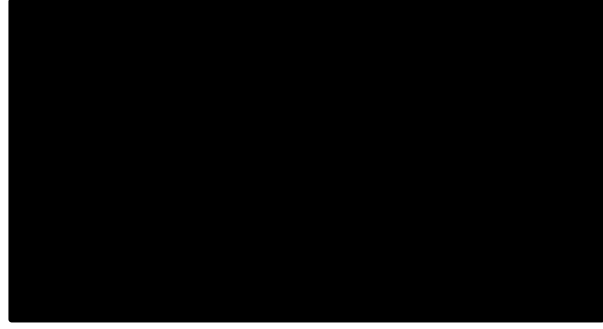




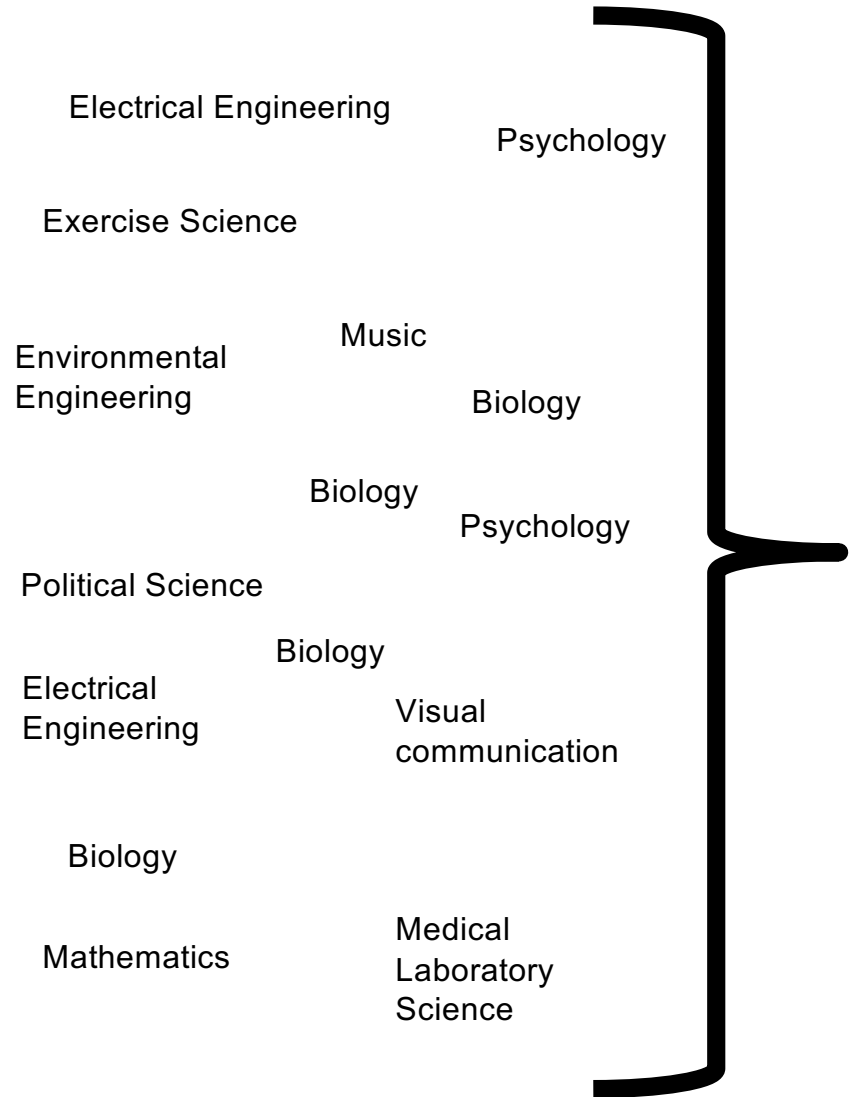
Biomedical Data Science Education at the University of Utah

Designing a Transfer Function from Novices to Professionals

AMIA 2018



Biomedical Data Scientist



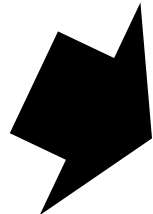
Biomedical Data Scientist

What is Data Science? Who can be a Data Scientist

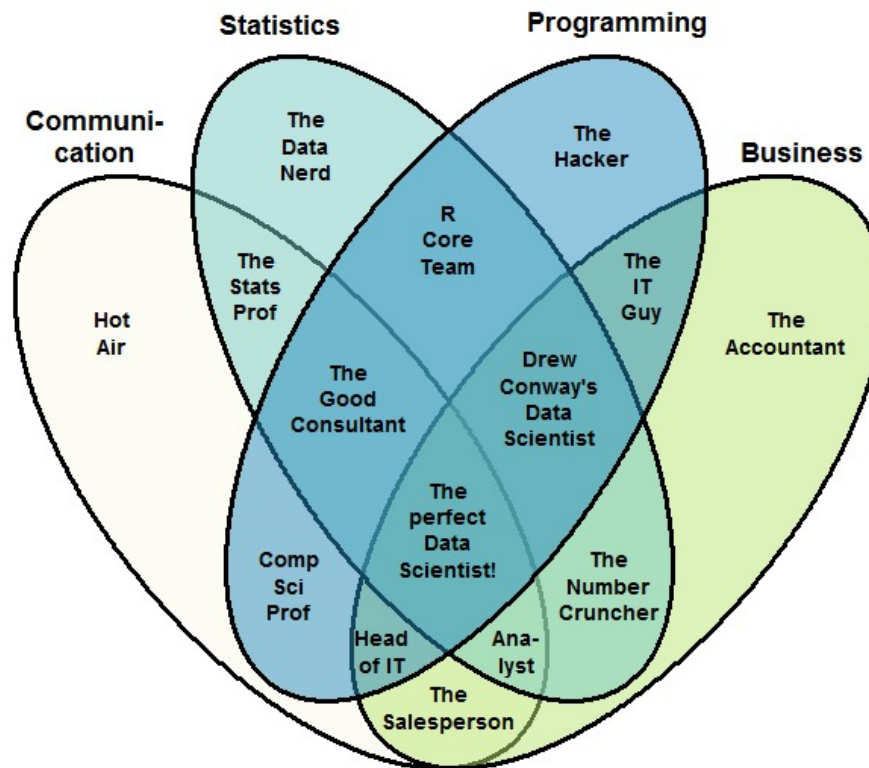
An art student in
Biomedical data science?



Visual Communication



The Data Scientist Venn Diagram



•Wikipedia

Teaching Novices Biomedical Data Science

- Adult learners want the why/context
- Data Science Integrated **Throughout**

Curriculum

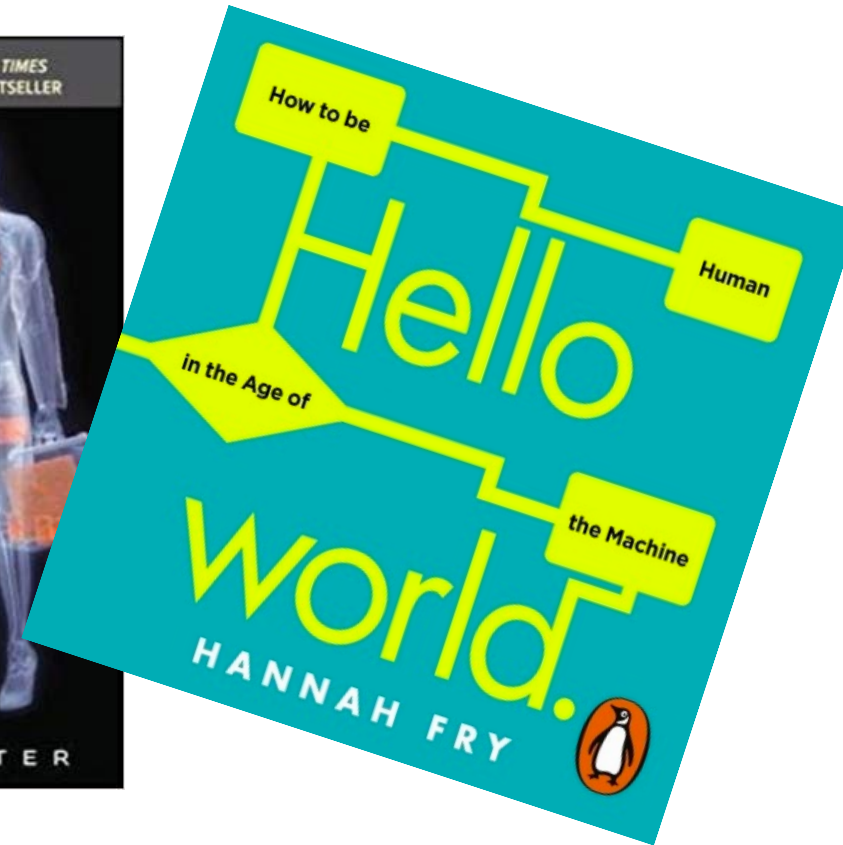
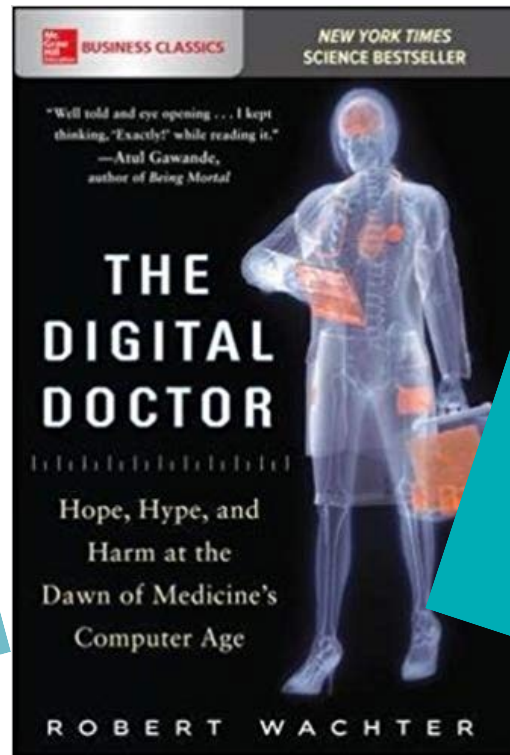
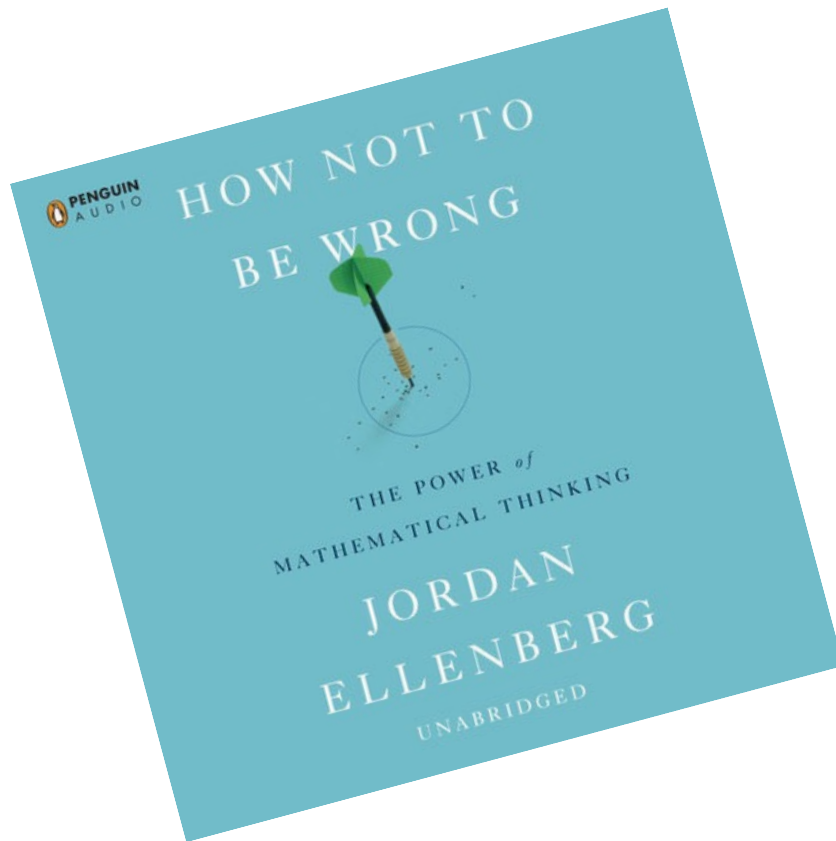
–Compare Harvey Mudd's 1st Year Curriculum

Greg Wilson (Software Carpentry)



- We teach kids to read and then write
- Why do we start off expecting students to write programs before “reading”

Reading Biomedical Data Science

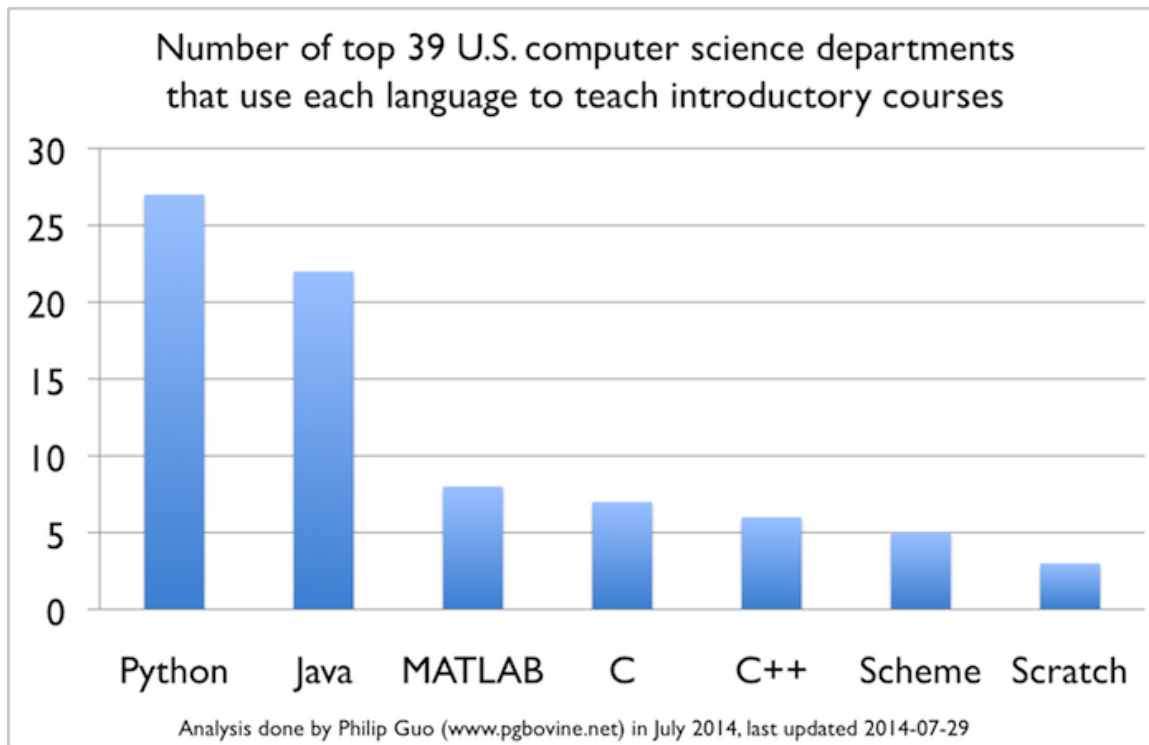


(Minimizing) Tool Troubles

- Git is trouble
 - Disheartening
- Jupyter->
Jupyterhub
 - nbstripout
- Docker/Singularity



Language Wars: What Should We Teach?



- The language you choose matters
- Some languages lead to faster competency and confidence
 - Koulouri, et al. 2014
 - Stefik and Siebert 2013

But “Employers want ‘Multilanguage’ employees”



Mark Guzdial

It’s hard for a novice to move from one programming language to another. Student knowledge of programming tends to be tied to the specific programming language and the textual details of that language ("surface representation").

“Employers want ‘Multilanguage’ employees”



Mark Guzdial

“Students are going to successfully transfer from one programming language to another only once they have developed deep understanding of the first programming language....

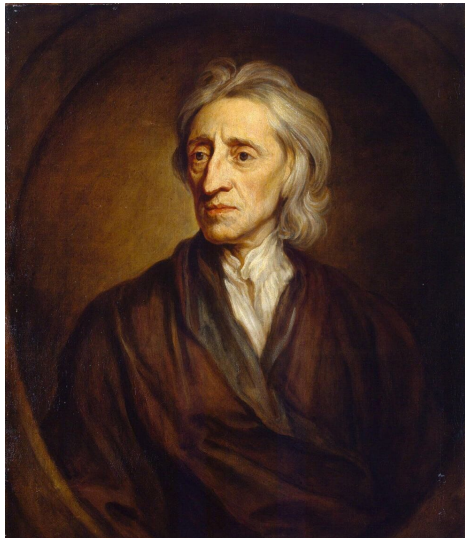
If you move between languages too soon, students just struggle with the different syntax elements, and the switch may actually *delay* their development of the deeper levels of understanding.”

Utah's Mistake (Our Only Mistake!)



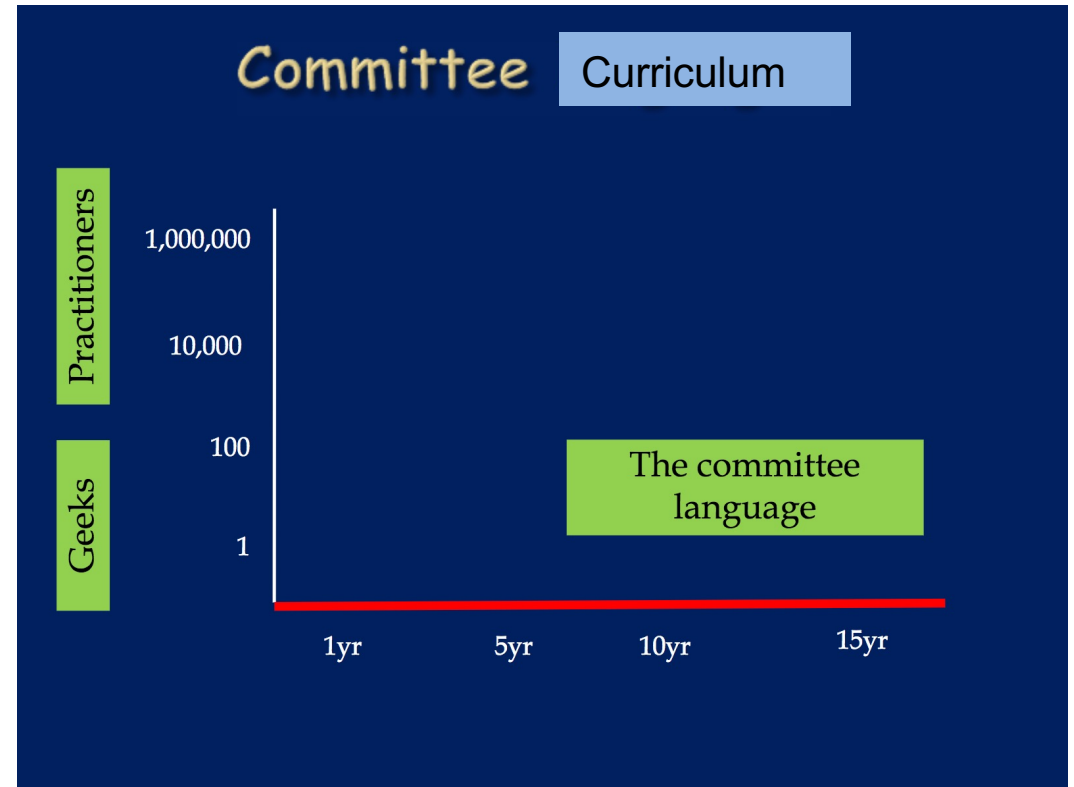
Even if R is superior to Python in practice for probability and statistics, having chosen Python as our initial language, Python is sufficient for novice instruction and it is more important to keep single language.

Designing a Biomedical Data Science Curriculum



“For every church is orthodox to itself; to others, erroneous or heretical.”

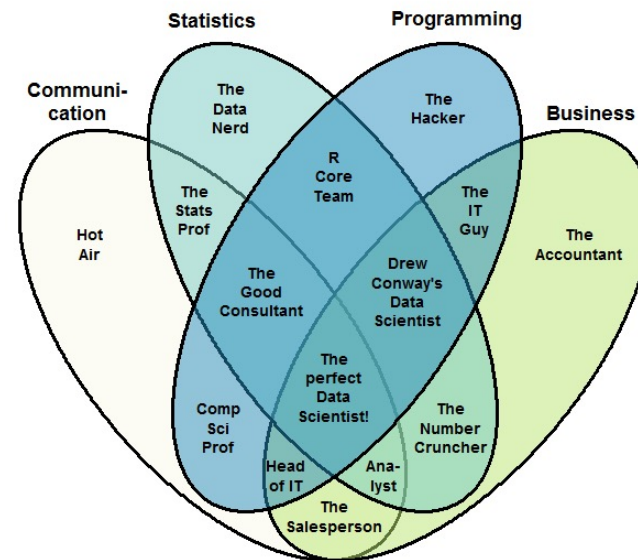
John Locke, *A Letter concerning Toleration*



Simon Peyton Jones

Data Science is Crowded with Competing Orthodoxies

The Data Scientist Venn Diagram



What Orthodoxies are We Willing to Sacrifice for a Feasible Biomedical Data Science Program?