Perspectives on Data Science Education at the School Level

William Finzer, Senior Scientist, Concord Consortium
Tim Erickson, Epistemological Engineering

March 21, 2018

If you can't chat, text Talya at (617) 821-5374

How can we give secondary-school students an authentic feel for doing data science? One aspect of this question is to ask where they should do data science—in a dedicated class? In math? In science? In history? Another is to ask for examples of what they can realistically do. What characterizes data science, and what kinds of activities are suitably accessible, complex, and don't require coding? In this webinar we'll use the free, browser-based Common Online Data Analysis Platform (CODAP) to illustrate our thinking on these questions. We'll also draw from our current Data Science Games project in which game-like plugins to CODAP generate rich and surprisingly interesting data.

What has changed?

Before 2000

Data are
- $$ Expensive
- In short supply
- Relatively inaccessible
- Relatively small datasets

Data skills are
- Important
- Concentrated in a few professions

After 2000

Data are
- ¢¢ Inexpensive
- Plentiful
- Accessible
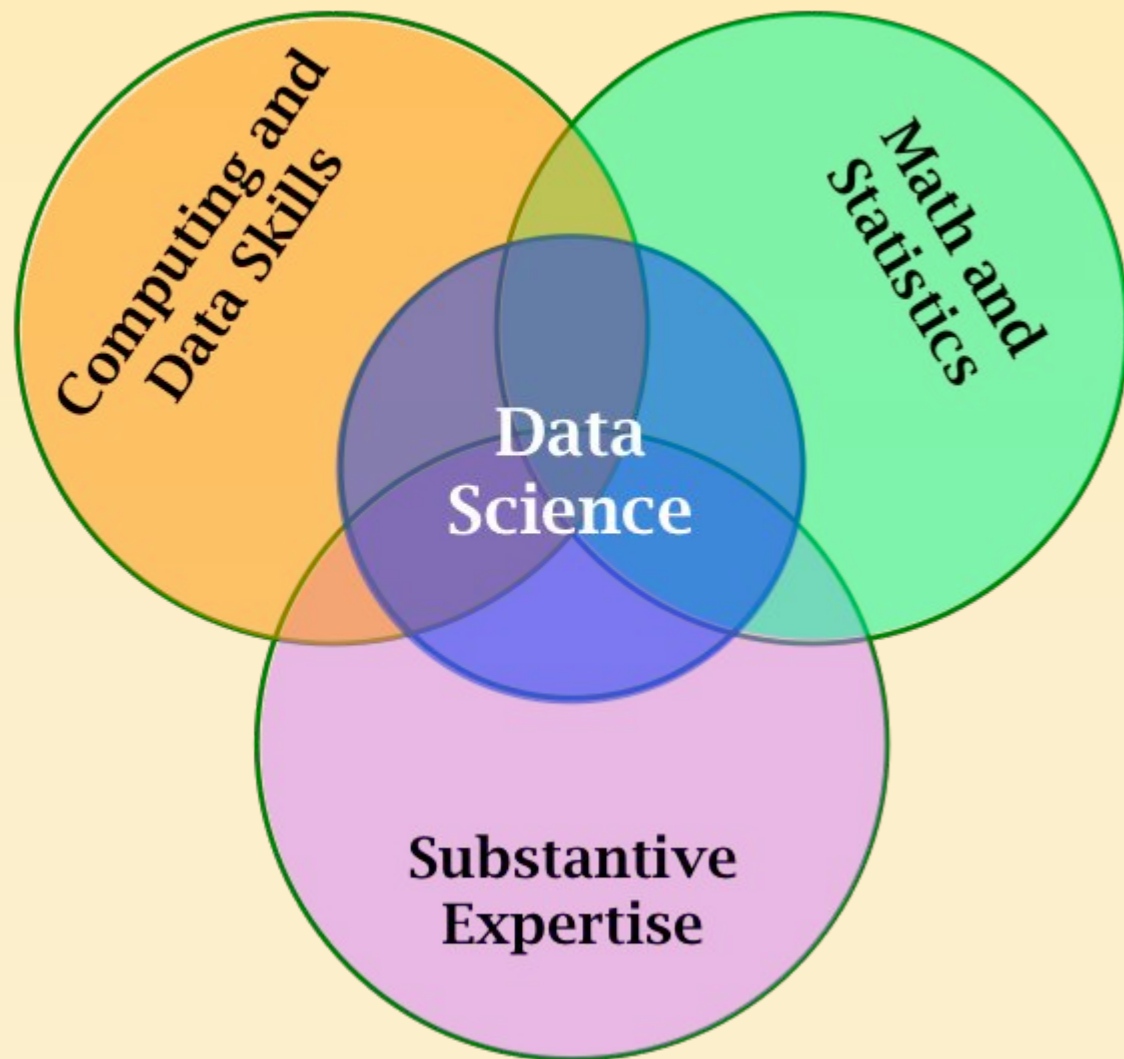- Ubiquitous
- Huge datasets

Data skills are
- Essential
- Required in every endeavor

⇒ Data Science
⇒ Data Science Education

What is

Data Science?

Computing and Data Skills

Math and Statistics

Data Science

Substantive Expertise

Data Science Education?

- Bring about an appropriate level of data science fluency
- Develop technology suited to learning with and about data
- Develop curricula that engage students with data
- Research how learners conceive of data and how these conceptions change with time and experience
- Figure out how DS fits with computational thinking

At the school level, will data science be seen as a subject, or as habits of mind brought about by data-immersive experiences across all disciplines?

Data Habits of Mind

Some Examples

- Ask for the data. Look for the data.
- Visualize! Look for patterns
  - Graphs as tools, not tasks
- Ask "Where did these data come from?"
- Ask "Will the intended uses of these data cause harm?"
- Look for bias
- Be aware of variability
- Consider data moves