NATIONAL ACADEMIES *Sciences Engineering Medicine*

Foundations of Data Science for Students in Grades K–12 A Workshop

September 13-14, 2022

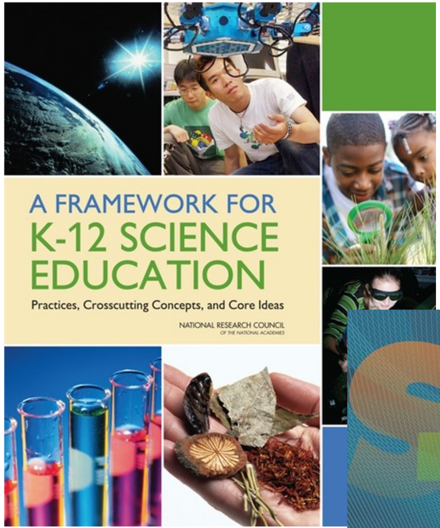
SEPTEMBER 2022

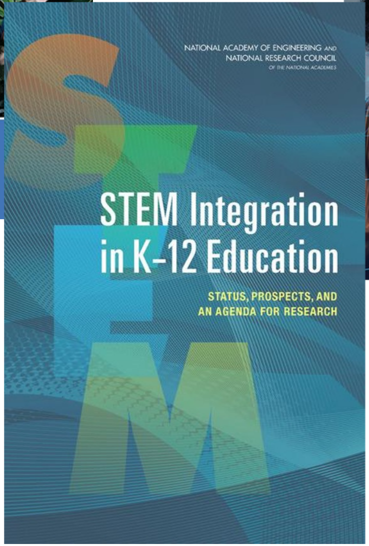
SEPTEMBER 2022

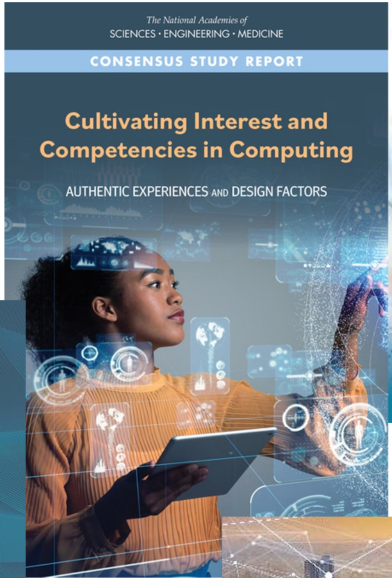
Welcome from the National Academies

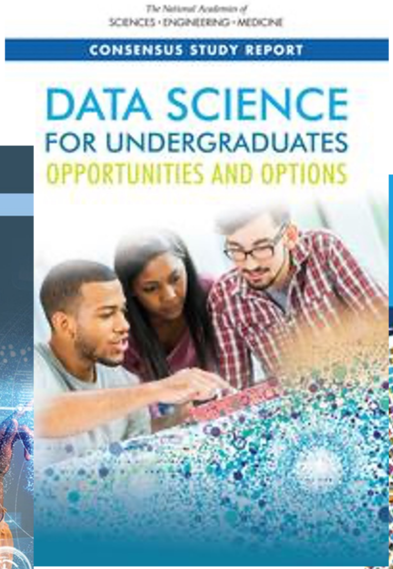
01

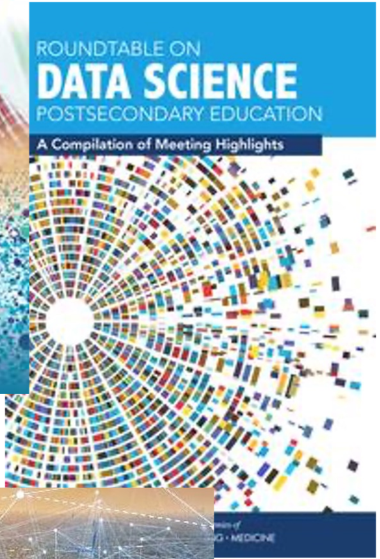
How Did We Get Here?











Norms for Participation

- Embrace diversity
  - Differences in opinion are welcomed
  - Be open, listen respectfully
- Strive to promote an inclusive environment where everyone feels welcomed, valued, respected, and supported
  - Be constructive in your comments
  - Remember, bullying behavior will not be tolerated

Policy on Preventing Discrimination, Harassment, and Bullying

- Maintain an environment free of harassment and intimidation
- Shared responsibility not to commit harassing or discriminatory acts, not to tolerate or ignore those of others, and to avoid knowingly placing others in situations where they may be harassed
- Compliance required in all settings at the National Academies in which work is performed
- Report any incident of harassment, discrimination, or bullying to NASEM staff

More information about our policies and procedures can be found at nas.edu/about/volunteers

Thank you!

Sponsor: Valhalla Foundation

Committee Members		Staff
Nicholas Horton (co-chair)	Michelle Wilkerson (co-chair)	Heidi Schweingruber
Tammy Clegg	Zarek Drozda	Amy Stephens
Tim Erickson	Hollylynne Lee	Janet Gao
Camillia Matuk	Leigh Peake	Lauren Ryan

Opening Remarks and Workshop Framing

02

What Were We Tasked to Do?

To bring visibility to the need for data science education at the K-12 level:

- **Goals and Outcomes**
  - What outcomes matter the most for learners in data science?
  - What are the competencies that make up data fluency?
  - How can these data fluency competencies be measured?
- **Tools and Instruction**
  - What kinds of learning experiences might help bring about these data fluency competencies?
  - What tools & data sets are needed to support young learners in acquiring data understanding & skills?
  - How can K-12 data science education be designed to specifically reach students who have been traditionally marginalized and/or underrepresented in STEM?

What Were We Tasked to Do? (cont.)

To bring visibility to the need for data science education at the K-12 level:

- **Integrating Data Science into the K-12 System**
  - How can learning with data be meaningfully integrated with K-12 education, in both STEM & non-STEM classes?
  - How well prepared is the current teacher workforce for teaching data science related content in K-12 settings? What strategies can be used to enhance teachers' expertise related to data science?
- **Evidence Base and Future Directions**
  - What bodies of research can be leveraged to gain insight on the development of data fluency & how best to support students?
  - What are the critical gaps in the current knowledge base?
  - What are the highest priority next steps for research & practice?

Commissioned Papers

- A Secret Agent: K-12 Data Science Learning Through the Lens of Agency
  - *What do the diversity of fields engaged in this work know about student learning of data science?*
- Critical Data Literacy: Creating a More Just World Through Data
  - *What are definitions, examples, and outcomes of critical data literacy activities with youth?*
- Previewing the National Landscape of K-12 Data Science Implementation
  - *How is K-12 Data Science Education being implemented nationally, as evidenced by policy, frameworks, case studies, and practitioner reports?*
- Tools to Support Data Analysis and Data Science in K-12 Education
  - *What are the strengths, needs, and current uses of technological tools to support data investigations at the K-12 level?*

Overview of the Agenda (Day 1)

Visioning and Outcomes

- A Vision for Data Science Education
- Where and How is Data Science Happening?
- Working Lunch: What are the Outcomes We Want?
- Invited Commentary & Report Out on Outcomes Working Lunch
- How are Tools and Resources Supporting Data Science Learning Experiences?
- Townhall

Overview of the Agenda (Day 2)

State of the Field

- Hearing from Practice: What is Happening in and Out of Schools?
- How is Data Science Integrated in Content Areas?
- What is the State of Educator Preparation in Data Science?
- Townhall
- Funder Reflection
- Final Reflections from the Planning Committee

Opportunities and Challenges

K12 Data Science is in its infancy: This is a chance to think deeply and creatively about how it looks and how it might develop

- Diversity and Inclusion
  - We must act now to ensure the disparities that have become all too common in STEM are not replicated here.
- Tools and Technologies
  - Pedagogical and professional tools have become more powerful, and costs have plummeted. What role should they play (or not play) in teaching, learning, and lessening disparities?
- Pathways and Developmental Progressions
  - What educational pathways are most promising? How can this be informed by what we know about student learning and development?
- Teacher Preparation
  - How can we equip future and current teachers with a foundation in data acumen that can allow them to confidently teach about data?

Some Guiding Questions

- What does learning data science look like? Where is it happening (in and out of school)?
- How do we design a data science education that:
  - Prepares all students to access, learn, and practice data science
  - Invites and supports diversity in who participates in data science, including along the dimensions of race/ethnicity, gender, disability status
  - Encourages more just and ethical approaches to data science as a field
- Where *should* data science education be located?
- What's special about data science?
- What's special about doing data science in K-12?
- How can the somewhat siloed communities involved in this space come together to collaborate, learn, and share findings?

Some Guiding Questions

Our goal is not to try and answer all of these questions, but to explore options that might shed light, share resources, and connect communities.

We encourage you to share your comments and wisdom over the next two days.

How to Engage

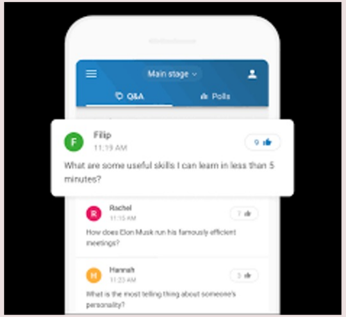
AUDIO

Ensure you are muted unless speaking.

CHAT

We encourage active engagement in the chat.

SLIDO



Interactive polls. Q&A questions.

Panel Discussion
A Vision for High Quality Data Science Education

03

The panelists will explore how we can be sure that we are successfully educating students and consider how those ideas can be applied to the emerging K-12 data science and data literacy efforts.

SESSION GOALS

Panelists

Moderator: Michelle Hoda Wilkerson, University of California, Berkeley

Rob Gould	University of California, Los Angeles
Josh Recio	Dana Center
Tricia Shelton	National Science Teaching Association
Alfred Spector (virtual)	Massachusetts Institute of Technology
Trena Wilkerson	Baylor University; National Council Teachers of Mathematics

Panel Discussion
Where and How is Data Science Happening?

04

Explore the research on the settings and contexts of K-12 data science education with an emphasis on what data science looks like in these contexts and connections relevant to K-12 learners' lives.

SESSION GOALS

Panelists

Moderator: Tammy Clegg, University of Maryland

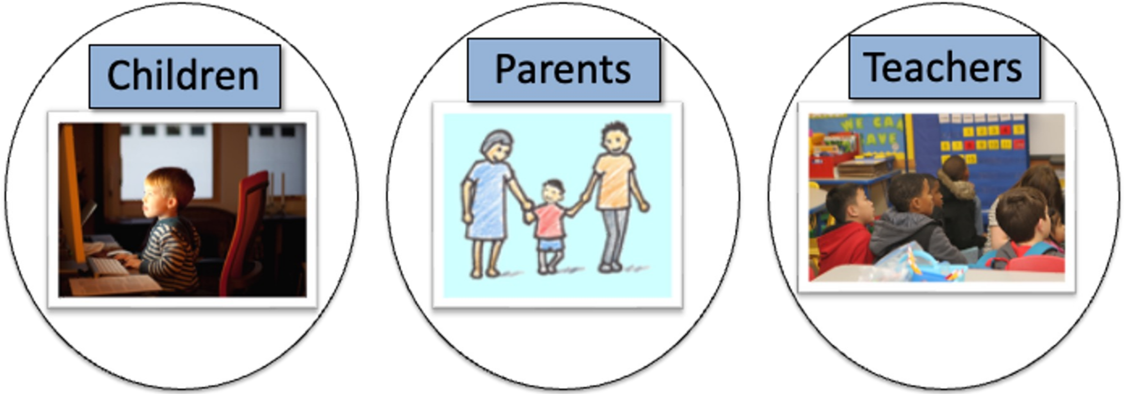
Marshini Chetty (virtual)	University of Chicago
Kayla DesPortes	New York University
Rafi Santo	Telos Learning
Stephen Uzzo	New York Hall of Science

Marshini Chetty

https://airlab.cs.uchicago.edu

I want
The Internet
to be more
Trustworthy and
Inclusive

How do I frame data science in my work?



Helping K-5 Children Learn About Online Privacy and Security (spe4k.umd.edu)

- Data access
- Data collection
- Data management
- Questions of who, why, when?

Where is data science learning happening in my work?

- Children rarely learning about how their data is being used
- Teachers do not have training or time to teach these concepts
- Parents defer to teachers on these topics
- K-5 Children lack **critical data literacy**
    - not just about getting insights from data
    - **but** what are the implications of getting that data
    - e.g. who is using the data from cameras in my neighborhood the South Side of Chicago?
    - how does that affect me as a student?
    - or my community

Why K-5?

Start early with foundations when things are easier

Data Literacy Through the Arts [DLTA]

Exploring the **co-design of data art units** with middle school art and math teachers to examine how various arts disciplines can couple with data literacy to be **mutually supportive**.

Camillia Matuk
PI NYU

Anna Amato
NYU PhD Student

Kayla DesPortes
Co-PI NYU
kd90@nyu.edu

Megan Silander
Co-PI EDC

Marian Tes
NYU PhD Student

Contemporary Dance

Digital Comics

Ralph Vacca
Co-PI Fordham

Peter Woods
MIT Research Sc.

NATIONAL ACADEMIES *Sciences Engineering Medicine*

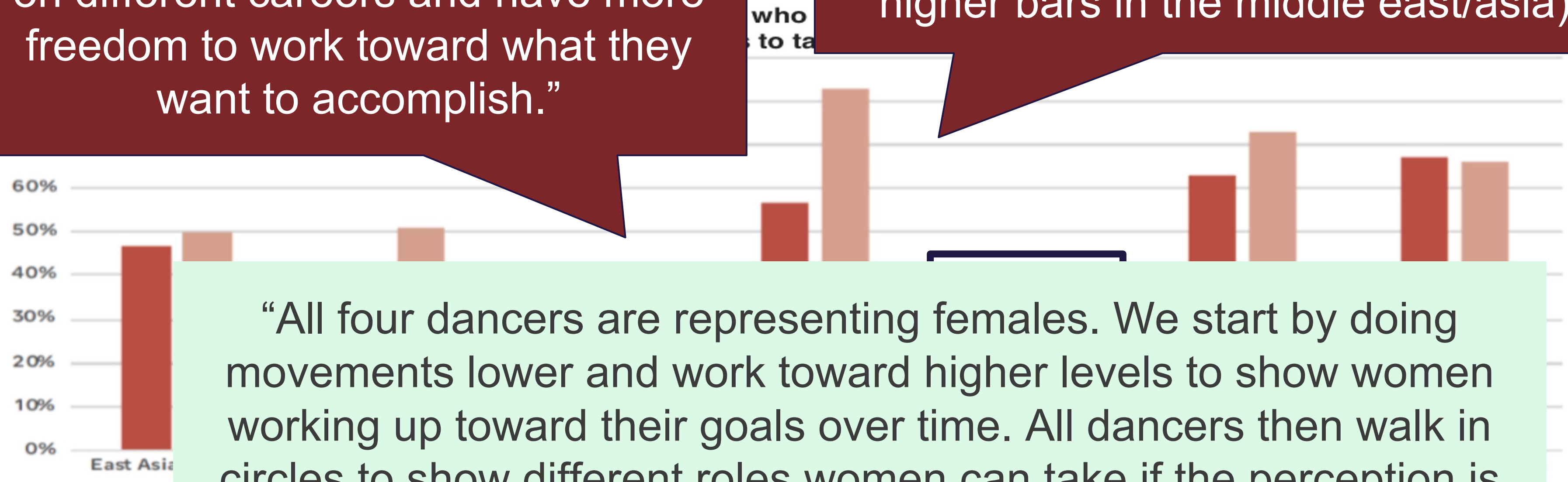
NSF NYU FORDHAM EDC Education Development Center

Graph Analysis for Choreography

Women's Rights Group

"In those regions, women now take on different careers and have more freedom to work toward what they want to accomplish."

"While some regions have not changed their perception of what women can do, there are regions who have changed their perception. (The lower bars in North America/Latin America vs the higher bars in the middle east/asia)"



"All four dancers are representing females. We start by doing movements lower and work toward higher levels to show women working up toward their goals over time. All dancers then walk in circles to show different roles women can take if the perception is changed like it has been in the North America and Latin America."

Gender Surve
Facebook 20

Emerging Practices for Data Dance Inquiry

Embodying numbers and values in dance movements

Exploring the shape of the graph or the numerical values

Embodying the context & implications of the data

Exploring characteristics or attributes of the variables and how they are situated in the world

Engaging in ensemble representations

Making meaning through collaborative movements

Iteration & Audience Interpretation

Critiquing & iterating on dance through perspective taking as an audience member

Transdisciplinary Work

Combining data science and dance changed how learners worked across both data science and their dance practice

Digital Comics About Friendship using Pixton

Socio-Emotional Learning (SEL) in Data Comics

Verbal Communication: Text, Speech Bubbles, & Captions

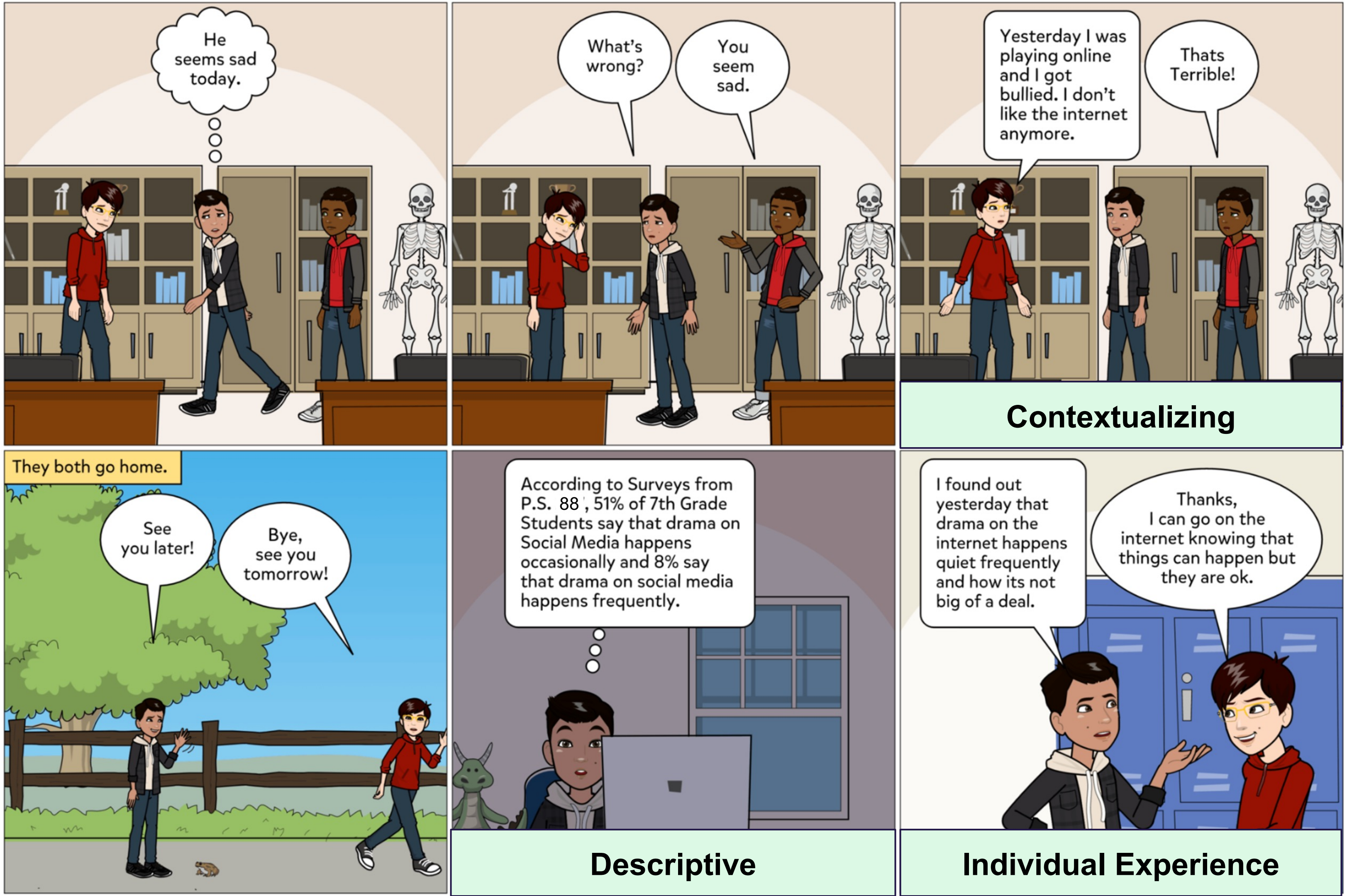
Non- Verbal Communication: Facial Expressions & Body Positions

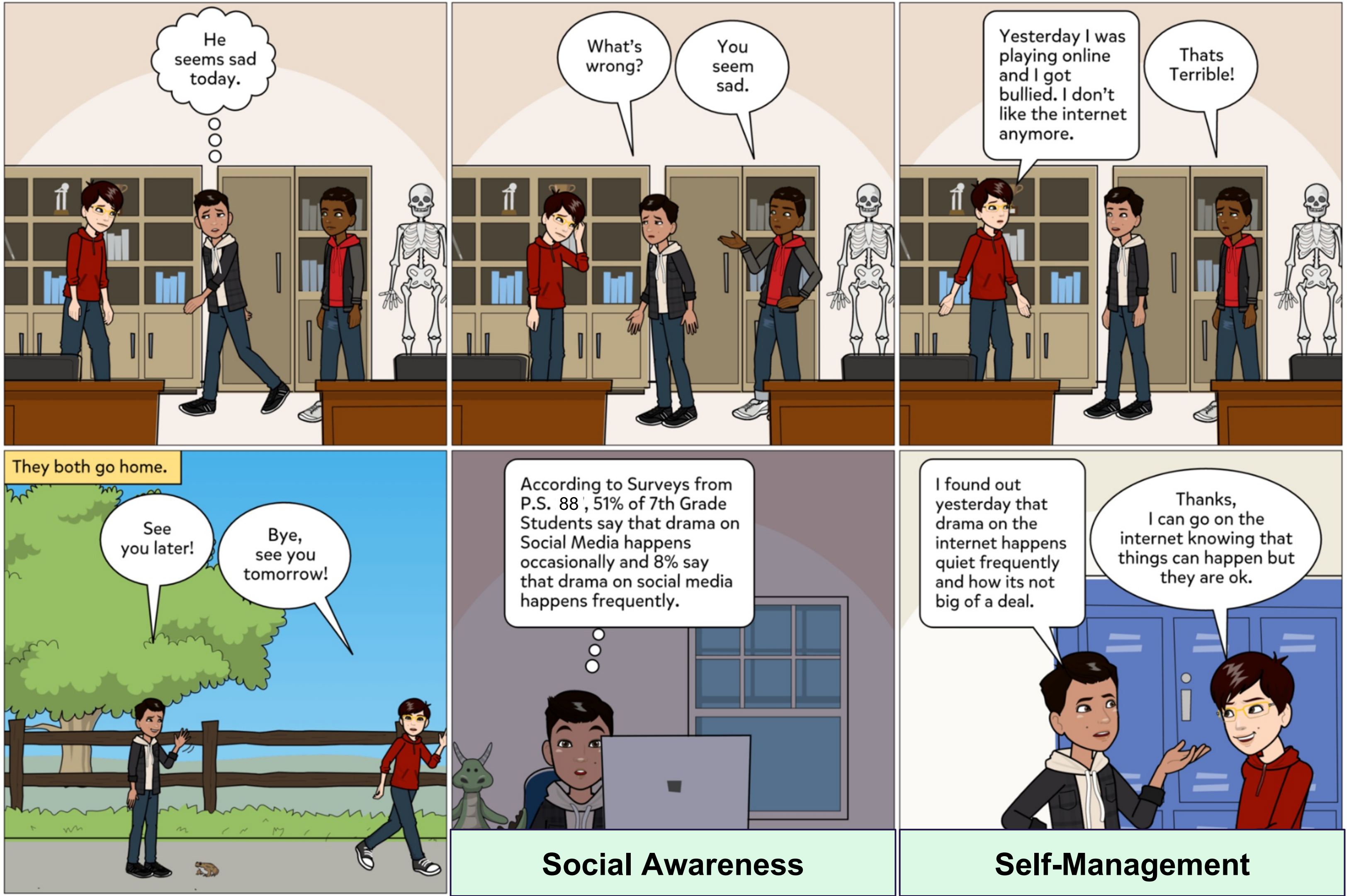
Data Reasoning Codebook

Incorrect Interpretation	Incorrect or partially incorrect interpretation of the data in support of an argument
Descriptive	Lists a data point, percentage, or descriptor of the data
Contextualizing the Data	Interpretation of the data relationships in relation to context. Generalizing to context
Individual Experience	Reflecting on an individual experience (or data point) in relation to the relationship communicated in the data
Proportional Comparisons	Uses percents to compare different categories in the data.
Implications of the Data	Reason about how data or other evidence support a hypothesis or claim
Inquiry about Related Data	Engages in inquiry or makes an assumption about a data relationship by connecting it to another variable or relationship that is not present in the data

Socioemotional Learning Codebook [Based on CASEL framework]

Self-Awareness	Any character describes or reflects on their own emotions, thoughts and values
Self-Management	Any character engages in or reflects on behaviors to deal with their emotions over time or across contexts
Social Awareness	Any character reflects on other characters or audience members perspectives
Relationship Skills	Reflection on a specific skill that is important for establishing and maintaining relationships
Responsible Decision Making	Reflection or demonstration of characters making caring and constructive choices about personal behavior and social interactions across diverse situations


He seems sad today.
What's wrong?
You seem sad.
Yesterday I was playing online and I got bullied. I don't like the internet anymore.
Thats Terrible!
Contextualizing
They both go home.
See you later!
Bye, see you tomorrow!
According to Surveys from P.S. 88, 51% of 7th Grade Students say that drama on Social Media happens occasionally and 8% say that drama on social media happens frequently.
Descriptive
I found out yesterday that drama on the internet happens quiet frequently and how its not big of a deal.
Thanks, I can go on the internet knowing that things can happen but they are ok.
Individual Experience


He seems sad today.
What's wrong?
You seem sad.
Yesterday I was playing online and I got bullied. I don't like the internet anymore.
Thats Terrible!
They both go home.
See you later!
Bye, see you tomorrow!
According to Surveys from P.S. 88, 51% of 7th Grade Students say that drama on Social Media happens occasionally and 8% say that drama on social media happens frequently.
I found out yesterday that drama on the internet happens quiet frequently and how its not big of a deal.
Thanks, I can go on the internet knowing that things can happen but they are ok.
Social Awareness
Self-Management

integrated Computational Thinking

Aman Yadav, PhD (PI)
MSU

Rafi Santo, PhD (co-PI)
Telos Learning

Carlos Leon (co-PI)
Telos Learning

Secil Caskurlu, PhD
MSU - Postdoc

David Phelps, PhD
Telos Learning-Fellow

MICHIGAN STATE UNIVERSITY

mouse_

Telos Learning

NATIONAL ACADEMIES Sciences Engineering Medicine

NSF

This work is supported by the National Science Foundation under grant number 1933933. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Motivation

K12 administrators are aiming to develop comprehensive, district-wide computer science initiatives, but middle school options are limited, largely focusing on integration into science and math.

Lack of meaningful integration approaches that *enhance* disciplinary learning in humanities and arts can lead to teacher backlash.

integrated Computational Thinking

Objectives

- **Identify design principles ("CT integration pathways") in Language Arts, Social Studies, and Arts** that enhance disciplinary learning in these subject areas by drawing on computational thinking.
- **Understand institutional and disciplinary factors** that mediate possibilities for meaningful integration of computational thinking in these subject areas.
- **Develop evidence-based professional development resources** that support CT integration in Language Arts, Social Studies, and Arts.

integrated Computational Thinking

Data science is present in integration pathways we identified in all three disciplines.

Language Arts	Social Studies	The Arts
Analyze Texts Through Computational Methods	Create Models and Representations in Social Studies	Create Computational and Computationally Enhanced Art
Enhance Writing Processes Through Computational Practices	Engage in Data Practices for Social Studies Inquiry	Explore Art Through Computational Thinking
Compose Interactive Computationally Enhanced Texts	Understand Computing's Impact on Society	See Data in Art & Make Data as Art
Critical Analysis of Computational Texts and Practices		

projects.ctintegration.org

NATIONAL ACADEMIES Sciences Engineering Medicine

integrated Computational Thinking

Epistemic Tensions between Computational Thinking and Humanities

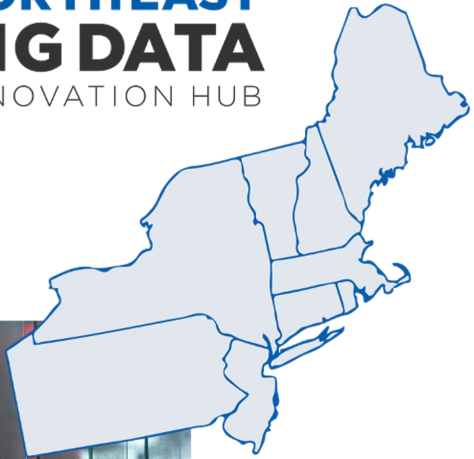
Epistemic Tension	**Definition**
Contextual Reductionism	Valuation of the abstract and quantifiable overrides the valuation of nuance and particularity in historical events or literary texts
Procedural Reductionism	Inappropriate application of algorithmic logics to knowledge production
Epistemic Chauvinism	Elevation of CT epistemologies at the expense of those related to a focal discipline leading to "sidelining" existing ways of knowing.
Threats to Epistemic Identities	Elevation of CT identities alienating learners who identify with epistemic identities associated with humanities disciplines
Epistemic Convergence	"Reskinning" certain humanities practices as CT foreclosing possibilities for substantial interdisciplinary novelty in problem solving

NATIONAL ACADEMIES Sciences Engineering Medicine

Data Science For All

NORTHEAST BIG DATA INNOVATION HUB

- Needs of multiple sectors
- Community of practice
- Data literacy essential principles



Stephen Uzzo

nysci

New York Hall of Science

Collaborative Inquiry

New York Hall of Science

Communities of Need

Data for Good Exchange & community/library workshops)

From panel discussion to action:
Elmcor, CBO in Corona, Queens; data analytics support

To state-wide impact:
NY State Office of Alcoholism and Substance Abuse Services

nysci
New York Hall of Science

Big Data for Little Kids

- *Structure:* 7-week workshop, meeting once per week for 1.5 hours
- *Theme:* Use data to help you design a new exhibit for the Hall of Science.
- *Families:* Each iteration involved 7-10 local families
- *Languages:* Facilitators spoke English, Spanish, Mandarin

nysci
New York Hall of Science

Data Science for High School Computer Science Workshop: *Identifying Needs, Gaps, and Resources*

- Need for inclusive tools, resources and curricula
- Must be available to all students
- Integration across the curriculum
- Need for teacher support and teacher enfranchisement in the process

nysci

New York Hall of Science

DataWise

- User-centered design process
- Users ARE the designers
- Involved in the design process at the outset, not brought in later
- Individual perspectives on design matter
- Goal is usability and flexibility in teaching and learning practice

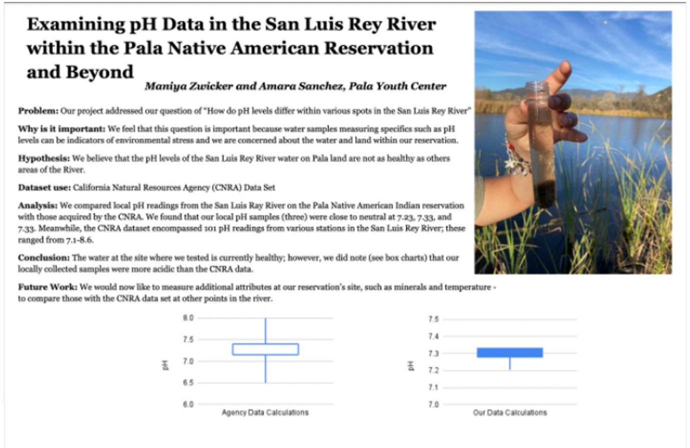
nysci

New York Hall of Science

NATIONAL ACADEMIES Sciences Engineering Medicine

DataJam: Online

- Form a team of 2 or more students at the participating school.
- Team brainstorms a research question, and formulate a hypothesis. *(with help from DataJam mentors)*
- Use big data to uncover answers, and make their own data visualizations. *(help from Data Jam mentors)*
- Present their findings.



Examining pH Data in the San Luis Rey River within the Pala Native American Reservation and Beyond

Maniya Zwicker and Amara Sanchez, Pala Youth Center

Problem: Our project addressed our question of "How do pH levels differ within various spots in the San Luis Rey River"

Why is it important: We feel that this question is important because water samples measuring specifics such as pH levels can be indicators of environmental stress and we are concerned about the water and land within our reservation.

Hypothesis: We believe that the pH levels of the San Luis Rey River water on Pala land are not as healthy as others areas of the River.

Dataset use: California Natural Resources Agency (CNRA) Data Set

Analysis: We compared local pH readings from the San Luis Ray River on the Pala Native American Indian reservation with those acquired by the CNRA. We found that our local pH samples (three) were close to neutral at 7.23, 7.33, and 7.33. Meanwhile, the CNRA dataset encompassed 101 pH readings from various stations in the San Luis Rey River; these ranged from 7.1-8.6.

Conclusion: The water at the site where we tested is currently healthy; however, we did note (see box charts) that our locally collected samples were more acidic than the CNRA data.

Future Work: We would now like to measure additional attributes at our reservation's site, such as minerals and temperature - to compare those with the CNRA data set at other points in the river.

PITTSBURGH DATAWORKS

nysci
New York Hall of Science

https://nebigdatahub.org/ds4all/

Funding provided by the NSF awards

1027752, 1139478, 1509079, 1636736 and 1922898

Additional Support from the Northeast Big Data Innovation Hub

NORTHEAST BIG DATA INNOVATION HUB

Bloomberg Philanthropies

NSF

nysci

New York Hall of Science

suzzo@nysci.org

Working Lunch
What are the Outcomes that We Want?

05

Small Group Discussion: Desired Outcomes for K-12 Data Science Education

Outcomes can be student-level (e.g., skills, proficiencies, interest/identity) or related to policy and practice (e.g., access, funding, preparation, standards)

- Participants have a number on their badge – this is your small group
  - You can eat inside or outside
  - Link to slide deck: NASEM K12 DS Outcomes
  - Be prepared to share reflections during next session
- **Where to meet your groups**
  - Groups 1 and 2: **Lecture Room**
  - Groups 3 and 4: **East Court**
  - Groups 5 and 6: **Lobby area in front of East Court**
  - Groups 7 and 8: **NAS 114**
  - Groups 9 and 10: **NAS 125**

Invited Commentary & Report Out on Outcomes

06

Explore evidence on learning and critical data literacy to consider what students should be able to do with data and how should those outcomes be measured.

SESSION GOALS

Panelists

Moderator: Nick Horton, Amherst College

<u>Presenters:</u>

Ryan Seth Jones	Middle Tennessee State University
Jo Louie	Education Development Center, Inc.

<u>Discussant:</u>

Jo Boaler (virtual)	Stanford University

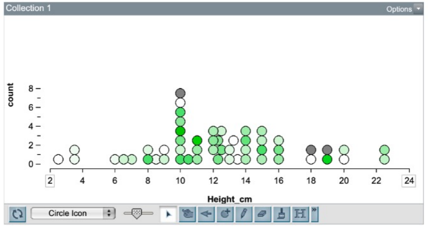
Ryan Seth Jones

Educational Leadership
MIDDLE TENNESSEE STATE UNIVERSITY

NSF

CAREER: Supporting Model Based Inference as an Integrated Effort Between Mathematics and Science

her...	Student_...	Plant_N...	Condition	Height_...	Number_of_Lea...	Number_of_Buds_and_Flo...	Number_of_Se
e ...	Yoelyn ...		Control	8.5	4	1	
e ...	Chastel...		Control	10.5	15	10	
e ...	Xander	A	Control	12	6	5	0

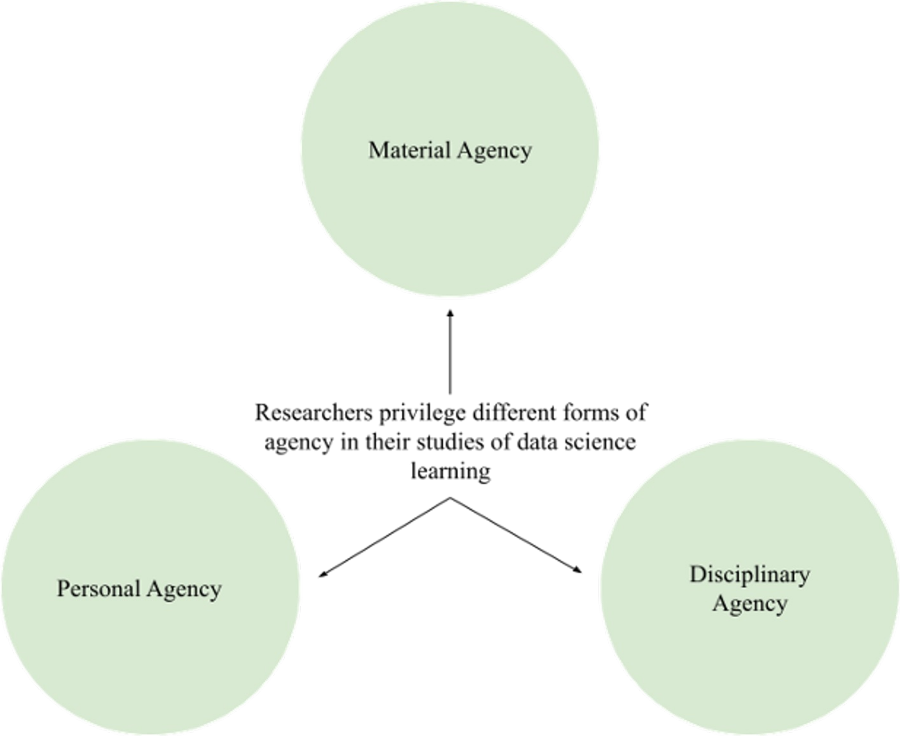


What Do We Know About Data Science Learning?

This is a difficult question because:

- Data Science, as a field, is emergent and diverse
  - Many different communities generate and revise knowledge using “data science” ideas and practices
  - Technologies, methods, and problems change quickly, and will continue to do so
  - It’s impossible to know what types of problems
- Research on Data Science learning is happening in many different communities
  - Diverse contexts (K-12 classrooms, undergraduate classrooms, museums, in homes, etc.)
  - Diverse theoretical traditions (cognitive, disciplinary focused, socio-cultural, etc.)
  - It’s difficult to synthesize research on student learning across these differences

Data Science Learning as a Dance of Agency



Rosenberg J.M. & Jones R. S., (2022). A Secret Agent? Data Science Learning Through the Lens of Agency. Commissioned paper for the National Academy of Sciences, Engineering, and Medicine *Foundations of Data Science for Students in Grades K-12: A Workshop.*

Data Science Learning as a Dance of Agency

- Engaging students with material agency provides opportunity to learn fundamental ideas about variability, measurement, sampling, etc.

- Material agency supports learners to understand the context through which data came to be.

- It can be productive for students to describe sources of variability, including variability due to measurement, the phenomenon under investigation, and random noise.

Data Science Learning as a Dance of Agency

- Learners possess epistemic assets, and it's important to recognize these assets.
- Though students are unlikely to reinvent all of data science, they can be positioned to see data as personally meaningful, construct coherent data stories, and develop data science approaches that resemble disciplinary approaches in meaningful ways.
- It is important to scaffold and sequence students' learning about data in such a way that permits learners to not only exert personal agency but also make progress toward generating answers to the questions on which they were working.

Data Science Learning as a Dance of Agency

- Research that prioritizes disciplinary agency has emphasized the roles of computing and highly valued disciplinary procedures and tools.
- Many studies have chosen specific programming languages that are valued in the discipline, especially R and python.
- Students can learn to use particular tools valued in a discipline when supported through careful instructional design.

Data Science Learning: Where Do We Go From Here?

A few thoughts on moving forward…

- To what ends can we orient our work?
- What is the role of systems level organizations and support?
- How much weight should be given to the different dimensions of data science learning? (material, personal, and disciplinary agency)
- Whose voice is highlighted in research on data science learning?
  - Where do you identify in these categories?

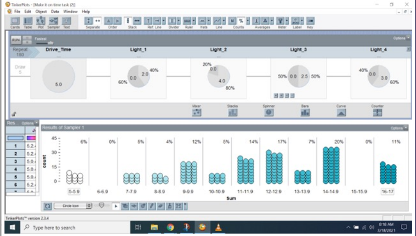
Data Science Learning: Where Do We Go From Here?

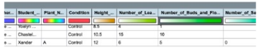
A few thoughts on moving forward…

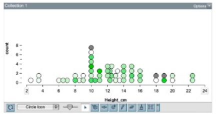
- To what ends can we orient our work?
- What is the role of systems level organizations and support?
- **How much weight should be given to the different dimensions of data science learning? (material, personal, and disciplinary agency)**
- **Whose voice is highlighted in research on data science learning?**

CAREER: Supporting Model Based Inference as an Integrated Effort Between Mathematics and Science

Level	Description
6	Compare competing models for the same distribution of observed values
5	Develop a model that accounts for the distribution observed
4	Construct explanations that account for sources of variability
3	Structure a collection of measures as a distribution and measure distributional features with statistics
2	Develop or appropriate a measure and apply to a collection
1	Describe qualitative differences in a collection







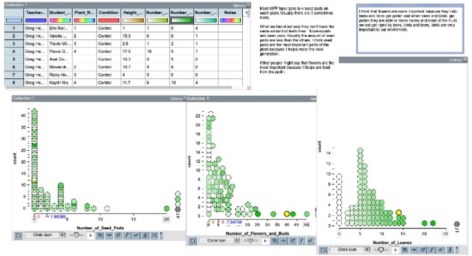
Learning To Explain Variability

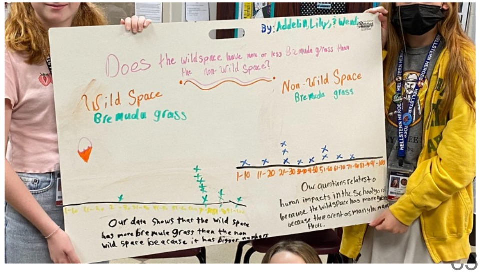
NATIONAL ACADEMIES Sciences Engineering Medicine

CAREER: Supporting Model Based Inference as an Integrated Effort Between Mathematics and Science

Helping Teachers Design Instruction Where Students Explain Variability

Sources of Variability	How Children Might Encounter The Variability	Questions This Variability Might Evoke or Inform	Resources Students Might Use to Make Sense of The Variability	Scientific Practices This Variability Might Support	Scientific Concepts This Variability Might Support





Josephine (Jo) Louie

Education Development Center

Critical data literacy:

Creating a more just world with data

Josephine (Jo) Louie

Education Development Center

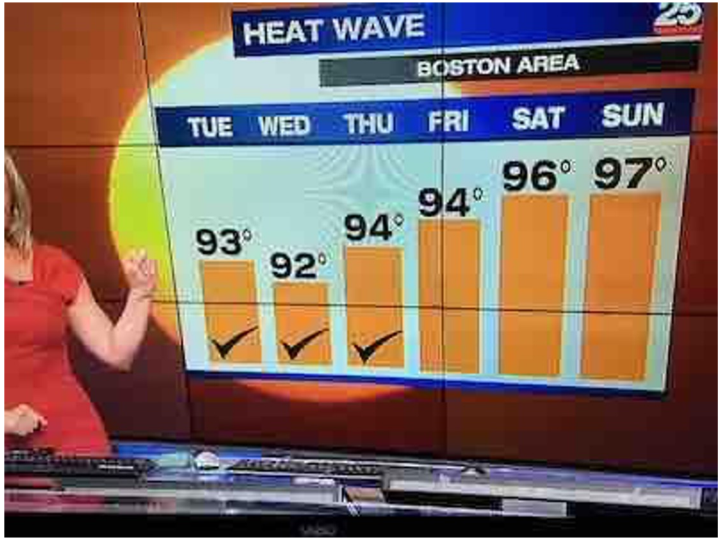
SDLC Strengthening Data Literacy across the Curriculum

A collaboration among EDC, California Polytechnic State University, and The Concord Consortium
https://sites.google.com/view/uss-data/home

WeatherX

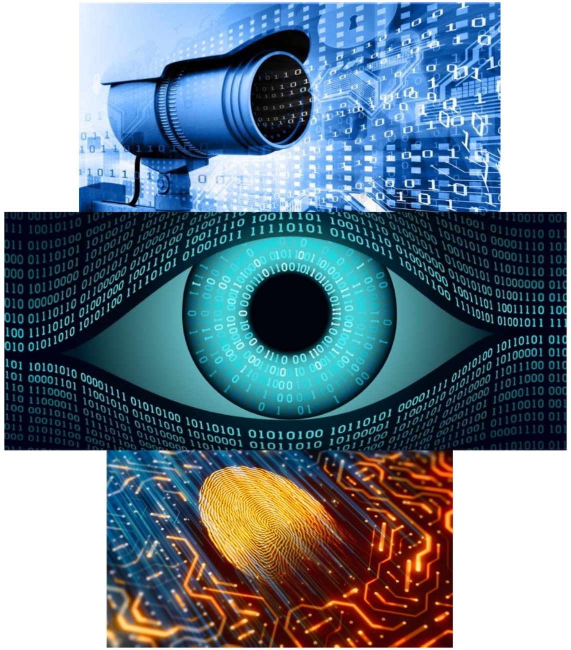
A collaboration among EDC, Mount Washington Observatory, University of Maine-Orono, University of Washington-Seattle, and The Concord Consortium
https://www.edc.org/weatherx

People have long called for stronger data literacy



- 64% to over 80% of people in business and public surveys have little confidence in their abilities to make sense of data (H+K Global, 2016; Qlik, 2018)

Big Data raises new questions about privacy and power

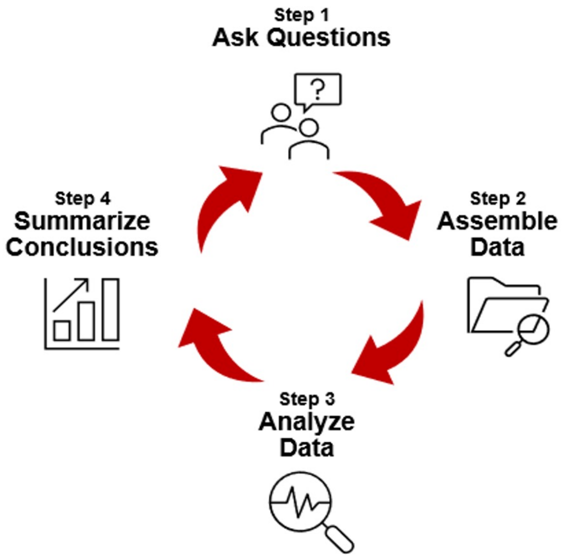


- “While digital technology users consciously volunteer masses of data on a daily basis, in many other instances data are collected without an individual’s knowledge.” (Pangrazio & Selwyn, 2019, p. 420)

- “This new form of information capitalism aims to predict and modify human behavior as a means to produce revenue and market control.” (Zuboff, 2015)

Growing calls for critical data literacy

An approach anchored in statistical literacy

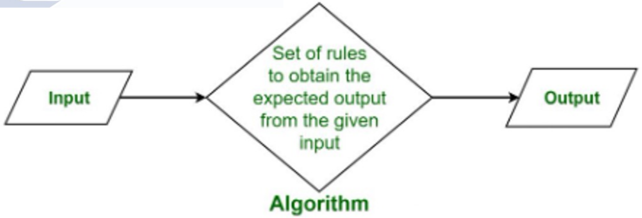


- Requires understanding of and fluency with the data investigation process
- Includes familiarity with multivariable reasoning
- Involves questioning throughout

\- Bargagliotti et al., 2020; Engel, 2016; Ridgway, 2015

Growing calls for critical data literacy

Must expand to address new data realities for individuals and society

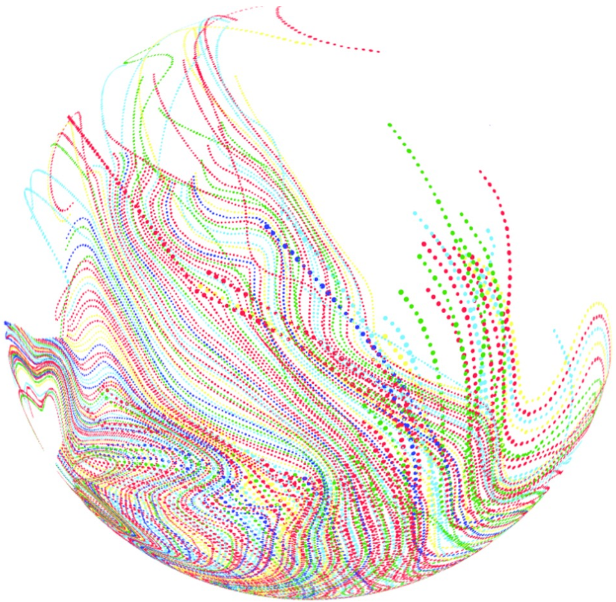


- Understand when data are collected from us
- Understand algorithms and what they do
- Consider ethical impacts of data collection and data-based decisions

- D'Ignazio & Bhargava (2015)

Possible social futures with critical data literacy

Read and **write** the world with data



- Raise social and political awareness with data
- Develop social agency with data
- Build pride in one's cultural and social identity

- Gutstein (2003, 2006)

Examples of critical data literacy interventions

READ the world with data	***WRITE the world with data***
• Uncover social inequality Kokka (2020), Louie et al. (2021a, 2021b), Rubel et al. (2016)	• Envision new public spaces Taylor et al. (2019), Van Wart et al. (2020)
• Confront the ethics of Big data Vakil et al. (2020)	• Craft one's family history Kahn (2020)
	• Express one's self Stornaiuolo (2020)

Reported successes

- *Reading the world with data*: More critical perspectives of society
  - More critical opinions of the lottery (Rubel et al., 2016)
  - Greater “civic empathy” over housing and economic inequality (Kokka, 2020)
  - Stronger awareness of the scale and persistence of income inequality (Louie et al., 2021a,b)
  - Shock and concern over data privacy violations (Vakil et al., 2020)
- *Writing the world with data*: Stronger agency in authoring with data
  - Develop alternative bus tour, park vision (Taylor et al., 2019; Van Wart et al., 2020)
  - Write own family geobiography (Kahn, 2020)
  - Design own data story on a T-shirt (Stornaiuolo, 2020)

Reported challenges

- Critical sociopolitical views may not arise naturally or easily
  - Student resistance (Brantlinger, 2013; Rubel et al., 2016;)
  - Persistence of existing views (Enyedy & Muhopadhyay, 2007; Kokka, 2020)
  - Focus on the personal rather than social (Stornaiuolo, 2020)
- Discussing social and political inequality can be difficult
  - Feelings of disempowerment are possible (Louie et al., forthcoming)
  - Negative stereotypes of non-dominant groups may arise (Kokka, 2020; Philip et al., 2016)

Reflections

- Learning to “read” and “write” the world with data may be a helpful framework for promoting critical data literacy in schools

- The example studies suggest promise in strategies to build critical awareness of unequal social structures, and agency in authoring one’s own stories and social visions with data

Future directions

- Can we have critical data literacy without quantitative reasoning?
  - We need frameworks and coordination if support must occur across time and contexts
- Promoting critical data literacy is interdisciplinary
  - We need spaces and support for such work to occur
- Scaling promising approaches is needed
  - We must identify and scale if we want critical data literacy for all
- We need frameworks and tools for assessing critical data literacy
  - What do we want students to know and be able to do, and how will we measure it?
- How promote critical data literacy among different groups?
  - We need sensitivity to historical positions of power and strategies against headwinds

Jo Boaler

- The Nomellini-Olivier professor of mathematics education at Stanford University
- Youcubed is a center at Stanford providing free mathematics resources - including a one-year data science high school course, partnered with Google.

Explorations in Data Science taken by:

Over 160,000 students
1830 teachers

47% girls/non-binary
57% students of color
68% not mathematically accelerated

Youcubed: Explorations in Data Science: 8 Units

Data tells a story	ACS Data Distributions	Bivariate Data	Probabilistic Modeling
Categorical Data	Modeling with Data	Data Predictions	Being a Data Scientist

In partnership with Google

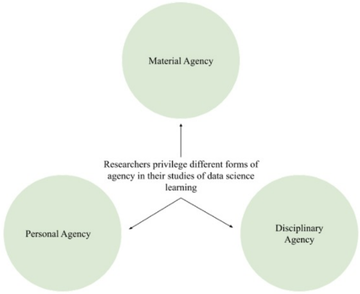
A Project Based One Year Course

NATIONAL ACADEMIES Sciences Engineering Medicine

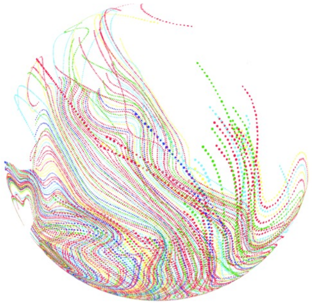
youcubed K-12 Data Science

What is important learning within data science?

And what does it mean for assessment?



Seth: The Dance of Agency



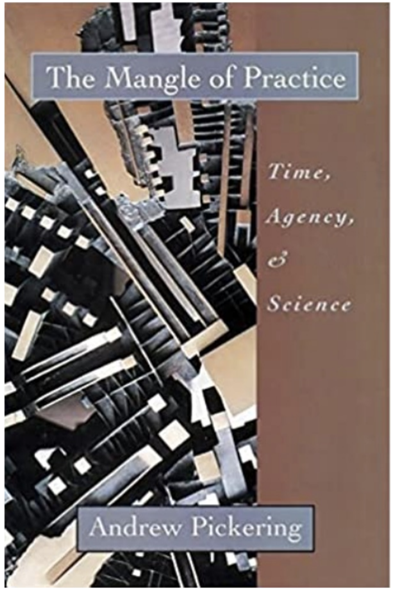
Jo: Critical data literacy

youcubed K-12 Data Science

NATIONAL ACADEMIES Sciences Engineering Medicine

Seth Jones: Agency

Andrew Pickering - Material, Personal & Disciplinary agency



In data science students need to ask questions, choose methods, and apply them, build models then interpret & communicate results - these are all acts of agency

In data science this personal agency comes in response to material agency (programming) and disciplinary agency (statistics) - in what Pickering calls a "dance of agency"

It is really important that we keep this "dance" in mind as we design assessments

NATIONAL ACADEMIES Sciences Engineering Medicine

youcubed K-12 Data Science

Jo Louie: Critical Data Literacy

- Privacy and power
- Questioning
- Multivariable reasoning
- Ethical implications of data collection

To read the world and answer it with data.

What is Data Science?

Content Knowledge	**Practices**
Statistical Knowledge	Agency - material, disciplinary, personal
Programming	Choosing, using, adapting, interpreting
Multivariate reasoning about variability	Critical Data Literacy

It is easy to assess content knowledge with narrow tests, but if we ignore the practices, we will miss the essence of data science and threaten the achievement of equitable outcomes

The Importance of what we assess for equity

SAT – More than a third of the variance in SAT scores is predicted by student race, after achievement has been factored out. (Geiser, 2015)

Smarter Balanced - a broader assessment – calculator in most questions, reasoning, mathematical practices

Using GPA Plus SBAC results in a more socioeconomically and racially diverse college pool than GPA plus SAT (Kurlaender & Cohen, 2019).

- **Assessments for Data Science need to include agency, and critical thinking**

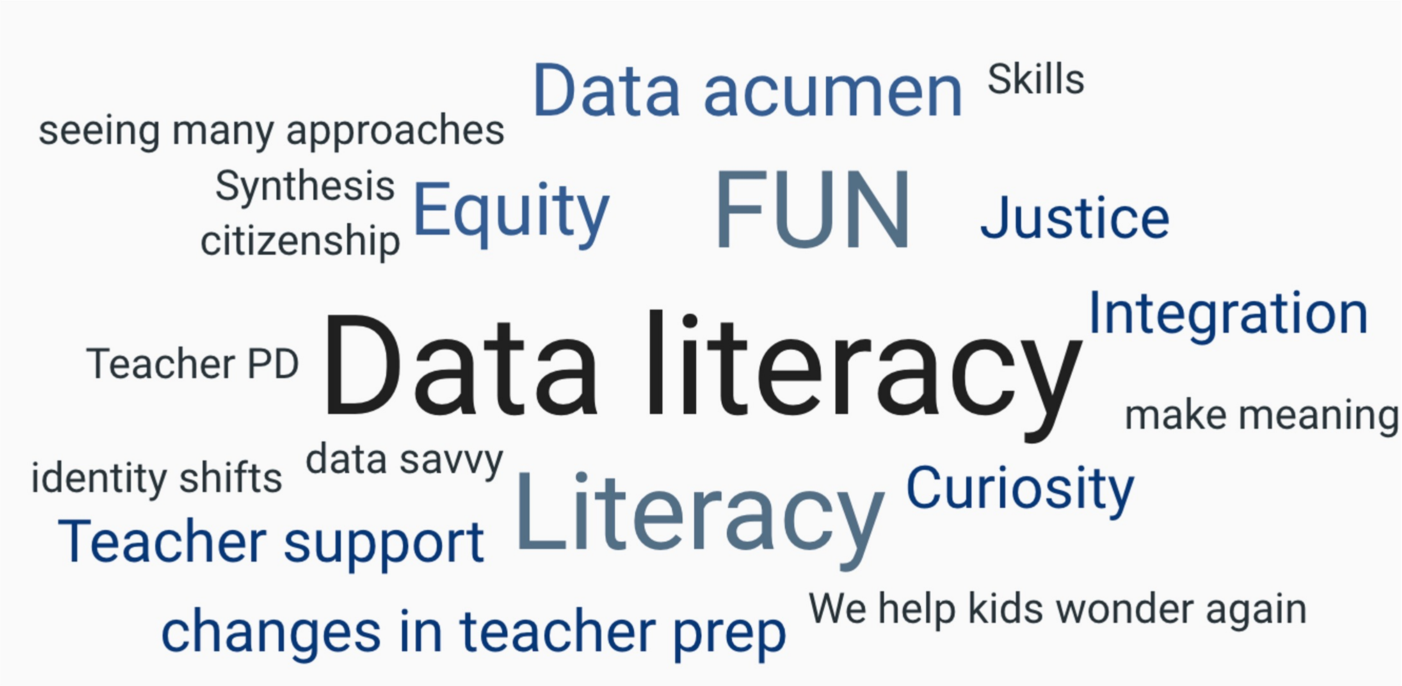
Students end of Course Data Projects

- Does the age of the congress members accurately represent the people in their state?
- Which is better for the environment – thrifting or fashion stores?
- How do untreated mental disorders affect people?
- Are education levels linked to poverty?
- Is climate change real?
- Who grew up with childhood trauma and how does it affect their plans?
- How often do oil spills happen in our environment? How educated are our students about them?
- Does preparation for AP exams cause a decline in mental health?

If we do not assess agency and critical literacy we will miss the essence of data science

Poll Results

What are some desired outcomes for K-12 data science education?



Panel Discussion
How are Tools and Resources Supporting Data Science Learning Experiences?

07

This session will explore the range of tools and data sets that exist and are needed to support students learning in acquiring data understanding and skills.

SESSION GOALS

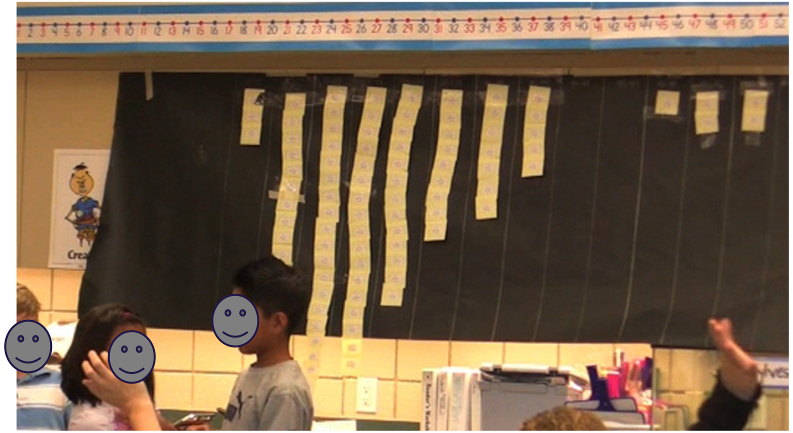
Panelists

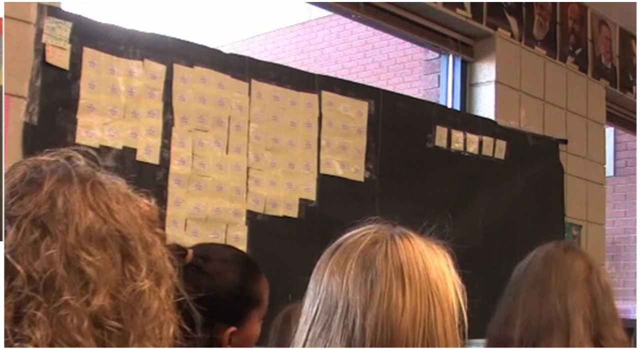
Moderator: Tim Erickson, Epistemological Engineering

Rolf Biehler (virtual)	Paderborn University
Chad Dorsey	Concord Consortium
Randy Kochevar	Education Development Center, Inc.
Victor Lee	Stanford Graduate School of Education
Andee Rubin	TERC

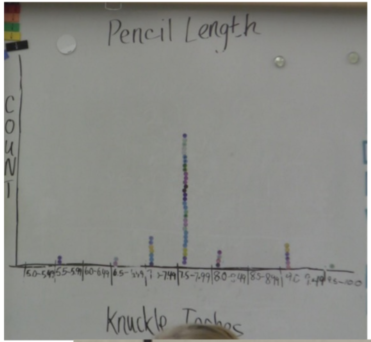
Quantified Selves in the Classroom - Analog Visualizations

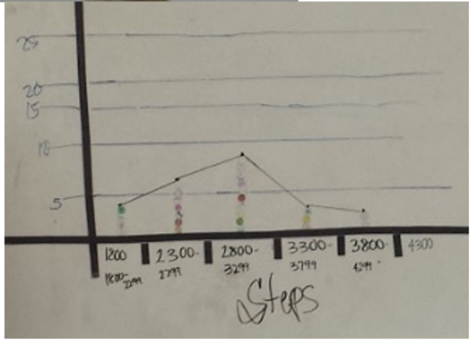
Comparing trade-offs of invented sticky-note 'histograms'





Magnetic whiteboards, dot magnets

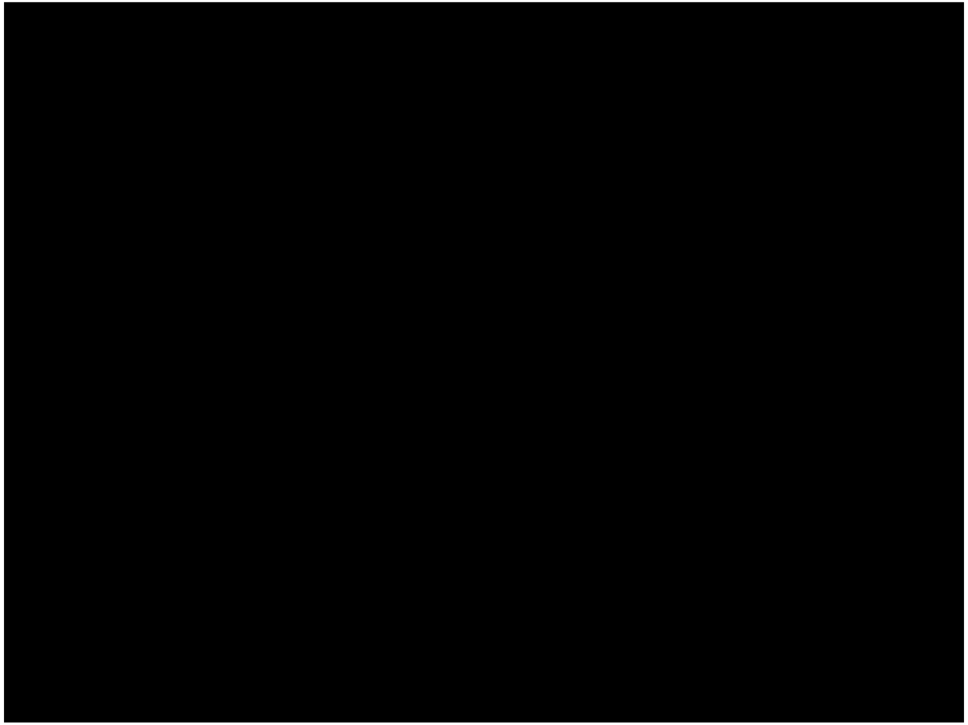




Line up by the number of years you’ve been teaching in your current school.

Then shift to the TOTAL number of years you’ve been teaching

CODAP does just what the teachers did.



We need to find “just right” datasets - not too small, not too large…

Shrimp for fall 2022

cases (3744 cases)

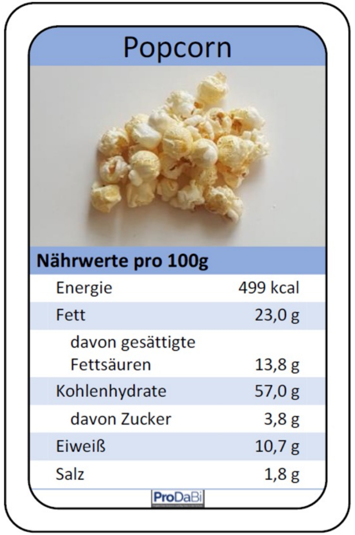
in-dex	week	season	species	pool	shrimp per trap	dis-charge	mean rainfall	mean …ir temp
1	1	winter	ATYACP...	P0	75.2	1.82	3.3	19.3
2	2	winter	ATYACP...	P0	57.3	1.81	4.8	21
3	3	winter	ATYACP...	P0	65.2	1.9	8.9	19
4	4	winter	ATYACP...	P0	65	1.15	4.9	20.3
5	5	winter	ATYACP...	P0	51.5	0.13	10.1	20.2
6	6	winter	ATYACP...	P0	61	0.13	4.6	18.8
7	7	winter	ATYACP...	P0	56.7	0.13	9.2	19.4
8	8	winter	ATYACP...	P0	52	0.13	3	19.3
9	9	winter	ATYACP...	P0	46.2	0.13	0.1	20.2
10	10	winter	ATYACP...	P0	74.3	0.13	12.9	18.5
11	11	winter	ATYACP...	P0	57.2	0.14	0.8	18.8
12	12	spring	ATYACP...	P0	54.7	0.18	9.4	19.6
13	13	spring	ATYACP...	P0	59.2	0.42	8.7	21.5
14	14	spring	ATYACP...	P0	49.3	2.97	33.2	20.5
15	15	spring	ATYACP...	P0	52.2	0.34	5.5	21.2
16	16	spring	ATYACP...	P0	60.7	0.53	10.6	20.8
17	17	spring	ATYACP...	P0	53.5	1.53	22.4	19.6
18	18	spring	ATYACP...	P0	46.8	1.13	10.9	20.4
19	19	spring	ATYACP...	P0		1.78	21.6	21.6
20	20	spring	ATYACP...	P0		1.83	21.5	21.8
21	21	spring	ATYACP...	P0	38.7	1.89	19.7	21.8
22	22	spring	ATYACP...	P0	38.5	2.25	24.8	21.8
23	23	spring	ATYACP...	P0	61	1.58	17.8	22.4
24	24	spring	ATYACP...	P0	15.2	1.46	16.9	22.5

Data Jam Shrimp (1993-2019)

cases (1348 cases)

in-dex	Week	Season	ATYACPUE-P0	XIPHCPUE-P0	MACCPUE-P0	ATYACPUE-P8	XIPHCPUE-P8	MACCPUE-P8	ATYACPUE-P9	XIPHCPUE-P9	MACCPUE-P9	Discharge (m3/sec)
24	[illegible]	summer	[illegible]	[illegible]	0				[illegible]	1	0	
25	30	summer	86.3	45.7	0				9.3	2.3	0	
26	31	summer	71	45.7	1				27	1.7	0	
27	33	summer	28.5	35.8	0.3				14	0.3	0	
28	34	summer	33.3	26.3	0				6	0.3	0	
29	35	summer	26.2	24.5	0.3							
30	36	summer	25	15	0.5				3	1	0	
31	37	summer	27.5	21.2	0				18.3	5.3	0	
32	38	fall	24	19.7	0.5				5	2	0	
33	39	fall	33.7	21.5	0				7	3.7	0	
34	40	fall	29.2	14.8	0				18.7	4	0	
35	41	fall	32	24.7	0				26.7	3.3	0	
36	43	fall	38.3	17.8	0.2				15	1.3	0	
37	44	fall	67	16.3	0				13	2.3	0	
38	45	fall	46.2	16.5	0.3				7.7	1.7	0	
39	46	fall	53	19.8	0.2				7	1	0	
40	48	fall	36.8	20.2	0.3				9.7	0.3	0	
41	49	fall	49.3	24.8	0.2				6.7	1.3	0	
42	50	fall	36	17.5	0				6.7	1.7	0	
43	51	winter	48.3	22.7	0.2				6	2.7	0	
44	52	winter	65.8	23.8	0.2				5.3	4	0	
1	1	winter	69.2	31.5	0.2				17	4	0	
2	2	winter	32.8	19.2	0.8				18.7	0.7	0	
3	5	winter	84.7	23	0.3				21.3	0.3	0	
4	6	winter	64.2	2.7	0.5				9.7	0.3	0	

Decision trees unplugged with data cards (grade 5 and 6)



ProDaBi
Project Data Science and Big Data at School

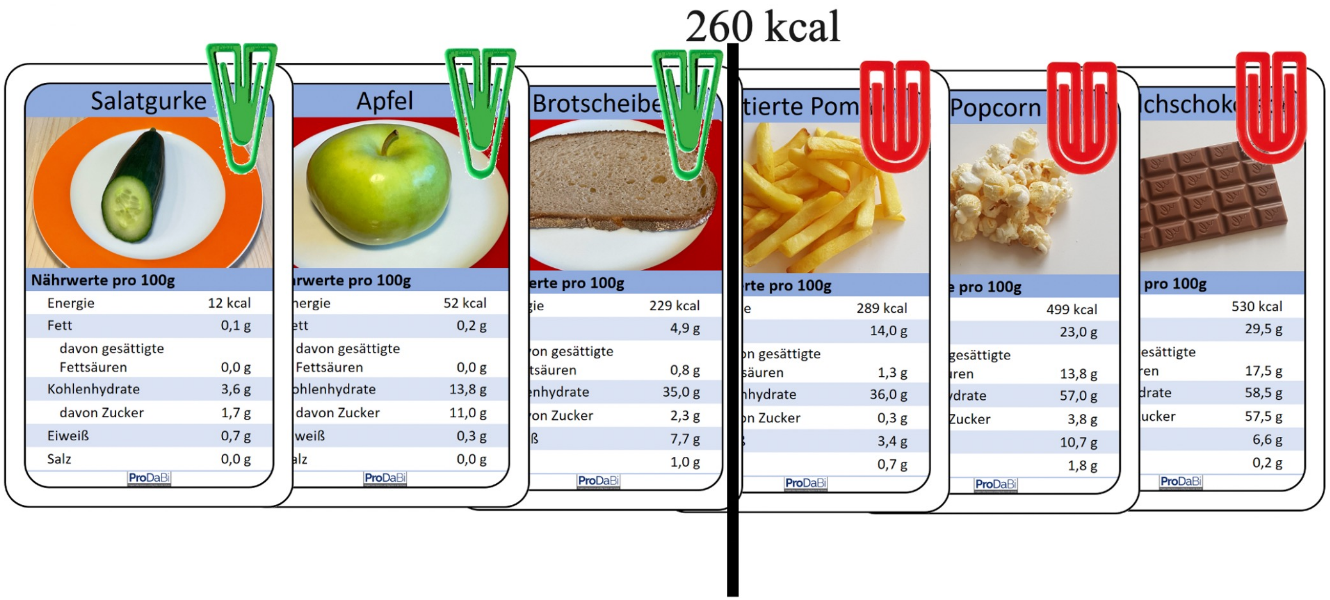
www.prodabi.de/en/

Rolf Biehler
PADERBORN UNIVERSITY
The University for the Information Society

Card game

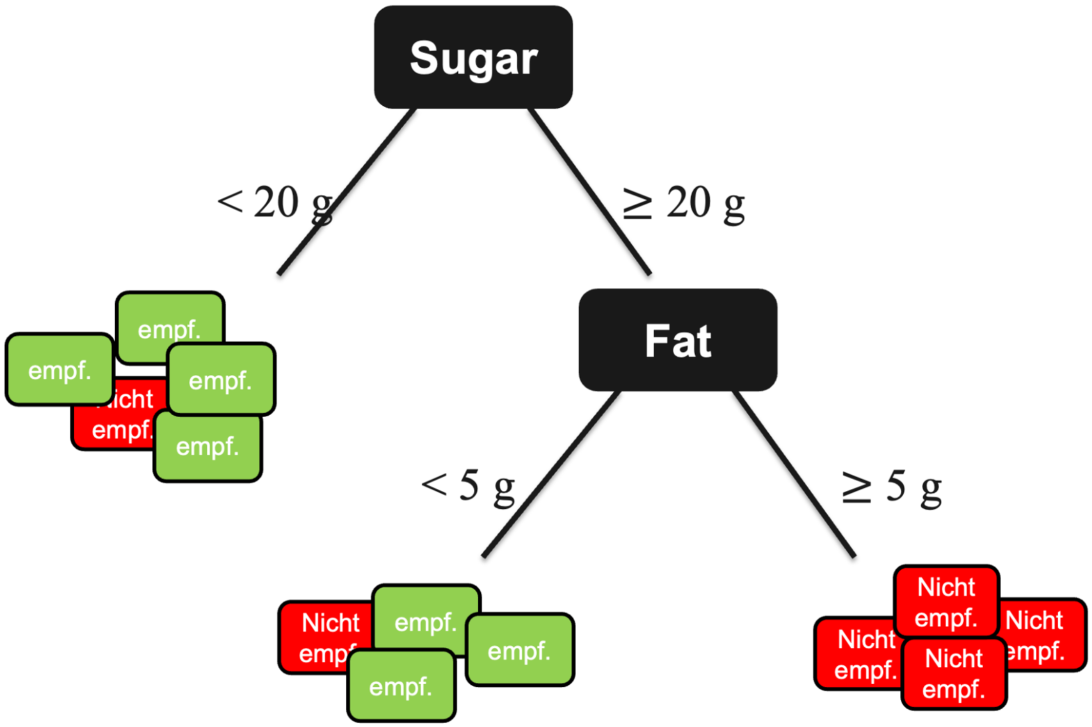

Haselnussschnitte
1
Popcorn
2
Apfel
3
Brotscheibe
4
Banane
5
Gummibärchen
6
Chips
7
Salzstangen
8
Nudeln
9
Reis
10
Frittierte Pommes
11
Erbsen
12
Eisbergsalat
13
Möhre
14
Frikadellen
15
Spiegelei
16
Brötchen
17
Knäckebrot
18
Zwieback
19
Marmorkuchen
20
Paprika (rot)
21

Finding „good“ splits in the data set



Decision trees unplugged with data cards

A multistep decision tree



Townhall

08

What are the highest priorities for additional research?

SESSION GOALS

Day 1
Wrap up and Adjournment

09

NATIONAL ACADEMIES *Sciences Engineering Medicine*

Foundations of Data Science for Students in Grades K–12 A Workshop

September 13-14, 2022

SEPTEMBER 2022

SEPTEMBER 2022

Welcome and Reflections on Day 1

01

Themes from Chat & Discussions

- We know about the **protective qualities** of some pedagogical approaches (e.g., identity-building): What do we know from other fields and how can we apply it here?
- How do we build **bipartisan support** and partnerships for more research and educational initiatives (taking the example of CS/computing ed)
- **Data Stories and confirmation bias**: Weighing alternative explanations with data; anticipating what data and analyses are needed to support (and reject!) claims
- Importance of **developmental perspectives**. Learning trajectories of students; how do these trajectories intersect with other disciplines.
- **Tools**: ideal tools don't exist, how to work with existing tools (and understand what they do well and what they don't)
- Data types beyond numeric: Chat was particularly active around **maps and geographic data** as particularly rich and important.
- Building **confidence**: how do we help instructors (and students) develop mastery amidst the cognitive load (so many moving parts)
- Building **student-centered** approaches where all are supported for success

Genres

- Technological Tools (always changing, but similar in type)
  - Spreadsheets; Visual Tools; Scripting Tools; (*and* Unplugged Experiences)

- Curricular Approaches (time scale, purpose, enactment spaces)
  - Mini-integrations; Replacement Units; Courses

- Looking *within* genres for best practices
- Looking *across* genres for complementarities, trajectories, coherences.

Overview of the Agenda

State of the Field

- Hearing from Practice: What is Happening in and Out of Schools?
- How is Data Science Integrated in Content Areas?
- What is the State of Educator Preparation in Data Science?
- Townhall
- Funder Reflection
- Final Reflections from the Planning Committee

How to Engage

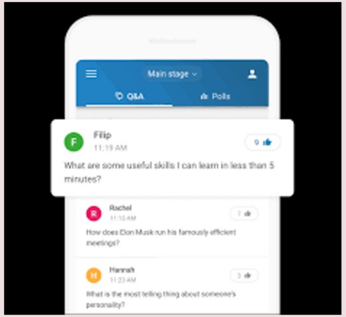
AUDIO

Ensure you are muted unless speaking.

CHAT

We encourage active engagement in the chat.

SLIDO



Interactive polls. Q&A questions.

Panel Discussion
Hearing from Practice: What is Happening in and Out of Schools?

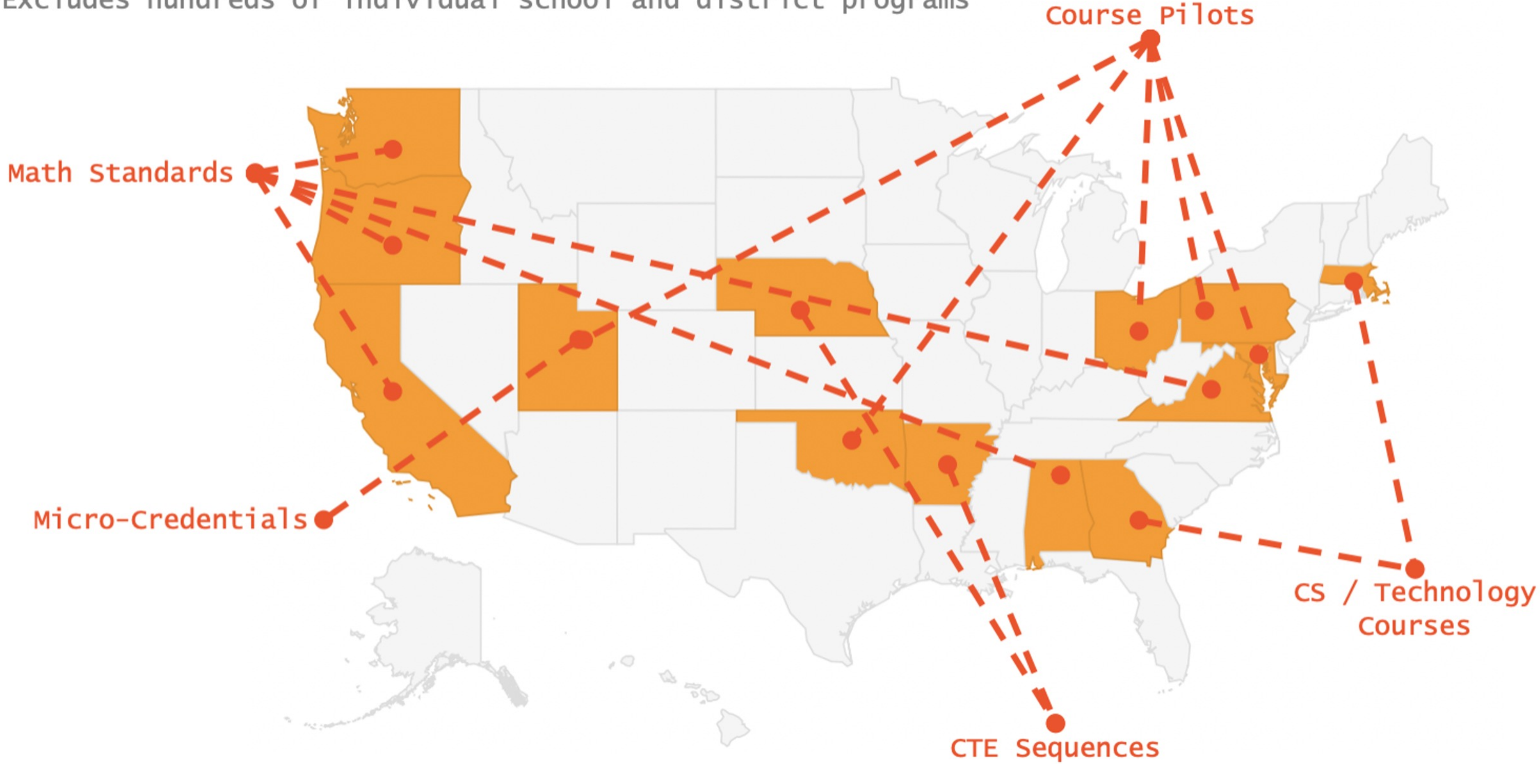
02

Explore the reality on the ground in K-12 data science education, with a deep focus on the design of student learning opportunities

SESSION GOALS

Discrete *Data Science* programs:

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



Discrete *Data Science* programs:

Estimated active K-12 Data Science programs:

1,600+ schools / districts

2,000+ teachers

180,000+ students

14 statewide programs

*Aggregated through curriculum providers and statewide programs.

**A lower-bound! Data is incomplete.

Discrete *Data Science* programs:

Estimated active K-12 data science programs:

1,600+ schools / districts (6.1% of 26,000 HS)

2,000+ teachers (0.1% of 1.8 million HS)

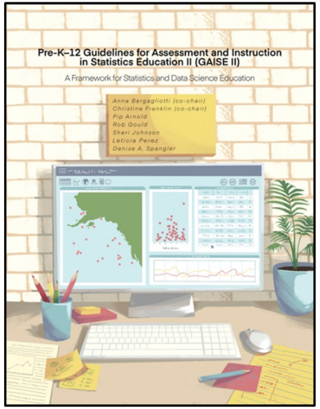
180,000+ students (3.6% of 15.1 million HS)

14 statewide programs

*Aggregated through curriculum providers and statewide programs.

**A lower-bound! Data is incomplete.

National context: data edu. in existing standards



1. Data collection (NGSS, K-12) (CCSS Math K-5, via measurement) (CS, K-2, tech / digital tools focus) (GAISE II, A)
2. Data tables (NGSS, 3-5) (CCSS, HS) (CS, 6-8, tables for data storage) (GAISE II, B)
3. Basic variance (mean / median / mode) (6-8 NGSS) (CCSS HS) (GAISE II, A)
4. Variables in data (GAISE II, A) (NGSS, P3, 3-5) (CCSS, HS)
5. Statistical or inference-based questions (CS, K-2) (CCSS, HS) (NGSS, P4, 3-5) (GAISE II, A)
6. Data visualizations (dot plot, histogram, box plots) (NGSS, 3-5) (CCSS HS) (CS, K-2) (GAISE II, A, as variability)
7. Outlier data (CCSS Grade 8, HS) (GAISE II, C)
8. Data types (qualitative, quantitative, etc.) (NGSS, K-2) (CCSS, 8) (CS 6-8, as characters, noms, bits) (GAISE II, A)
9. Correlation vs. causation (NGSS, P4, 6-8) (CCSS, HS) (CS, 9-12, in prediction) (GAISE II, C)
10. Function / Model Fit (NGSS, P4, 9-12) (CCSS HS, emphasizes linear, quadratic, and exponential models; residuals) (CS, 6-8, focus on tech fit)
11. Slope / intercept / correlation coefficient (NGSS, P4, 9-12) (CCSS, HS) (GAISE II, C)
12. Sample selection, sample vs. population (NGSS, P4, 9-12) (CCSS, HS) (GAISE II, B)
13. Bayesian probability; updates to priors (NGSS, P4, 9-12) (GAISE II) (CS K-12, 6-8)
14. Units, ratios, percents, compound units, basic algebra (NGSS, P5, 9-12) (CCSS, throughout)
15. Analysis types: sample surveys vs. experiments vs. observational studies (CCSS HS) (GAISE, C) (U.T. Dana Center)
16. Sampling: random sampling, sample vs. population (CCSS 6-8) (CS K-12 6-8, simulation) (GAISE, B)
17. Randomization (CCSS, HS) (GAISE II, B)
18. Compare two things with data:
    a. NGSS: two alternate solutions (P4, 3-5)
    b. CCSS HS: two treatments & significance (or two distributions) (6-8)
    c. GAISE II: two groups & association between two variables (A)
19. Data collection from modern tech devices (NGSS, devices only) (CS, devices & online data) (GAISE II, devices & online data)
20. Data cleaning (GAISE II, A) (CS, 6-8)
21. Limitations of data (NGSS, P4, 6-8) (GAISE II, B) (CS, 3-5)

National context: data edu. in existing standards

1. Data collection (NGSS, K-12) (CCSS Math K-5, via measurement) (CS, K-2, tech / digital tools focus) (GAISE II, A)
2. Data tables (NGSS, 3-5) (CCSS, HS) (CS, 6-8, tables for data storage) (GAISE II, B)
3. Basic variance (mean / median / mode) (6-8 NGSS) (CCSS HS) (GAISE II, A)
4. Variables in data (GAISE II, A) (NGSS, P3, 3-5) (CCSS, HS)

The word "data" appears **53 times** in the
NCSS *College, Career, and Civic Life (C3) Framework*

17. Randomization (CCSS, HS) (GAISE II, B)
18. Compare two things with data:
    a. NGSS: two alternate solutions (P4, 3-5)
    b. CCSS HS: two treatments & significance (or two distributions) (6-8)
    c. GAISE II: two groups & association between two variables (A)
19. Data collection from modern tech devices (NGSS, devices only) (CS, devices & online data) (GAISE II, devices & online data)
20. Data cleaning (GAISE II, A) (CS, 6-8)
21. Limitations of data (NGSS, P4, 6-8) (GAISE II, B) (CS, 3-5)

Possibly not enough? Teachers told us why:

- **Not actually taught**: existing data standards often cut, brushed-over, or in the background;
  - Low teacher confidence
  - Perception of assessments not prioritizing
- **Technology is missing:** students enjoy authentic technology, but few and far between;
  - DS programs have found students (and teachers!) enjoy the innovation
  - Standards treatment of tech is skimpy: brief mentions of spreadsheets, calculators, hand-held sensors

Possibly not enough? Teachers told us why:

- **Not actually taught**: existing data standards often cut, brushed-over, or in the background:
  - Low teacher confidence

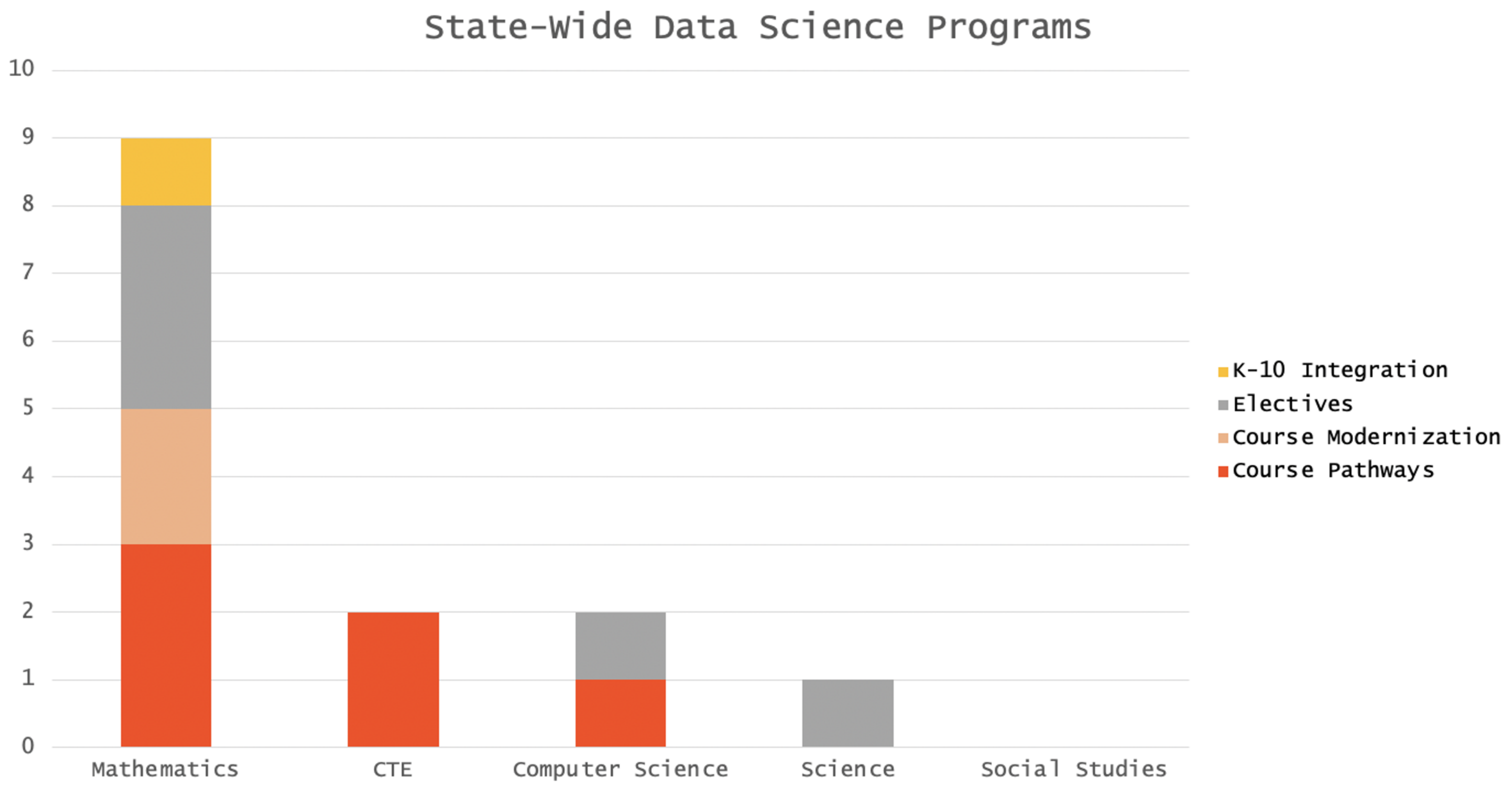
NGSS, Practice 4:
“use digital tools (e.g., computers) to analyze very large data sets for patterns and trends”

Possibly not enough? Teachers told us why:

- **Not actually taught**: existing data standards often cut, brushed-over, or in the background;
  - Low teacher confidence
  - Perception of assessments not prioritizing
- **Technology is missing:** students enjoy authentic technology, but few and far between;
  - DS programs have found students (and teachers!) enjoy the innovation
  - Standards treatment of tech is skimpy: brief mentions of spreadsheets, calculators, hand-held sensors

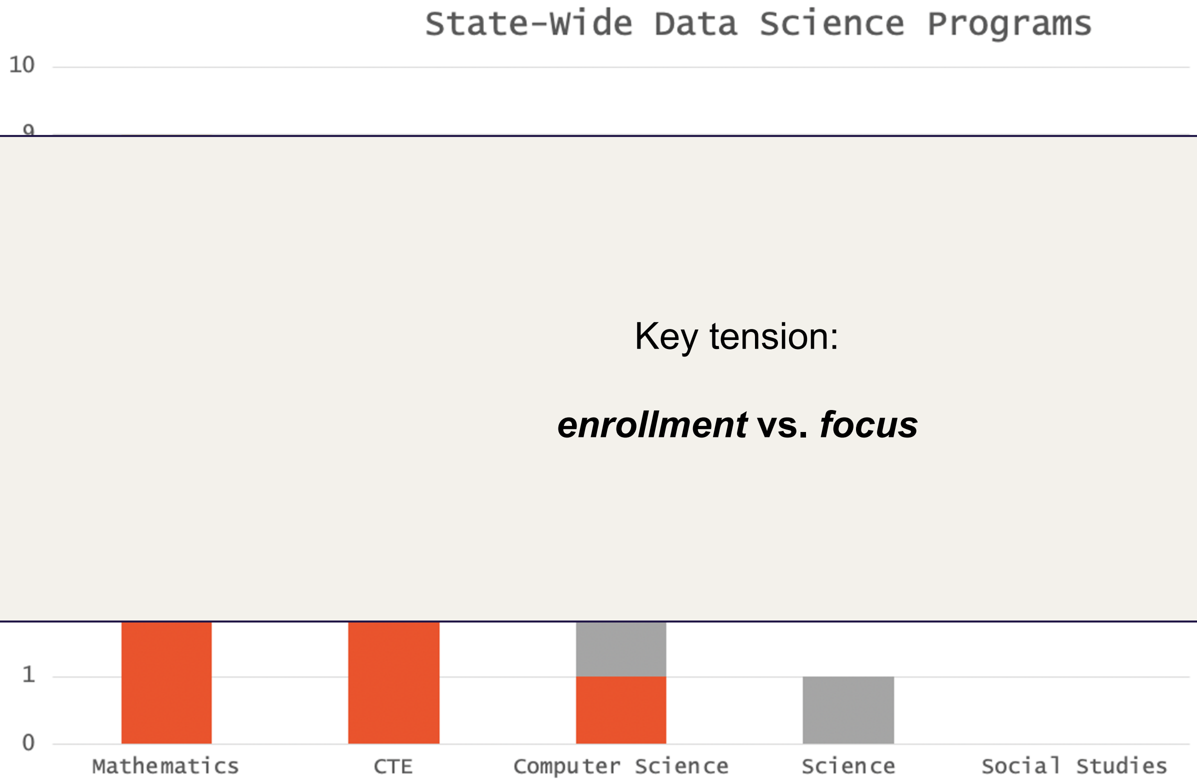
- **Data ethics is missing:** existing standards only highlight data privacy;
  - Little mention of bias in datasets, ML models, and the resulting algorithms
  - Little mention of “fact-based authority” of data and ways it can be manipulated
- **Methods are too simplistic**: we need modern methods and processes > finite list of skills

A variety of models & learning pathways:



NATIONAL ACADEMIES Sciences Engineering Medicine

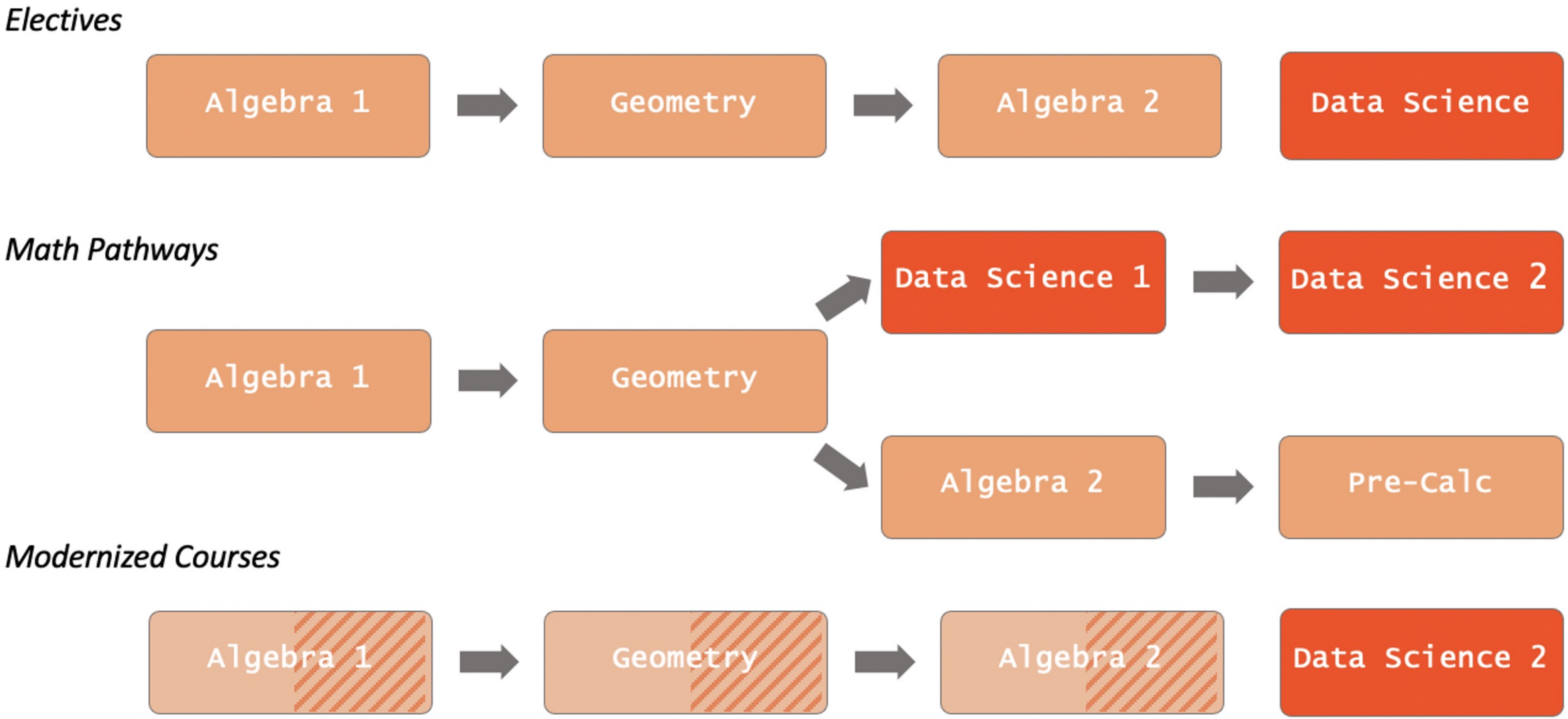
A variety of models & learning pathways:



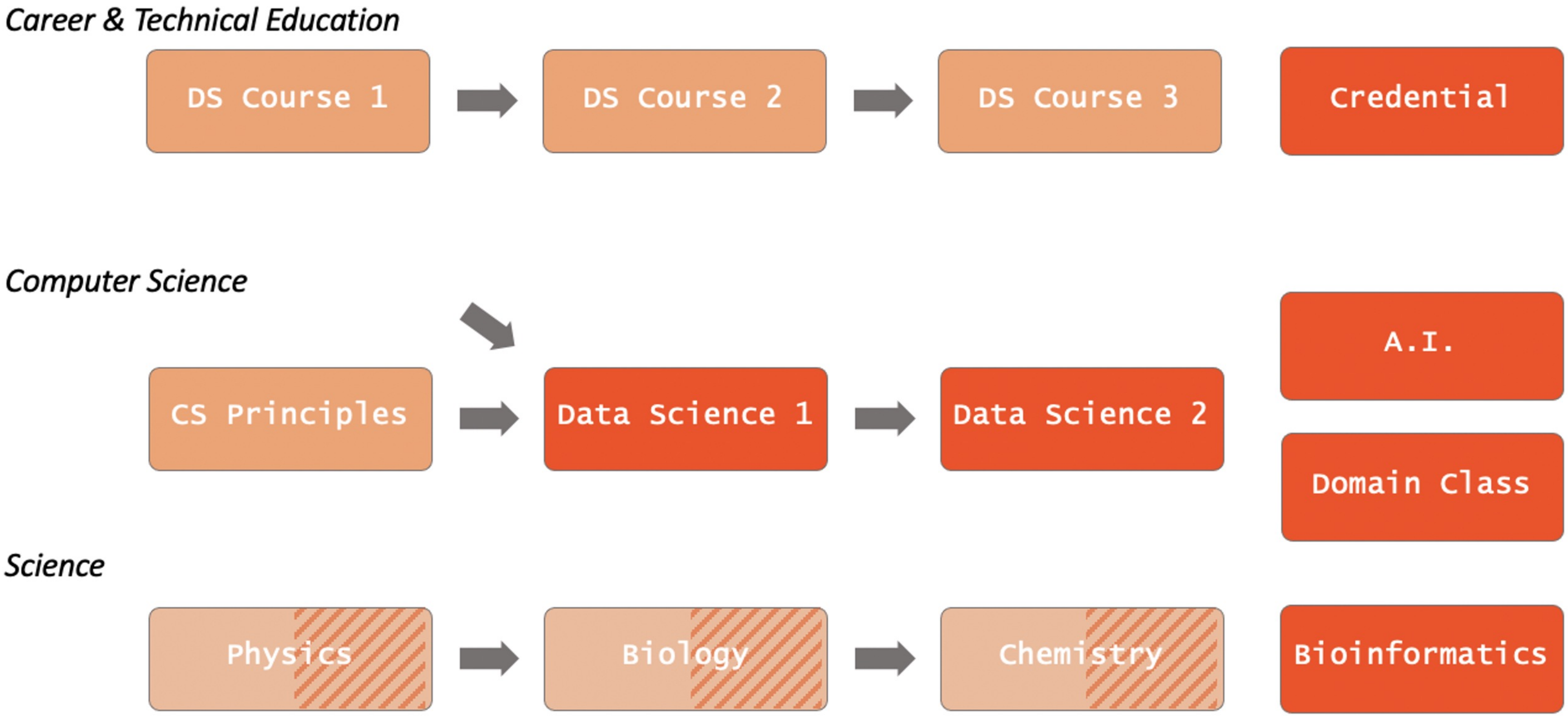
Key tension:

enrollment* vs. *focus

A variety of models: mathematics



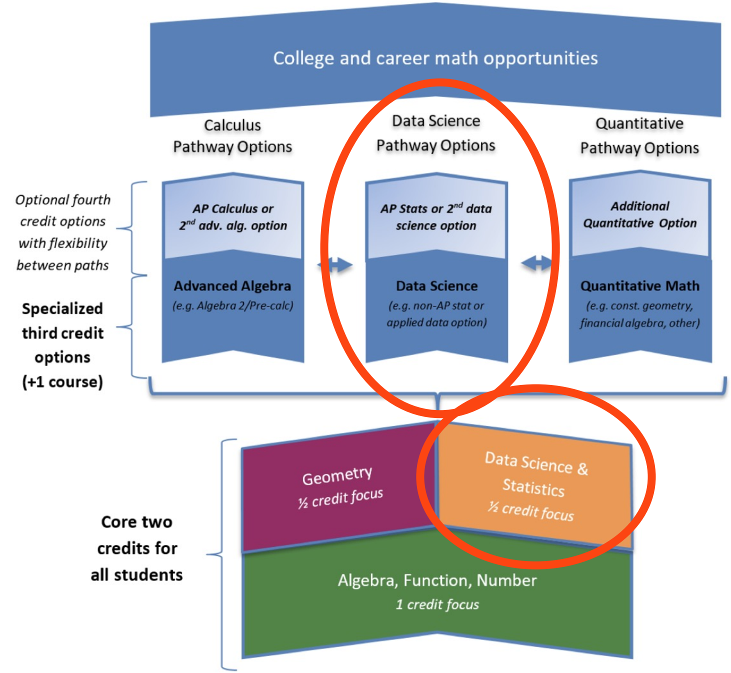
A variety of models: other K-12 subjects

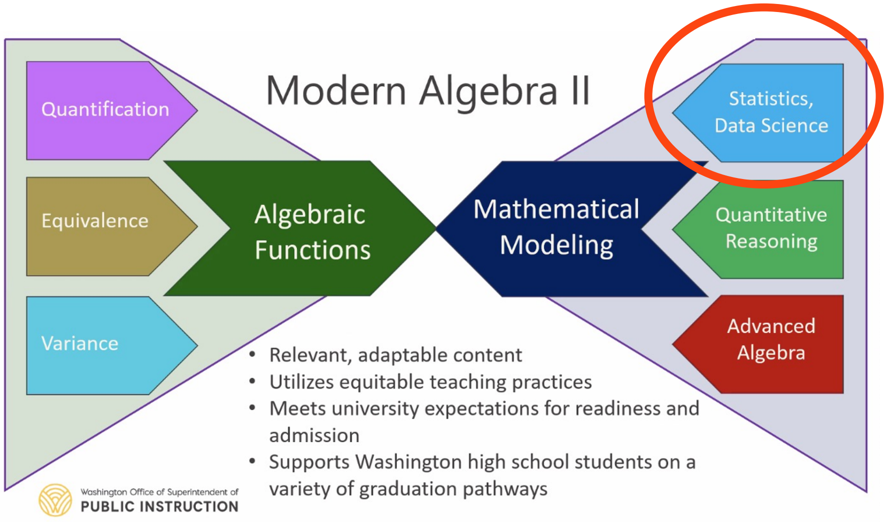


NATIONAL ACADEMIES Sciences Engineering Medicine

A variety of models: state examples

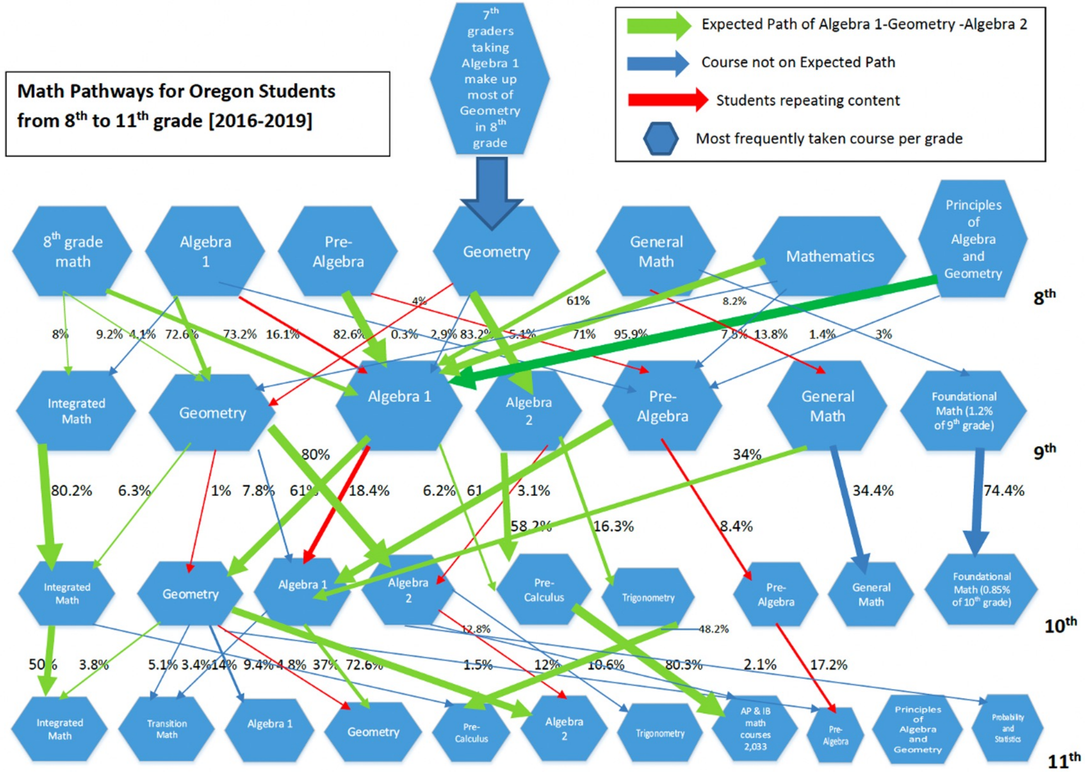
Figure 1 - Diagram of the 2 + 1 Model showing one possible sequence for the first three credits in high school mathematics





Is there really one model?

- Even in a standard “geometry sandwich,” there is significant diversity.
- Increasing **student, family options** is likely productive
- Increasing **integration in K-8** is likely productive - how do we **make middle school a launchpad > a filter?**



NATIONAL ACADEMIES Sciences Engineering Medicine

Hearing from practitioners:

Suyen Machado
IDS / UCLA CenterX (CA)

> Computer Science / Stats Ed.
> Curriculum design

Stephanie Melville
San Diego Unified (CA)

> Mathematics
> Teacher / Resource Teacher

Paul Strode
Fairview HS (CO)

> Biology
> Teacher / Teacher Educator

Katie Headrick Taylor
University of Washington (WA)

> Geospatial Data
> Researcher

Panel Discussion

How is Data Science Integrated in Content Areas?

03

Examine the ways in which data science has been integrated with other subjects beyond mathematics and implementation across settings

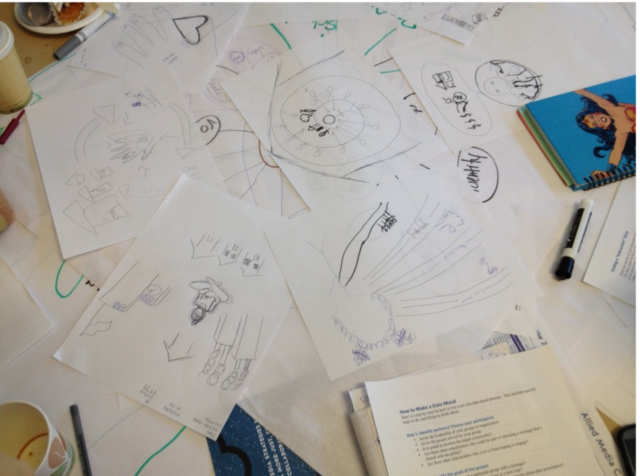
SESSION GOALS

Panelists

Moderator: Camillia Matuk, New York University (virtual)

Rahul Bhargava	Northeastern University
Angela Calabrese Barton	University of Michigan
Josh Radinsky	University of Illinois at Chicago
Emmanuel Schanzer	Bootstrap
Lissa Soep (virtual)	Vox Media, LLC

Rahul Bhargava



What, and who, is data science for?

As data becomes more central to civic and community decision making, we need to broaden how we invite people to the table around data.

- build critical data practices
- embrace epistemological pluralism
- learn arts-based approaches
- social impacts

Angela Calabrese Barton

Data as Sites of Youth Science-Justice Work

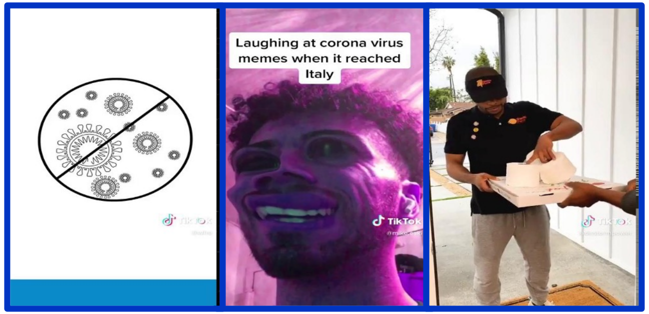
- “I don't see myself in the data.”
- “It’s not numbers that will solve the pandemic, but the stories people tell with numbers. . . If we don’t have different perspectives on the numbers, it will only offer one story. That won’t help everyone.”
- “It’s never only about COVID19.”

- DRL #2028370 - *How People Learn Rapidly: COVID-19 as a Crisis of Socioscientific Understanding and Educational Equity*
- DRL# 1502755 - *Tools for Teaching and Learning Engineering Practices: Pathways Towards Productive Identity Development in Engineering*

Data as Sites of Youth Science-Justice Work

How and why do young people engage with data to make sense of, make decisions about and take action on science-justice related concerns in their everyday lives and communities?

- Datafication does not impact all people equally
- Learning and engaging with/about data involves more than cognitive processes
- Epistemic (in)justice



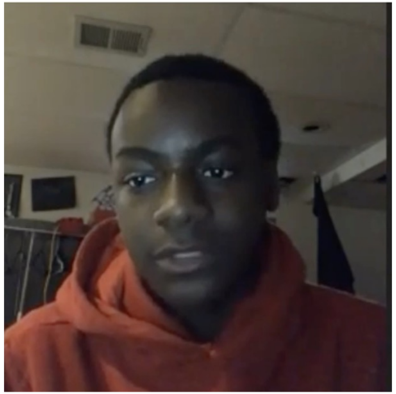
What counts as data and who decides? What gets datafied, by whom and in what ways? Who benefits and gets hurt in the process?

Data as Sites of Youth Science-Justice Work

UNIVERSITY OF MICHIGAN

To enact data justice in learning environments requires attention to how young people engage with and use data towards affecting their own and others' lives, social relations and possibilities.

- Critical Data Practices
- Alternative Data Infrastructures
- Data Agency



"On the CDC website, I looked at its spread-per-day worldwide. . . **What I figured out really scared me.** The CDC website and the WHO, those had the most amount of information I really need on the topic. [B]ut one of the things that really helped me **get over my anxiety, over like the entirety of COVID, was watching YouTube videos of YouTubers** saying to calm down, just, like, wear your mask, and be safe, and make sure you don't go out a lot." Prez

Data as Sites of Youth Science-Justice Work

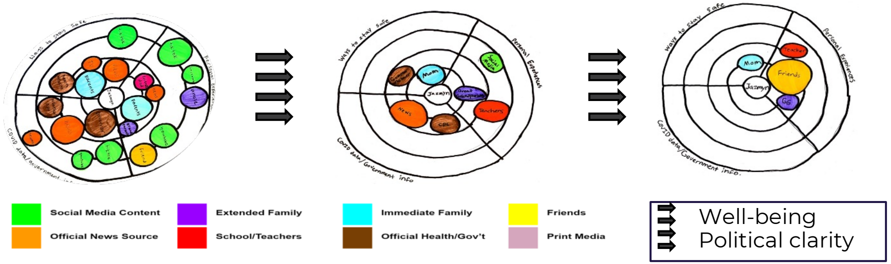
UNIVERSITY OF MICHIGAN

To enact data justice in learning environments requires attention to how young people engage with and use data towards affecting their own and others' lives, social relations and possibilities.

- Critical Data Practices
- Alternative Data Infrastructures
- Data Agency

- Individual/collective engagement
- Acting with and on data/infrastructures
- Encounter & counter dominant data practices & narratives
- Powerful forces for social transformation and justice

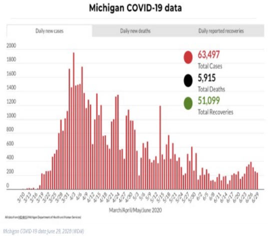
Youth-built infrastructure maps

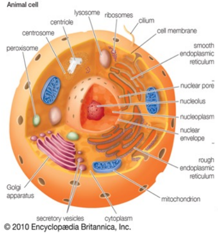


NATIONAL ACADEMIES Sciences Engineering Medicine

Data as Sites of Youth Science-Justice Work

> “I had to decide whether to protect myself and my family against injustice by protesting, or to protect myself and my family by not going.”

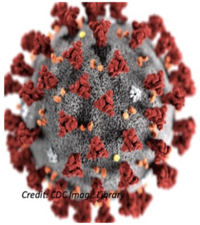




Cases by Race

Race	Percentage of Overall Cases by Race	Percentage of Deceased Cases by Race
American Indian or Alaska Native	<1%	0%
Asian/Pacific Islander	1%	1%
Black or African American	34%	40%
Caucasian	24%	29%
Multiple Races	2%	1%
Other	3%	2%
Unknown	36%	26%

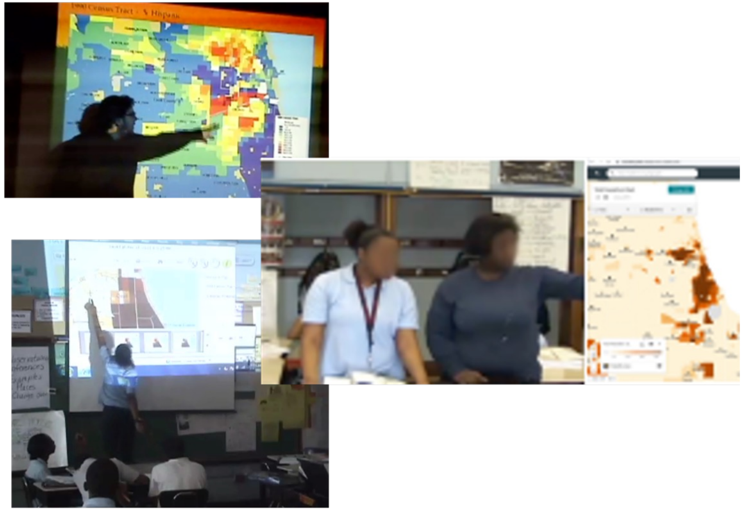
Credit Michigan Department Of Health And Human Services / Https://Www.michigan.gov/Coronavirus/0,9753,7-406-98163_98173---,00.Html / Https://Www.michigan.gov/Coronavirus/0,9753,7-406-98163_98173---,00.Html

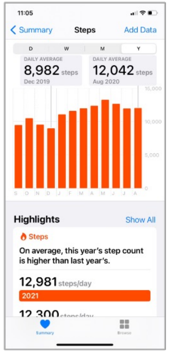


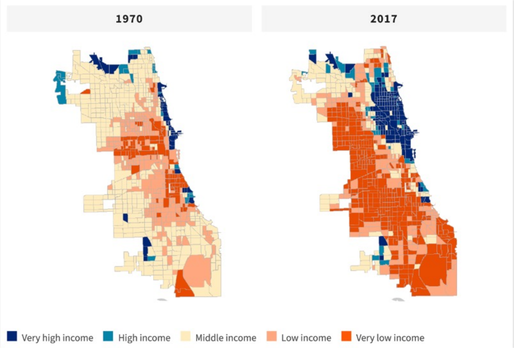
COVID-19 data unjustly overlooked challenging realities J, her peers + broader community navigated on a daily basis.

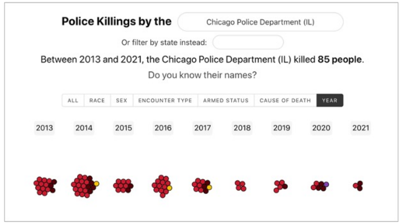
- **Political Clarity:** Repositioned self into new aggregations of data by researching viral transmission, examining mask wearing on protest images, searching infection rate/spread patterns, and following Black female doctors on Tik Tok for information on safety
- **Well-being:** Built new data infrastructures to share learning across a curated coalition of young Black women, including viral mitigation protocols + mental wellness protocols (e.g., daily vlogs uploaded to Snapchat group + private text groups)

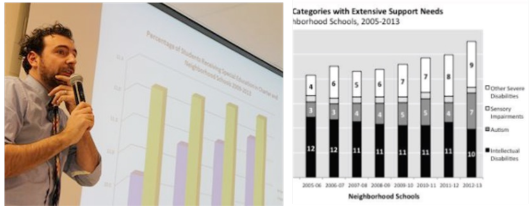
Josh Radinsky

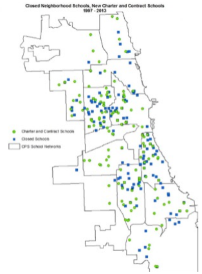




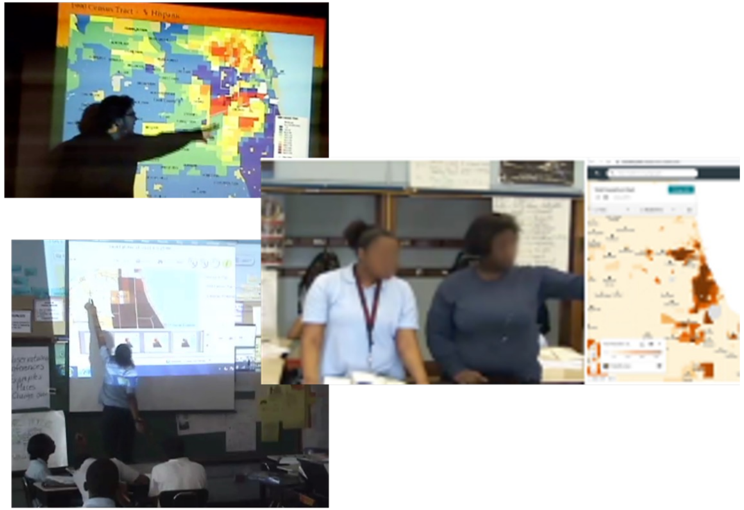








Josh Radinsky



Creating, enacting, improving & exploring curriculum and instruction

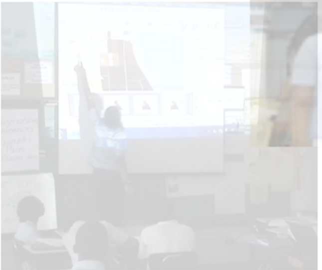
using data interactives

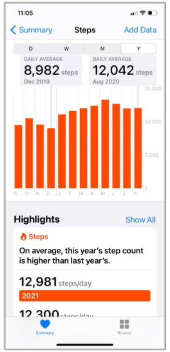
in social science classrooms

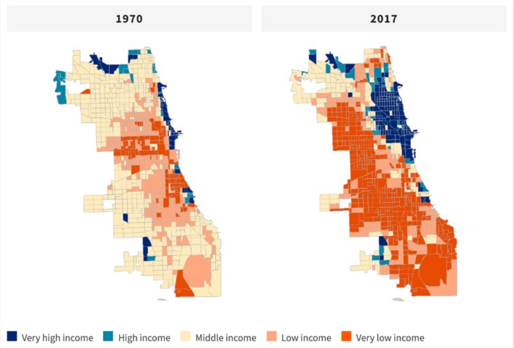
middle school through college

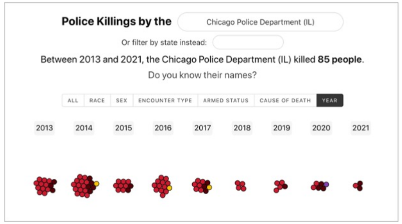
Josh Radinsky

Studying how we narrate our social worlds (especially Chicago) using data texts and tools



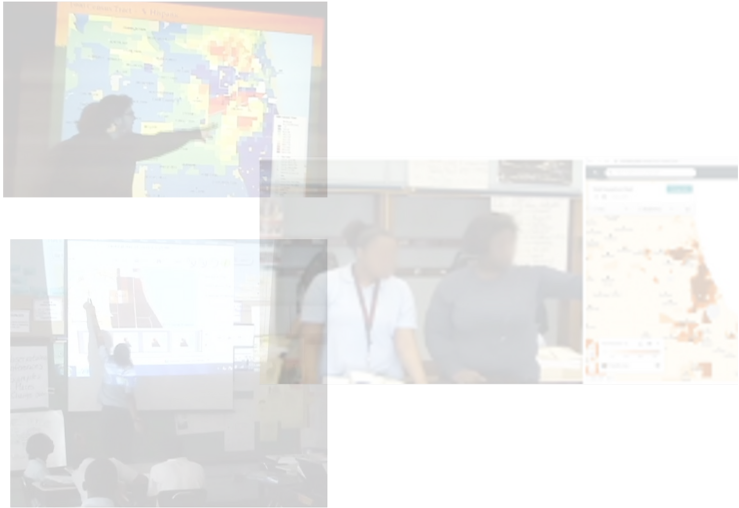




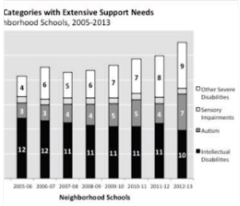


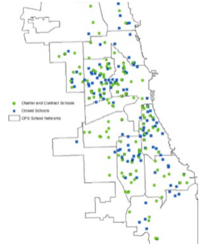
NATIONAL ACADEMIES Sciences Engineering Medicine

Josh Radinsky

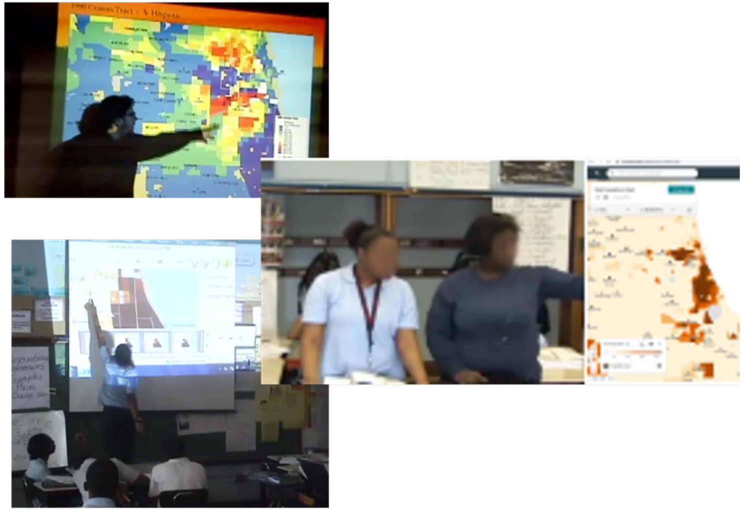


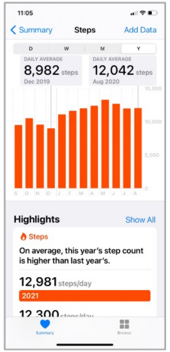
Engaging in the ongoing struggle for educational justice in Chicago – with data

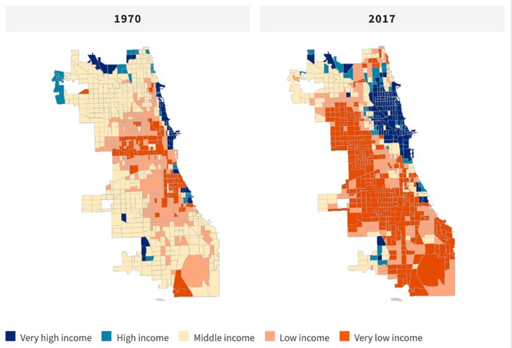


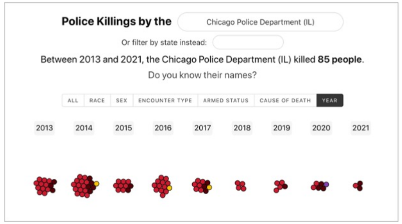


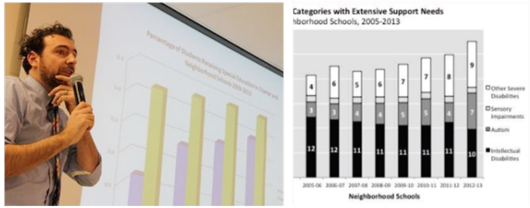
Josh Radinsky

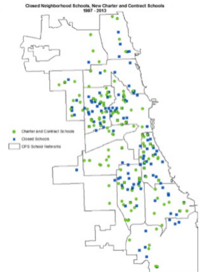












Data texts and tools are central to how we learn about the social world, and change it

- ***Narrating data*** should be understood as a way we engage with and change the world, not just a secret code for describing it
  - Data, tools, narratives - and the worlds they represent - are always “on the move”
- As teachers, we need to develop our ***social science learning objectives*** – clarify the skills, concepts, dispositions and information with which we make sense of society, history, ourselves, and one another
  - Promising practice: using *Design Patterns* (Mislevy) within an Evidence Centered Design approach to creating assessments
- We need to teach and learn with data as part of (not separate from) our human and more-than-human relationships - “***relational data literacy***,” and a relational understanding of data science.
  - We teach, learn, and do our data practices within relationships

Emmanuel Schanzer | Bootstrap

Data Science is a balance of four key ingredients - don’t leave one out!

Statistics

Civic Responsibility

Computing

Domain Investment

Emmanuel Schanzer | Bootstrap

Use all 4 ingredients, choose tools wisely, & support existing standards

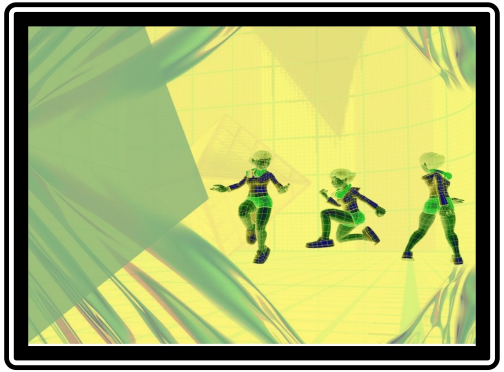
- **Mathematics**
  - Working with multiple states on *Data Science-infused curriculum for Algebra 2*
- **Physics**
  - Partnered with NSF/AAPT to create DS/Modeling materials for 9th grade Physics
- **History and Social Studies**
  - Developed elementary grade history and social studies materials for KIPP network
- **Biology, Chemistry, Earth Science, etc.**
  - Exploring domain-specific DS activities (tools and standards!)

Data Science for Storytelling and Social Justice: Insights from a Youth Newsroom

Lissa Soep

Surveillance U

Erase Your Face

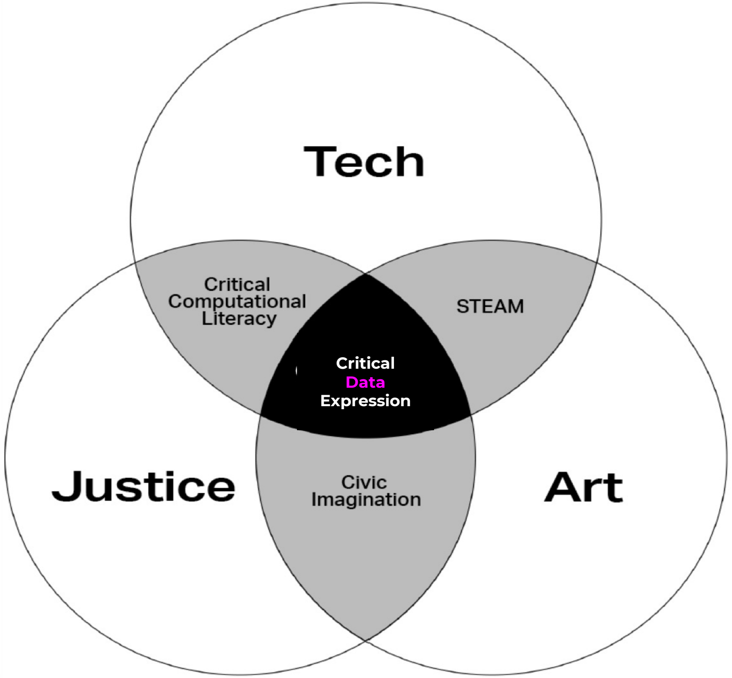


Can You Teach AI to Dance?

Frameworks to think and make with ...

What are some conceptual approaches that embed data learning into justice-driven storytelling with youth that reaches peers and the public sphere?

- Data as “Storytelling Matter”
- Critical Data Expression (w/ Lee)
- “Humanizing Collegial Pedagogy” (w/ Clark + McBride)
- Digital afterlife



Adapted from Lee and Soep, Code for What? 2023

"Through the digital media products they create, young people can reimagine what's possible in their lives and futures, and work with others to pursue that vision for change. Critical [Data] Expression synthesizes vibrant traditions from cultural studies, STEM, and the arts, honoring the interdisciplinarity and deep humanity of a society that depends on, but cannot be reduced to, its near-constant use of technology."

(adapted from Lee+Soep, 2023)

Lunch

04

Panel Discussion
What is the State of Educator Preparation in Data Science?

05

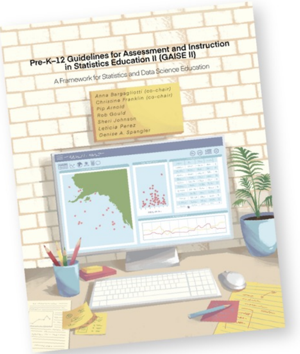
Examine issues on educators' teaching with data and the preparation needed to teach statistics/data science/computation for prospective and practicing educators

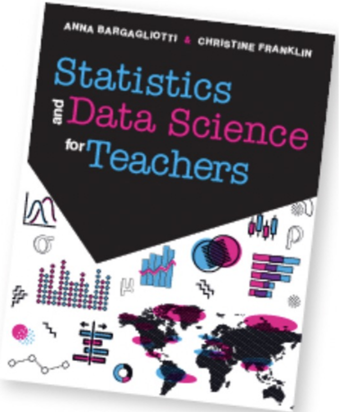
SESSION GOALS

Anna Bargagliotti

Professor, Department of Mathematics, Loyola Marymount University

- Author of GAISE II report
- Author of *Statistical Education of Teachers* (SET) report
- Author of *Statistics and Data Science for Teachers* book
- Current Projects
  - PI on *Undergraduate Data Pathways* NSF- funded
  - PI on E*quity in Computer Science* NSF - funded
  - PI on *Examining Disparities in Mathematics Achievement to Promote Educational Equity* funded by the John Randolph Haynes Foundation





Stephanie Casey

Professor of Mathematics Education

Eastern Michigan University, Ypsilanti MI

- High school mathematics teacher for 14 years at Deerfield HS, Deerfield IL
  - Started teaching AP Statistics in 1998, led Chicago-area AP Statistics teachers network
- Became a mathematics teacher educator specializing in statistics education in 2011
- Authored statistics teacher education curriculum materials for two recent NSF-funded projects

ESTEEM

Association of Public & Land-Grant Universities

NSF

MODULE(S²)

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

The University of Arizona · Eastern Michigan University · Middle Tennessee State University · University of Nebraska Lincoln · Utah State University

The Mathematics Of Doing, Understand, Learning, and Educating Secondary Schools (MODULE(S2)) project is made possible through funding from the National Science Foundation IUSE (Improving Undergraduate STEM Education) multi-institutional collaborative grant #1726707 (APLU), #1726098 (University of Arizona), #1726252 (Eastern Michigan University), #1726723 (Middle Tennessee State University), #1726744 (University of Nebraska - Lincoln), and #1726804(Utah State University).

Anne Leftwich

Barbara B. Jacobs Chair in Education and Technology, School of Education
Adjunct Professor, Computer Science
Indiana University

Current Work: Co-Design with Teachers

- Supervise and Teach Preservice Educational Technology Course (K-12)
- K-2 integration of CT into Literacy Activities: Rethinking Circle Time
- 3rd-5th grade integration of AI into life sciences: PrimaryAI
- 6th-8th grade AI education in rural schools: AI Goes Rural
- Expanding computer education pathways to broaden participation in K-12 computing: ECEP

Previous Work: Teacher beliefs and teacher technology integration

- Teacher technology beliefs; Teacher technology integration education; Teacher adoption and diffusion

Gemma Mojica

Research Scholar, Friday Institute, NC State University

- Former middle & high school teacher
- Mathematics and statistics teacher educator
- Co-director of a Hub for Innovation and Research in Statistics Education (Hi-RISE)
  - Multiple funded initiatives to build foundations for innovation and research in statistics education
- Research Focus
  - design, implement & research various models of professional learning
    - small settings & at scale
    - online curricular materials
    - technology tools & online learning platforms for teachers
- Current Funded Projects
  - *Invigorating Statistics Teacher Education* (InSTEP) [NSF DRL 1908760)
  - *Enhancing Statistics Teacher Education with E-Modules* (ESTEEM) [NSF DUE 162571

InSTEP
Invigorating Statistics and Data Science
Teaching through Professional Learning

ESTEEM

Leticia Perez

WestEd - Boosting Data Fluency Development Lead

Former Work:

- STEM+C3 Integration of Data Science and Computational Thinking into UCLA’s Math and Science Teacher Education Program
- Designed and refined 4 course Computer Science credentialing pathway for Math and Science Teachers in the State of California.
- Computational Thinking for Equity Framework
- GAISE II
- Former High School Science Teacher and Curriculum writer

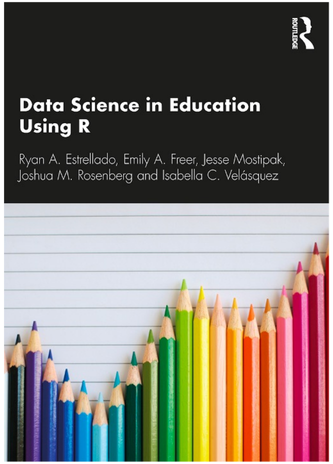
Current Work:

- Co-PI Boosting Data Fluency NSF Funded Project to develop a PL model that integrate data rich instruction into learning spaces of middle and high school aged students.

Joshua Rosenberg

Assistant Professor, STEM Education, and Faculty Fellow, Center for Enhancing Education in Mathematics and Sciences University of Tennessee, Knoxville

- Former high school science teacher, current science teacher educator
  - Also instructor of educational data science courses for graduate students
- Current funded projects:
  - *Advancing Bayesian methods for pre-collegiate learners* (part of a larger CS education-focused project)
  - *CS for Appalachia*
  - *The Learning Analytics in STEM Education Research (LASER) Institute*
  - *Advancing computational grounded theory for audiovisual data*
- Co-Author of *Data Science in Education Using R* (https://datascienceineducation.com)



Panel Discussion
What is the State of Educator Preparation in Data Science?

05

Townhall

06

How can practice inform current and future research needs?

SESSION GOALS

Funder Reflection

07

California Education Learning Lab

GRAND CHALLENGE: BUILDING CRITICAL MASS FOR DATA SCIENCE

A GRANT OPPORTUNITY OF THE CALIFORNIA EDUCATION LEARNING LAB

LEARNING LAB

Award Types	4
Number of Awards	Up to 18
Total Award Budget	$6.3 - $8 million
Duration of Grant	2-4 years

PATHWAYS DEVELOPMENT

~$1.3 million

Duration over 3-4 years

Up to 3 awards

FACULTY DEVELOPMENT

~$200K to ~$350K

Duration over 2-3 years

Up to 5 awards

INTERDISCIPLINARY COLLABORATION

~$100K to ~$200K

Duration over 2-3 years

Up to 9 awards

GRAND CHALLENGE COHORT COODINATOR

Up to $500K

Up to 5 years

1 award

California Education Learning Lab

https://calearninglab.org/grant/data-science-rfp/

Email: info@calearninglab.org

Final Reflections from Planning Committee

08

Reflections from Planning Committee members

- Hollylynne Lee
- Tammy Clegg
- Tim Erickson
- Zarek Drozda
- Camillia Matuk

Reflections from Nick

- Tremendous excitement and energy: opportunity to make a difference
- "Pyramid" of data science (vs. "ragged mountain" metaphor): what are the implications (and connections) for graduate data science? Undergraduate data science? Associate's programs? Experiences from NSF-funded DSC-WAV project
- Data literacy vs. data fluency vs. data acumen: supporting critical thinking grounded in data analysis cycle (and answering a question) in a developmentally sound manner
- Teacher prep remains a major challenge (+ transdisciplinary component)
- Building bridges and partnerships: hard work, good start, much to do to foster connections in a sustainable fashion
- Engaging with substantive questions, data realities and critical data literacy: unclear about how to make this real?
- Broadening participation and ensuring equitable access

Reflections from Nick

https://par.nsf.gov/biblio/10075971-keeping-data-science-broad-negotiating-digital-data-divide-among-higher-education-institutions

Consequently, as data-driven decision-making becomes more commonplace, having the skills to understand and make sense of data can provide a sense of power to the larger citizenry or conversely powerlessness to communities without these skills. This “Data Divide” separates communities that have access to devices and services that provide rich, data-driven services from those that don’t; it separates data-savvy individuals, and communities that have understanding and awareness of how their data is being collected and used to provide individualized services (and thus informed protections), from those that do not. The economic and social consequences of the Data Divide stratify populations, and severely limit the opportunities of those who are unable to take advantage of the data revolution (page 7)

Reflections from Nick

https://par.nsf.gov/biblio/10075971-keeping-data-science-broad-negotiating-digital-data-divide-among-higher-education-institutions

If we do not make diversity and inclusion a priority now, we will not have it in the future. We do not want to repeat the mistakes of the past, so we must reverse the trend for the growing divide to make and keep data science broad. Diversity will bring a lot of ideas and voices to the table, which may lead to significantly fewer models producing biased results when trained using algorithms on biased data sets. (page 30).

Reflections from Michelle

- There is not just one kind of “data expertise.”
  - Data Science PhD/ Google
  - Data Journalist; Computational Biologist/Geoscientist/…; Art that critiques (and critiques through methodology) datafication
  - Possible futures with/for data science (given different questions)
- What counts as success (and failure) should be framed carefully and with nuance
  - Data advocacy, impact, scalability…
- COLLABORATION
  - Across teachers, across levels of education/administration; with community members, librarians, practicing scientists, journalists, artists, families, youth
  - Incentives and supports for collaboration

SEPTEMBER 2022

Day 2
Wrap up and Adjournment

09

The recording and accompanying materials will be available on the project page:

https://www.nationalacademies.org/our-work/foundations-of-data-science-for-students-in-grades-k-12-a-workshop