning of
data science & statistics?

Let's Discuss

Given a standard that a secondary math teacher in 2023 would be expected to teach, *where* in your teacher preparation program do they (or could they) have an opportunity to develop *content*, *technological*, and *pedagogical* knowledge for teaching this standard?

bit.ly/teachsds

ESTEEM II Goals

Goal A: Investigate the current systems in undergraduate teacher preparation for teaching data science and statistics [DS&S].

Goal B: Build and sustain a DS&S teacher education networked improvement community.

Goal C: Reach a broad, large, and diverse teacher education audience through developing, curating and disseminating high quality DS&S teacher education curriculum materials.

Mathematics Teacher Educator Resources
for
DS&S Teacher Preparation

Teacher Education Curriculum Materials

EsTEEM

InSTEP

MODULE(S²)

Mathematics Of Doing, Understanding, Learning
and Educating for Secondary Schools

Statistical Investigations of Systemic Racism

https://pages.charlotte.edu/datamodulesandsystemicracism/

ESTEEM

The ESTEEM materials are designed to be used with **preservice middle and secondary mathematics teachers** and are delivered in learning management system formats (Canvas, Blackboard, Moodle, & Common Course Cartridge). There are 3 modules available: Foundation, Inferential Reasoning, and Association.

1.2.i Supporting Statistical Discourse with the Roller Coaster Task

While selecting a statistically rich task that ties together the learning goal, data, context, and investigative cycle is foundational in providing students opportunities to develop more sophisticated statistical thinking, it is as important that teachers consider the implementation of the task and how that implementation might promote reasoning that builds on productive habits of mind. Teachers can support students in developing statistical thinking by encouraging them to communicate their own ideas about engaging with data and consider the thinking of others through discourse.

Part 1. Learn about the Five Practices model for productive classroom discourse

Smith and Stein (2011) developed a model for supporting classroom discourse about students' work on tasks which involves the following Five Practices: *anticipating* students' responses to a task; *monitoring* students' responses to a task; *selecting* specific students to present mathematical ideas; *sequencing* students' responses that will be publicly displayed; and *connecting* between student responses and to key ideas. To learn more, read this three page paper.

Part 2. Watch a Classroom Statistical Investigation

Watch the following video, where a teacher launches a statistical investigation about roller coasters in a seventh grade classroom and students use CODAP for the first time. The sixth and seventh grade students in the video were doing an investigation similar to the one you did with older roller coasters in assignment 1.1.g. The videotaped class session was at the beginning of their school year; they had not yet engaged in a formal statistics unit. As you watch the video, note how the teacher implements the 5 Practices model as she monitors student work, selects and sequences several students' findings to discuss, and leads a whole class discussion connecting students' statistical ideas.

Read the transcript.

go.ncsu.edu/esteem

MODULE(S2)

- Develop deep understanding of both statistics and teaching statistics.

- Focus on study design, exploratory data analysis, inference, and quantitative and categorical association.

- Attention is given to developing an understanding of equitable teaching practices and equity issues in education.

Which 3rd variable (coloring) best helps predict which schools get less than their fair share?
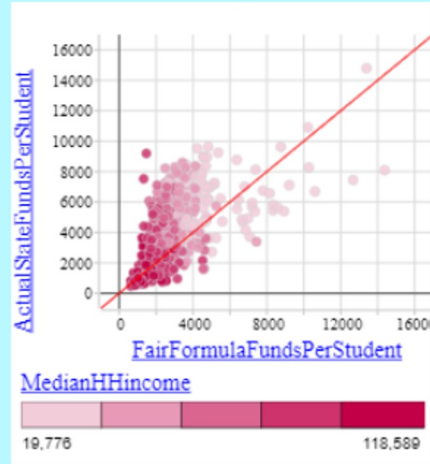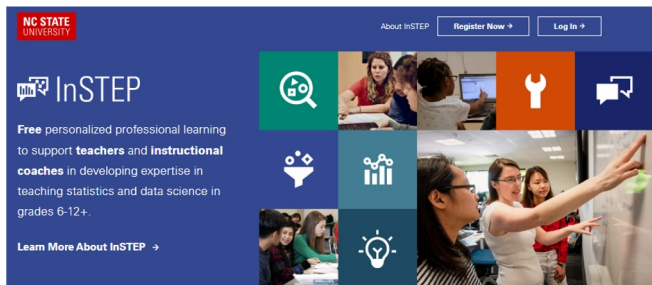
MedianHHincome: Median Household Income in the district
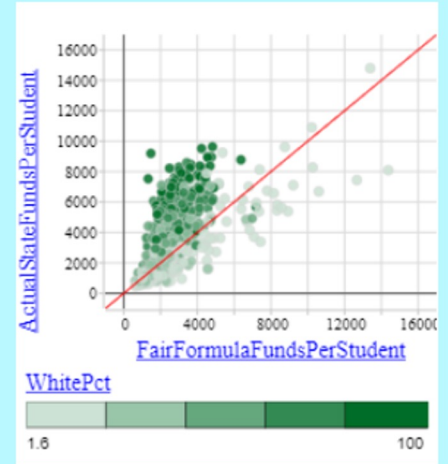
WhitePct: Percentage of White-identifying students in the district

www.modules2.com

InSTEP with Data
Personalized Learning for Teaching Statistics and Data Science

instepwithdata.org

FREE for All!

Teaching Statistics Through Investigations of Issues of Systemic Racism

Anthony Fernandes, Travis Weiland, Ksenija Simic-Muller

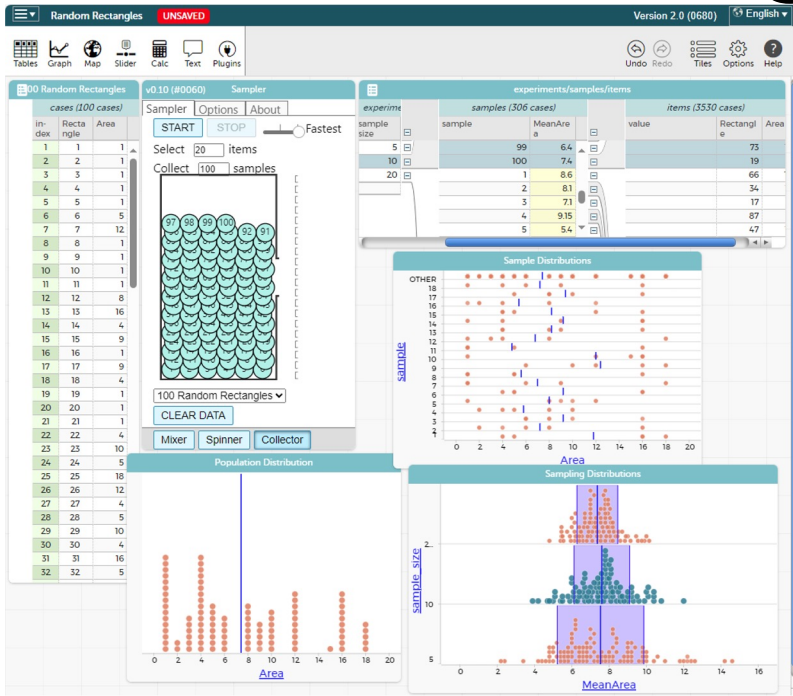Two modules - Traffic stops and School discipline. Materials include,
- Detailed instructor guide with a deep dive into the context
- Slide decks that can be adapted to your class
- Resources for understanding race and racism, and talking about these issues in class

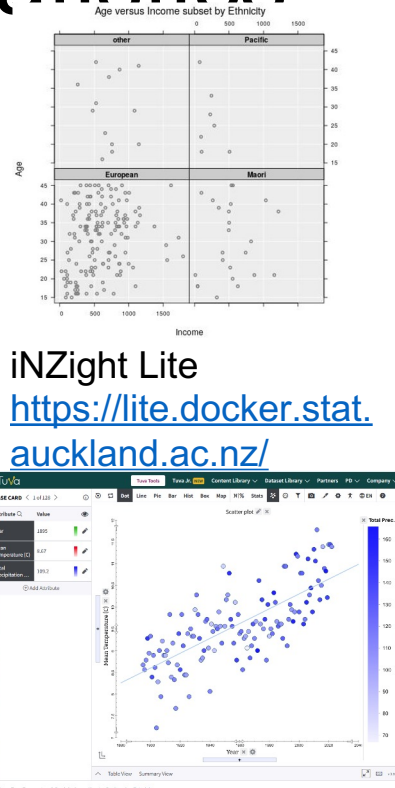Website to learn about the project and request the modules:
https://pages.charlotte.edu/datamodulesandsystemicracism/

Technology Tools for Teaching Statistics

CODAP

http://codap.concord.org

iNZight Lite

https://lite.docker.stat.auckland.ac.nz/

Tuva

https://tuvalabs.com/

- Excel/Google Sheets
- Tableau Public
- NZ Grapher
- TinkerPlots 3.0
- Plotly
- The Art of Stat Web Apps
- Rossman/Chance Applets
- Desmos Stats Activities
- StatKey Applets
- Mathigon Polypad
- Shodor Interactives
- NumWorks Online Calculator

ESTEEM II Goals

Goal A: Investigate the current systems in undergraduate teacher preparation for teaching data science and statistics [DS&S].
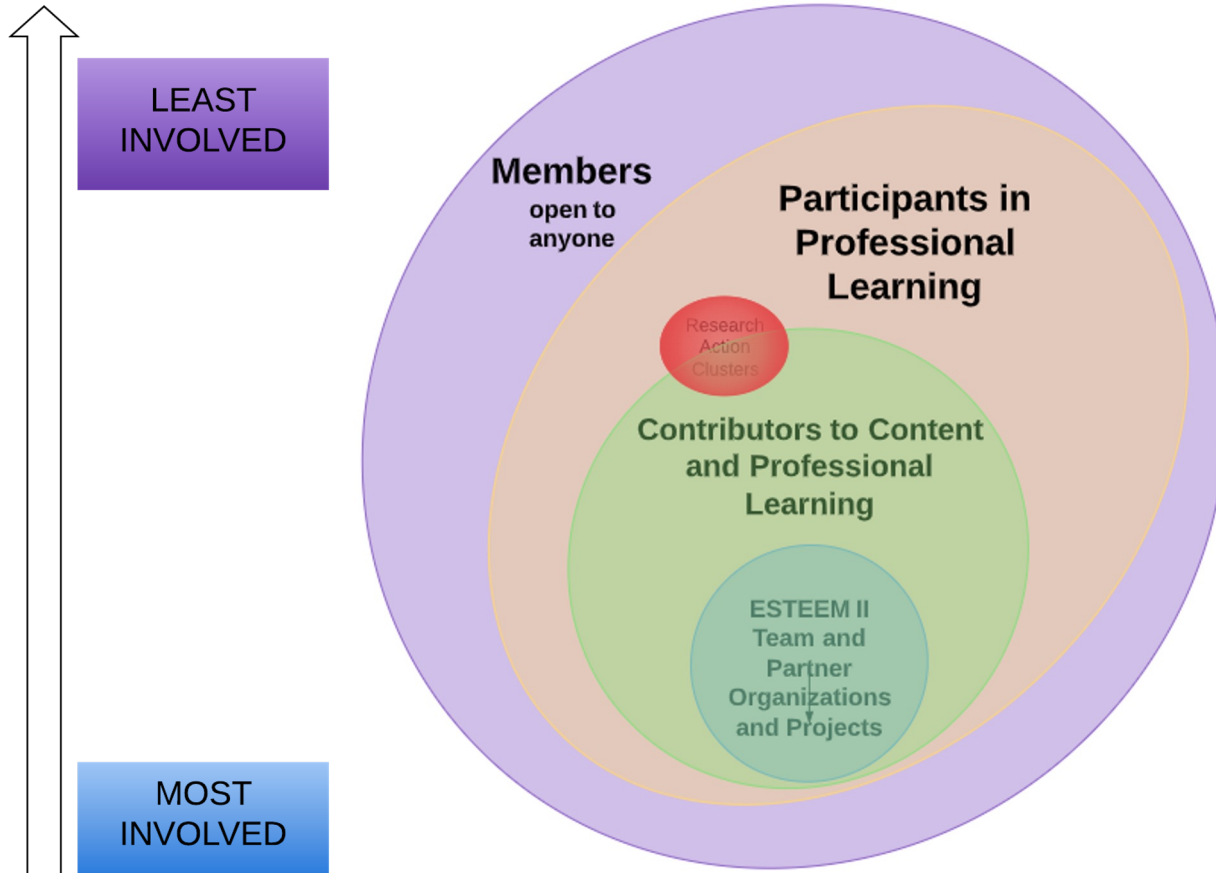
Goal B: Build and sustain a DS&S teacher education networked improvement community.

Goal C: Reach a broad, large, and diverse teacher education audience through developing, curating and disseminating high quality DS&S teacher education curriculum materials.

DS&S Teacher Education
Networked Improvement Community

Communities

Networked Improvement Community:
     Contact Emily Thrasher (epthrash@ncsu.edu) to join

Statistics and Data Science Teacher Education
 Community Circle:
     Contact Travis Weiland (tweiland@central.uh.edu) to join