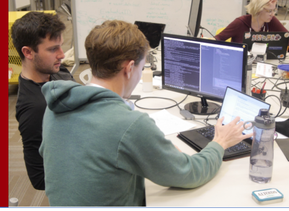
Improving data science education

Ben Stenhaug, PhD education student & data science scholar at Stanford University



Data science education is at a defining moment

Essential Question: How can we get more students access to better data science learning opportunities?

" University STEM teaching is a lot like medicine was in the 1800s. For any students in the room, I just want to say you're basically at the tail end of getting the pedagogical equivalent of bloodletting." **— Carl Wieman, Stanford University**

What we know: Key instructional components

1. Effective, empathetic instructor
2. Measured learning outcomes
3. Clear explanation of new concepts
4. Good Tasks
5. Active Learning in groups
6. Thoughtful Feedback
7. Growth Mindset Culture

Our task: To implement in data science education

Key Questions:

1. What do these instructional components look like in DS Edu?
2. How can we prepare teachers to execute?

" Concept inventories started a revolution in Physics education. Can we replicate this in computer science? We're sure as hell going to try." ***— Craig Zilles, University of Illinois***

Working towards a solution: Research ideas

Helping teachers develop pedagogical content knowledge (PCK)

Question: Teachers with more PCK will make better decisions. How can we help teachers develop PCK?

One example: I conducted talk aloud protocols.

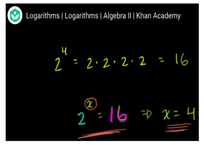
```
hamsters_data_frame
A tibble: 6 x 3
name   what      count
<chr>  <chr>     <dbl>
Amy    hamsters      2
Amy    cages         4
Amy    wheels        1
Bianca hamsters      8
Bianca cages         3
Bianca wheels        2
```

```
spread(hamsters_data_frame, what, count)
A tibble: 2 x 4
name   cages hamsters wheels
<chr>  <dbl>    <dbl>  <dbl>
Amy        4        2      1
Bianca     3        8      2
```

Experimenting with explanation

Question: How can we find the best way to explain something?

One example: Researchers crowdsourced videos explaining logs. 80% were accurate and the top videos were as effective as a Khan Academy video.



Data Science for Social Good programs: The future of data science training?

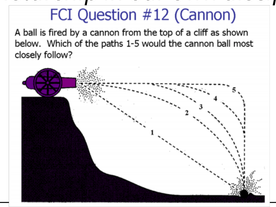
Question: How can we form teams of students learn effectively while solving important problems?

One example: Data Science for Social Good Programs

Data Science For Social Good
Summer Fellowship
THE UNIVERSITY OF CHICAGO

Defining outcomes and measuring learning

"We should apply the same rigor to assessment of student learning that we do to research and scholarship in our own disciplines" — Teresa Sullivan



30 years of computer science education research

Question: What can we learn from CS Edu research?

Examples:
Programming should be taught as "an exploratory process where programs are created opportunistically and incrementally" (Green, 1990)

Good novices are movers instead of stoppers. When struggling, movers keep trying by experimenting and modifying their code. (Perkins, 1989)

Coding to learn: R in K-12 education

Question: How can we create a synergy between learning data science in R and mastering more traditional K-12 objectives?

One example: There are a few high school teachers using R in their classrooms and they're winging it!

"(Teaching) involves using skill, love, and knowledge to maximize deliberately the probability that students will learn worthwhile things and will flourish as human beings from being in that learning environment."
— Deborah Loewenberg Ball, University of Michigan