Incorporating data science into undergraduate microbiology

Kim Dill-McFarland
U. of British Columbia
@kdillmcfarland

June 21, 2019


reproducible
research
education
practice
data science
open
approaches
community
teaching
incorporate
undergraduate
microbiology
science
new
computational
statistics
infrastructure
informal
graduate
strategies
incorporate
undergraduate
computer
teaching
adoption
tools
aligned
skills
field
practical
workshops
experience
Jupyter
everyone
introduce
data
programs
practices
computation
researchers
interdisciplinary
publishing
math
prominent
disciplines
results
increasingly
analyzing
educators
non-classroom
developing
linking

$>$ 90% of researchers in the biological sciences work with or plan to work with big data.

Williams & Teal 2017

Poll

When in your career did you *start* acquiring data science skills?

a. K12
b. Undergraduate
c. Graduate
d. Postdoc
e. PI / Scientist / Researcher
f. I do not use data science

$>$ 60% of researchers in the biological sciences report a need for more training in data science.

Meta-analysis 2013 - 2016
Attwood *et al* 2017

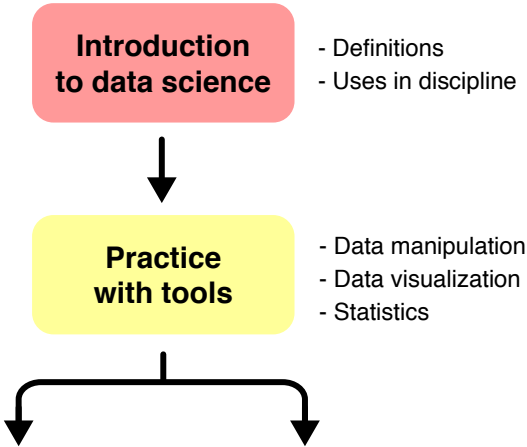

data science
reproducible
research
education
practice
open
community
approaches
incorporate
undergraduate
microbiology
science
teaching
new
educators
non-classroom
analyzing
developing
increasingly
linking
computational
results
disciplines
prominent
statistics
informal
infrastructure
graduate
strategies
math
publishing
big
interdisciplinary
researchers
computer
computation
practices
skills
adoption
tools
aligned
field
programs
practical
data
introduce
workshops
experience
Jupyter
everyone

Experiential
Data science for
Undergraduate
Cross-disciplinary
Education

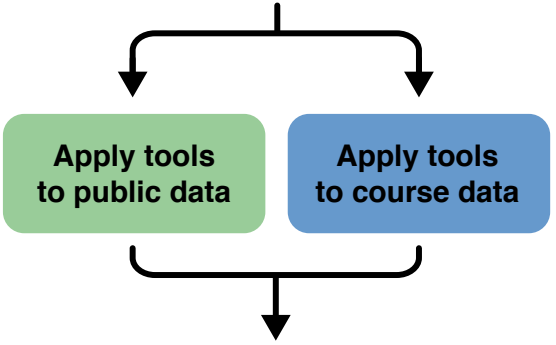
Our goal:

Modular integration of data science curriculum
into existing microbiology courses

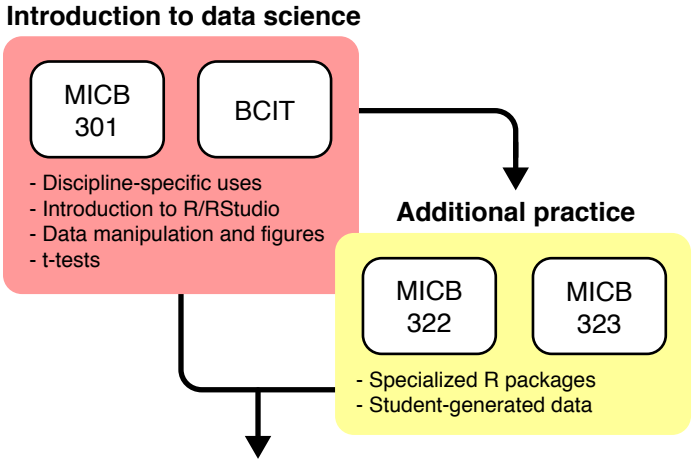
Content overview



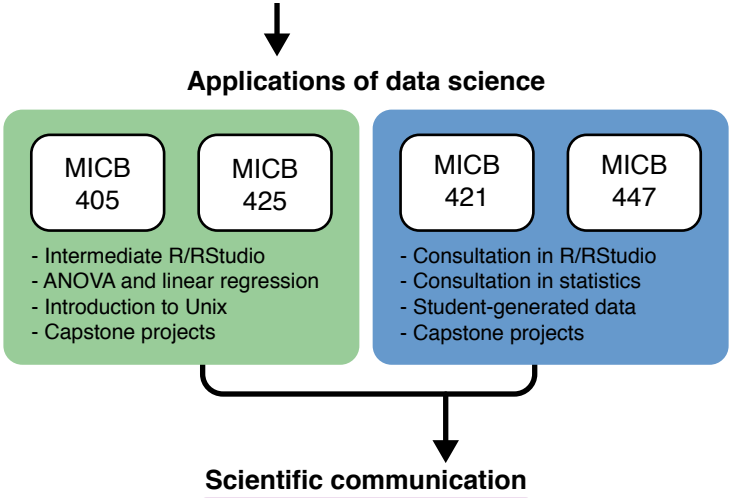
Content overview



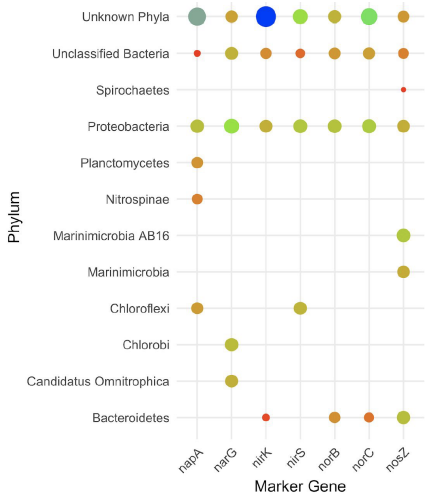
Course overview



Course overview



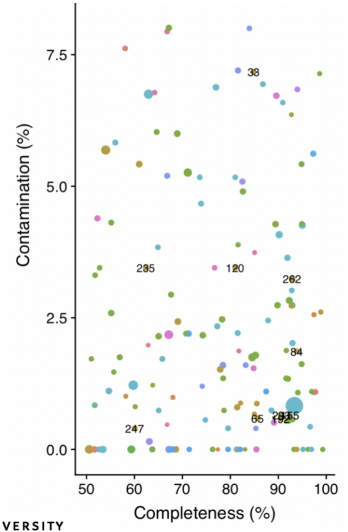
Examples



Genomic and transcriptomic analysis of denitrification processes at 120m in Saanich Inlet

Kevin Chan, Dionzie Ong, Adam Mesa, Gurkiran Parmar, Raaghul Anand

Examples



A Shallow Dive into Saanich Inlet: Metagenomics Analysis at 10m Depth

Virginia Pichler, Jessica Ha, Vivian Li (Greg Morgan, William Zeng)

Does EDUCE effectively teach
data science skills
to microbiology students?

Poll

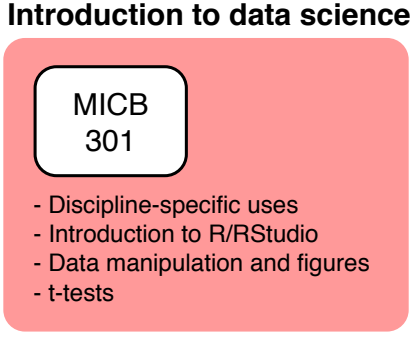
What aspect of 'data science' do you think undergraduate students find most difficult?

a. Programming / scripting
b. Statistics
c. Installing software
d. Interpretation within the context of microbiology
e. Other (Provide your response in our open GitHub issue!)

Does EDUCE effectively teach
data science skills
to microbiology students?

MICB 301 as a case study

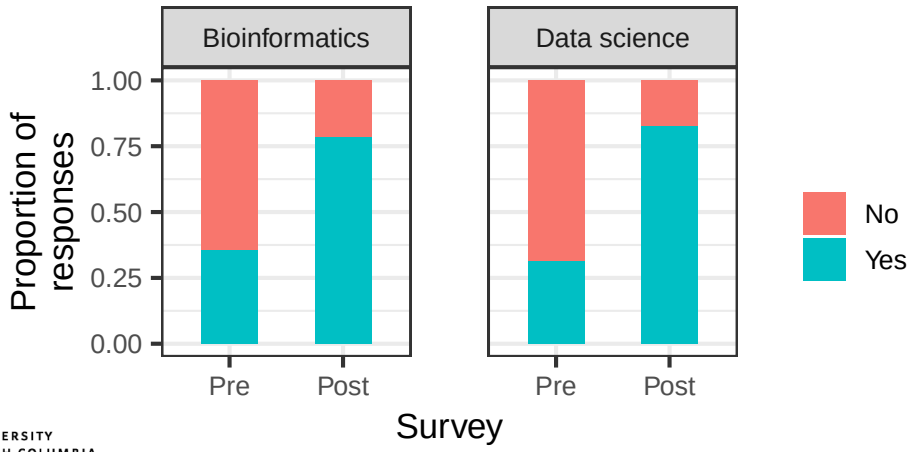
EDUCE in MICB 301



- Required for MICB majors
- ~120 students / yr
- 5 x 50 min class sessions across 5 weeks

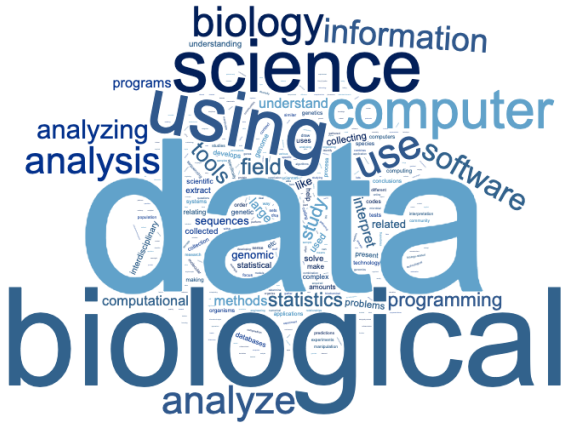
Increased awareness of data science

Have you heard the term _____?



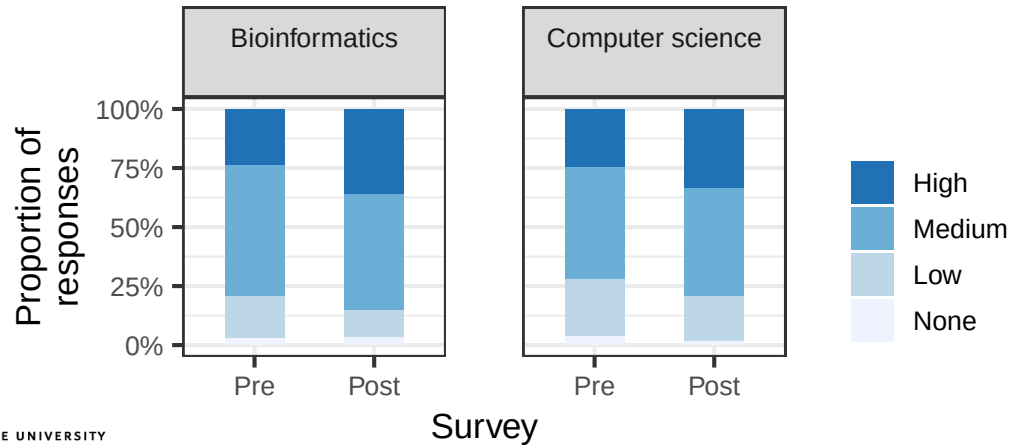
Definitions of data science

Definitions of bioinformatics



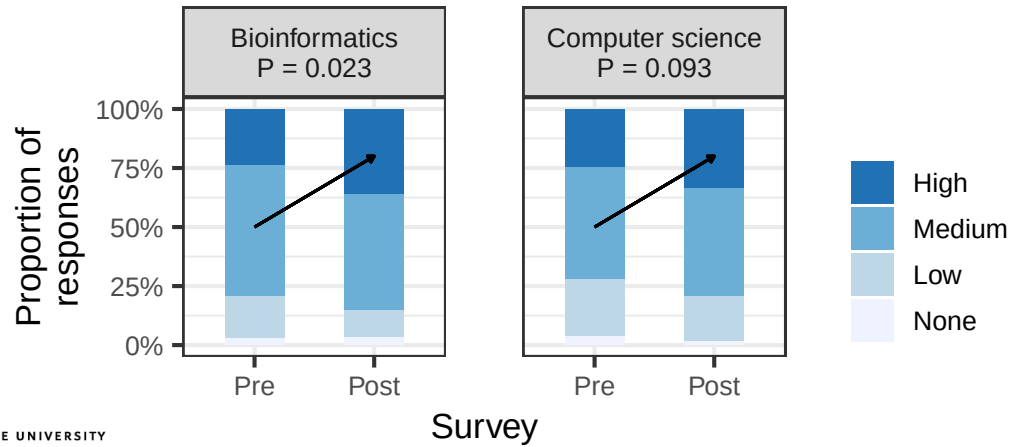
Increased interest in data science

How would you rate your interest in...



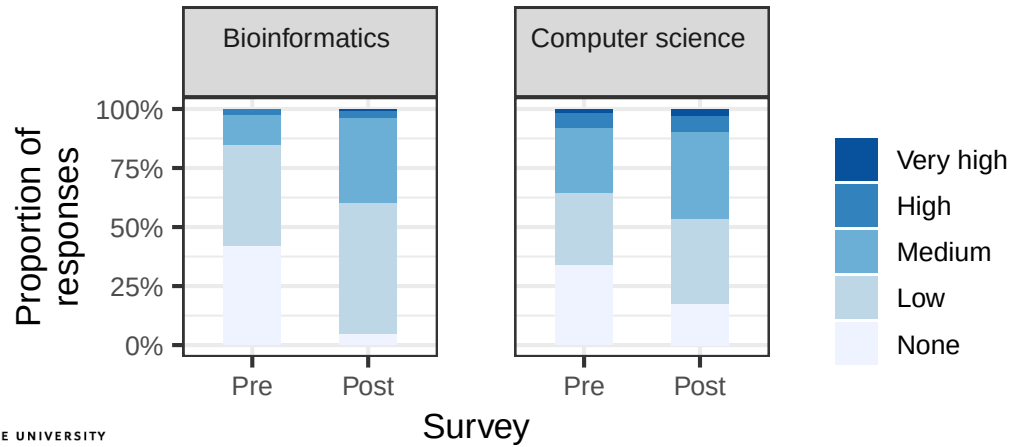
Increased interest in data science

How would you rate your interest in...



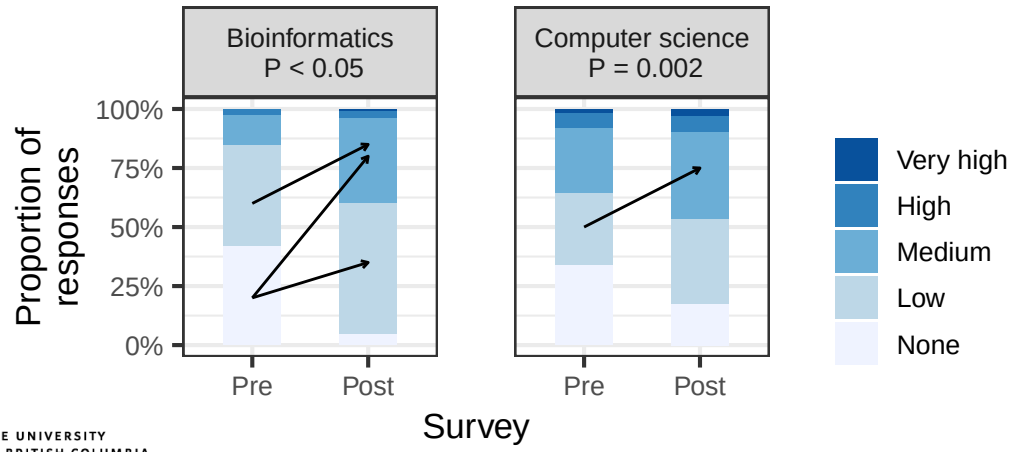
Increased experience in data science

What level of experience do you have in ...



Increased experience in data science

What level of experience do you have in ...



No changes in interest or experience
in “microbiology” or “statistics”

Conclusions

- Data science literacy is needed in the life sciences

Conclusions

- Data science literacy is needed in the life sciences
- EDUCE provides a flexible, modular approach for integrating data science into undergraduate curriculum

Conclusions

- Data science literacy is needed in the life sciences
- EDUCE provides a flexible, modular approach for integrating data science into undergraduate curriculum
- Even minimal exposure (5 hours) can increase student self-reported knowledge, interest, and experience in data science areas

The future

- A wealth of survey data
- More courses? Other departments?
- Faculty of Science Data Science Curriculum Committee
- Independent learning tools: https://ubc-educe.shinyapps.io/course_knitter/

Acknowledgements

Steven Hallam
Jennifer Bonderoff

EDUCE TAs
Julia Beni
Kris Hong
Jonah Lin
Yue Liu
Florent Mazel
Lisa McEwen
Ryan McLaughlin
Connor Morgan-Lang
Nolan Shelley
David Yin

Course instructors
Sean Crowe
Lindsay Eltis
Jennifer Gardy
Marcia Graves
Martin Hirst
Bill Mohn
Dave Oliver
Jen Sibley

Collaborators
Gaby Cohen-Freue (STAT)
Patrick Walls (MATH)
Biljana Stojkova (ASDa)

Funding

UBC Teaching and Learning Enhancement Fund (TLEF)

NSERC CREATE Program (ECOSCOPE)

Department of Microbiology & Immunology

UBC Microbiome Research Network (MRN)

UBC Skylight and the Center for Teaching, Learning and Technology (CTLT)

Want to learn more?

EDUCE GitHub: https://github.com/EDUCE-UBC

EDUCE website: https://educe-ubc.github.io/

EDUCE email: info.educe@ubc.ca

Undergraduate Journal of Experimental Microbiology and Immunology (uJEMI): https://jemi.microbiology.ubc.ca/

References

Attwood TK *et al* 2017. *A global perspective on evolving bioinformatics and data science training needs.* Brief Bioinform. 20(2):398-404. doi: 10.1093/bib/bbx100

Williams JJ & Teal TK. 2017. *A vision for collaborative training infrastructure for bioinformatics.* Ann N Y Acad Sci. 1387(1):54-60_ doi: 10.1111/nyas.13207

Students impacted per year