

STATISTICS AND THE EMERGING DISCIPLINE OF DATA SCIENCE



ADAM SULLIVAN
BROWN UNIVERSITY



JESSICA UTTS
UNIVERSITY OF CALIFORNIA, IRVINE



ANDREW ZIEFFLER
UNIVERSITY OF MINNESOTA



BEN WENDER
THE NATIONAL ACADEMIES OF
SCIENCES, ENGINEERING, AND
MEDICINE

DATA SCIENCE EDUCATION ROUNDTABLE

Project statement of task:

The Roundtable on Data Science Post-Secondary Education will convene critical stakeholders from data science training programs, funding agencies, societies, foundations, and industry to discuss data science education and practice, needs of the community and employers, and ways to move forward. The Roundtable will convene 4 times per year and create a venue for exchange of ideas and a mechanism for joint strategic planning among key stakeholders and experts in data science and education.

DATA SCIENCE EDUCATION ROUNDTABLE

Engaging through structured meetings and targeted dissemination

Each Roundtable meeting has a theme and targets slightly different audiences.

Core audience: Industry, government, foundations, and academic groups doing work related to data science post-secondary education.

Goal: Improve coordination among stakeholder and increase awareness of current and future efforts in the broader data science community.

Integrating societal and ethical concerns into data science education

December 8, 2017
(Washington, DC)

Data science education mechanisms: Matching educational approaches for the audience

October 20, 2017
(Chicago, IL)

You are here!

Videos, materials, and summary available online
<http://www.nas.edu/statistics>

Tailoring courses and training to diverse data communities

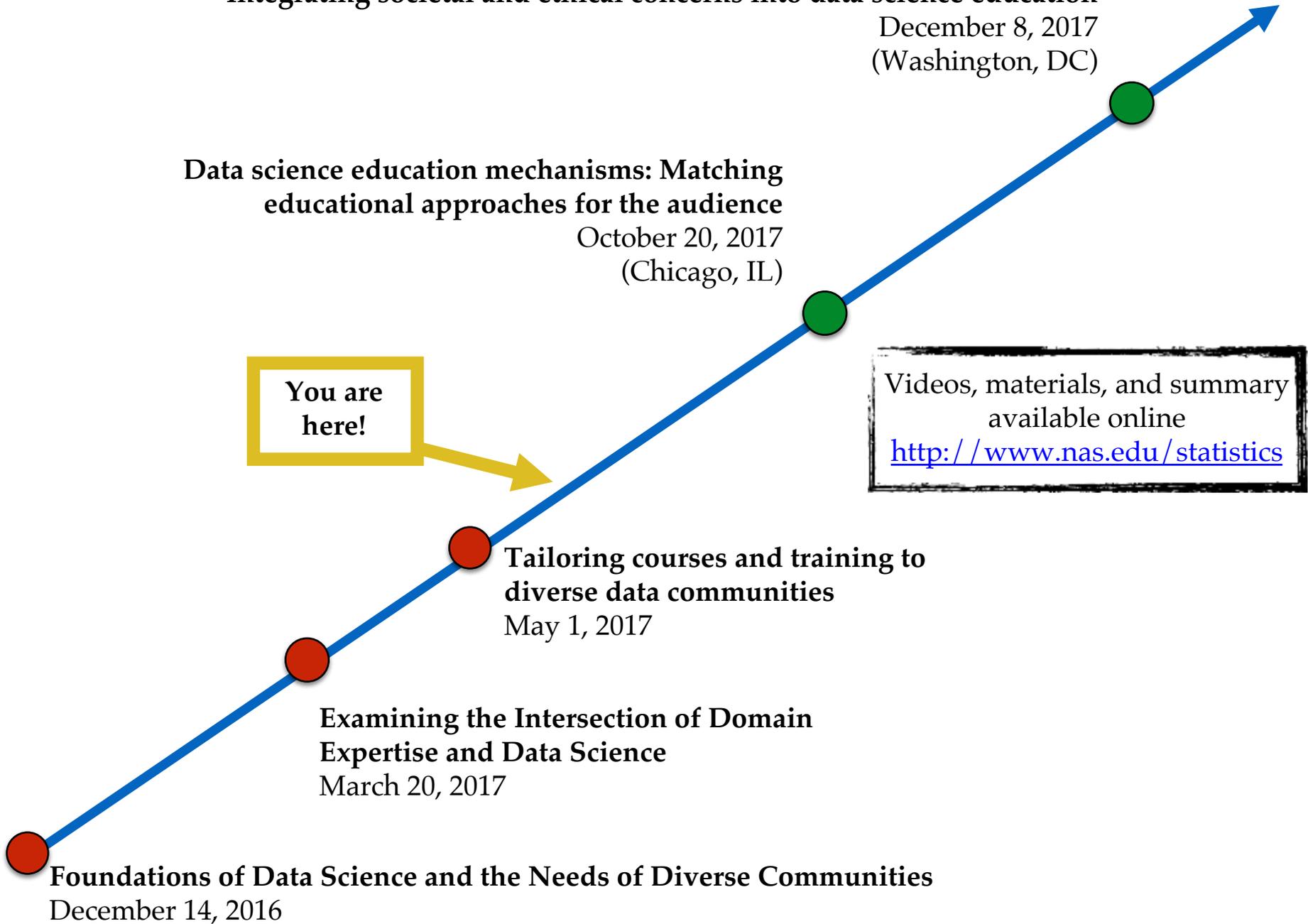
May 1, 2017

Examining the Intersection of Domain Expertise and Data Science

March 20, 2017

Foundations of Data Science and the Needs of Diverse Communities

December 14, 2016



ENVISIONING THE DATA SCIENCE DISCIPLINE: THE UNDERGRADUATE PERSPECTIVE

Carry out a **consensus study** to

- Identify core underlying principles, intellectual content, and pedagogical issues specific to data science, including core concepts that distinguish it from neighboring disciplines.
- Focus on undergraduate level, but address topics related to middle and high schools, community colleges, and draw on experiences of Master's-level programs.
- Consider opportunities to engage underrepresented student populations.



ULTIMATE GOAL TODAY

Hear about each others' opinions and experiences

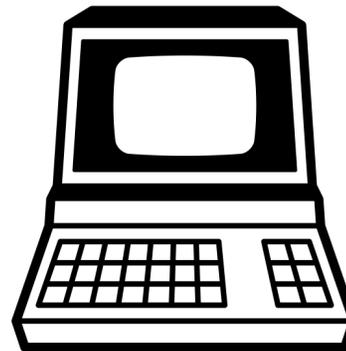
Identify common themes *and* differences across different groups

Explore and inform development of data science syllabi / courses and curriculum.

SMALL-GROUP ACTIVITY

In your small-groups from the warm-up, take **10 minutes** to discuss the questions on the next slide.

Have one person volunteer to take notes and send these to Ben after your discussion.



THOUGHT-PROVOKING DISCUSSION QUESTIONS

- What is data science?
- Is data science a "stand alone" discipline that should have its own undergraduate major?
- How should data science be taught in data-intensive domain fields such as astronomy, humanities, and biology?
- Does statistics education need to change in the context of data science and how?

DATA SCIENCE CONTENT

DISCUSSION QUESTIONS

20 MINUTES

- What topics are essential to include from statistics? Not just specific content, but what principles? How should this material be structured?
- Which Computer Science principles are essential to include in data science?
- Which (*fill-in-the-blank*) principles are essential to include in data science?
- What order should content be introduced (e.g., sequencing) and in what context?
- What infrastructure is necessary? (Not just hardware, but also which people and expertise?)
- Where should data science programs be housed?
- How do we keep data science from having the same gender and racial gaps as other STEM disciplines?
- How do we maintain a system of collaboration within the curriculum and design of the program?

HOW TO PARTICIPATE

Data Science Roundtable

To receive updates, please join the *Committee on Applied and Theoretical Statistics* (CATS) mailing list at: <http://www.nas.edu/statistics>

Envisioning the Data Science Discipline

Contact Ben Wender at: bwender@nas.edu

REFERENCES AND ATTRIBUTIONS

- *building* by Yo! Baba from the Noun Project
- *retro computer* by Tinashe Mugayi from the Noun Project

- Eric Kolaczyk for the Data Science Roundtable slides
- Laura Haas for the Envisioning the Data Science Discipline slides