Teaching data science authentically

using open source educational resources

Tiffany A. Timbers

2021-08-03

Why create and use open educational resources in general?

- cost savings for learners 💸
- quick iteration and extension on materials for instructors ♻️ + 🏗️
- and raising the quality of the resources by facilitating collaboration 👥👥👥
- your teaching materials (and pedagogies) become more discoverable 🔍

Why create and use open educational resources in data science?

All the reasons I just stated, and...

... it mirrors the practices, tools and workflows used when practising data science.

Practice what you preach!

Example of OER's we have built using Data Science Open Source tools

1. *Data Science: A First Introduction* - an open textbook aimed at undergraduates students taking their first course in data science
2. Syllabi, lecture notes, labs and lecture videos from courses for a professional Master's in Data Science program
3. Key Capabilities in Data Science online courses - Interactive online learning modules aimed at mid-career learners

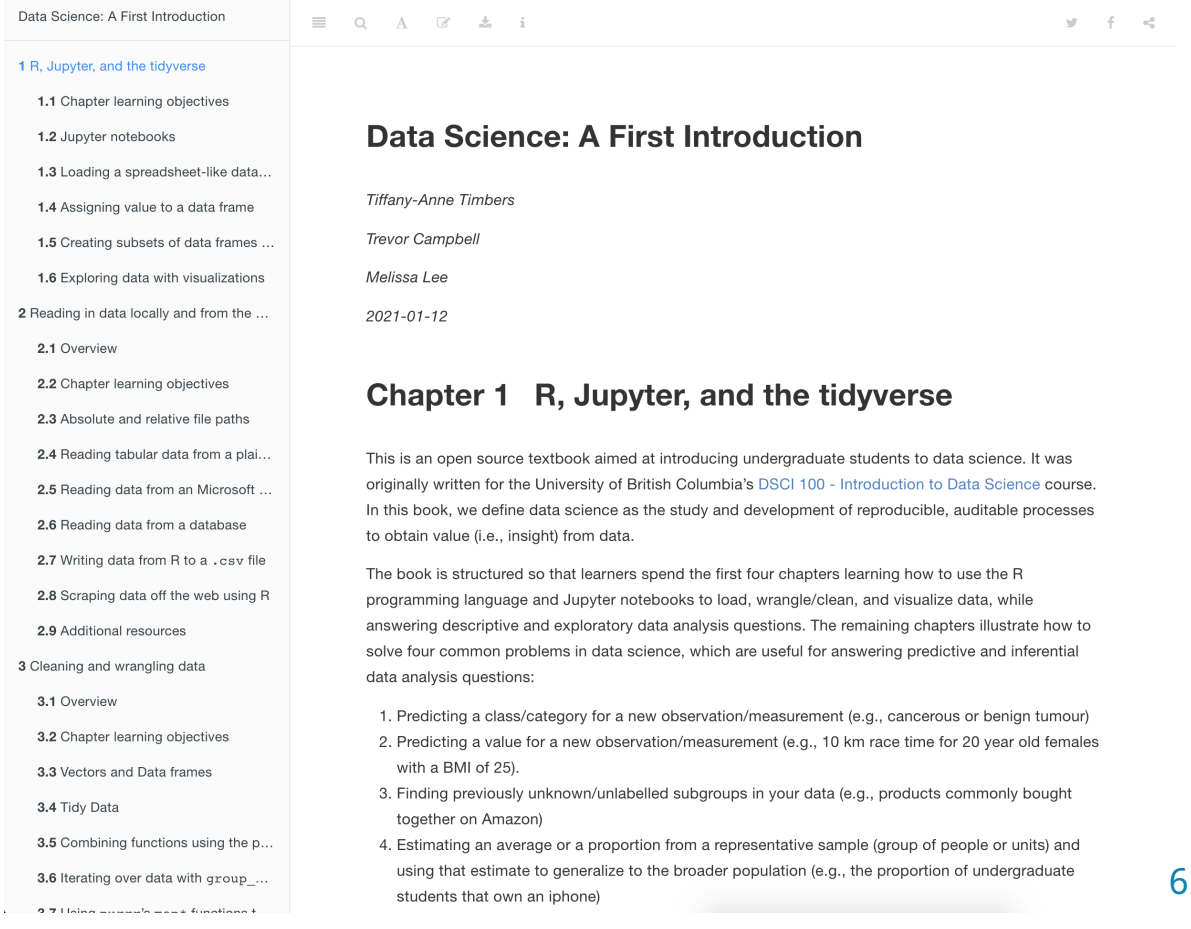
Data Science: A First Introduction

Aimed at first year undergraduates, from any discipline.

Tiffany Timbers

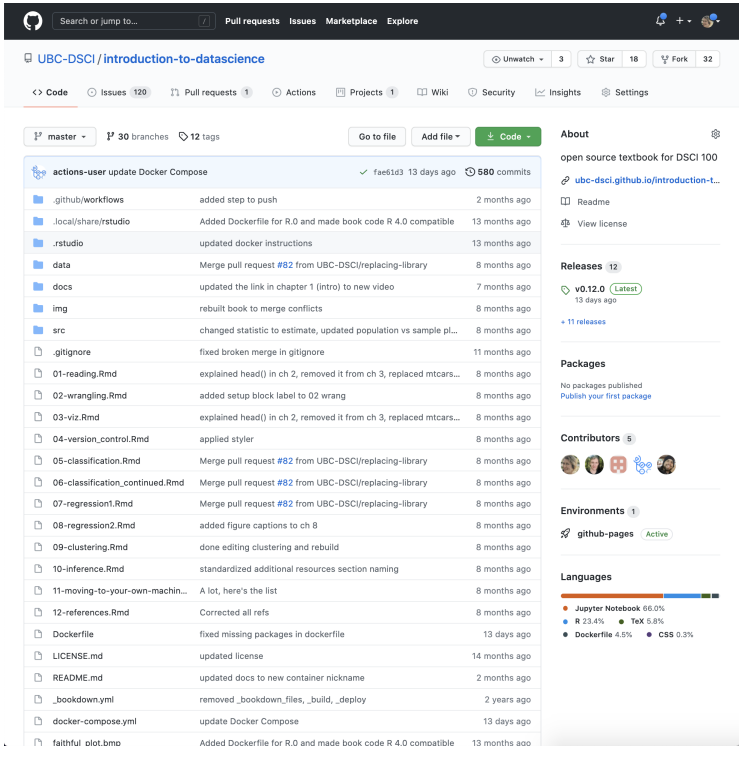
Trevor Campbell

Melissa Lee



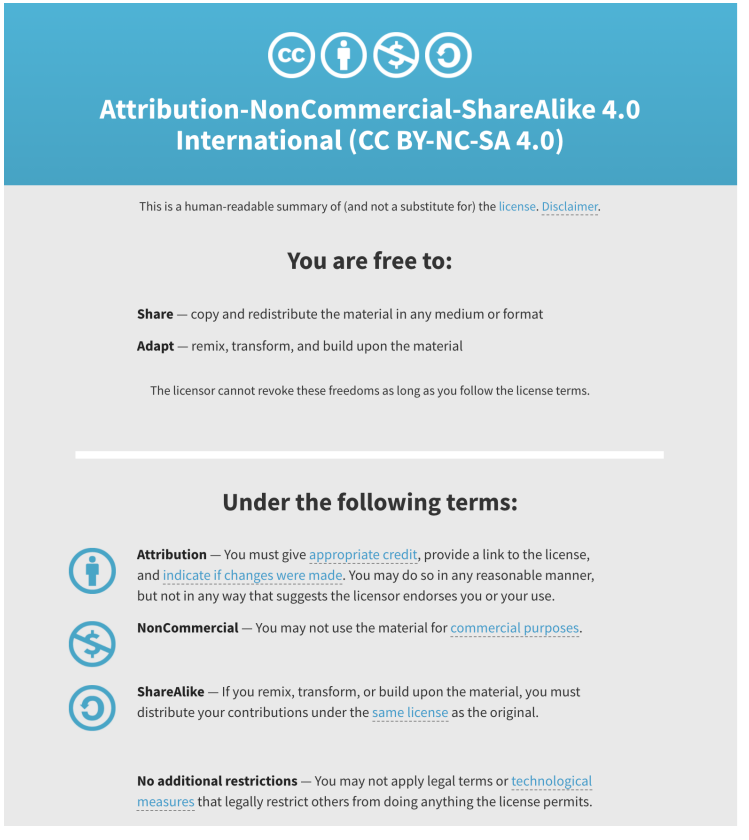
Data Science: A First Introduction

- Created via R + bookdown R package.
- Source code is available on GitHub: github.com/UBC-DSCI/introduction-to-datascience



Data Science: A First Introduction

- Book source and rendered HTML version openly licensed under Creative Commons BY-NC-SA 4.0 and will remain that way indefinitely.
- Partnering with CRC Press to publish and sell a print version.



Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0)

This is a human-readable summary of (and not a substitute for) the license. Disclaimer.

You are free to:

Share — copy and redistribute the material in any medium or format

Adapt — remix, transform, and build upon the material

The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms:

Attribution — You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

NonCommercial — You may not use the material for commercial purposes.

ShareAlike — If you remix, transform, or build upon the material, you must distribute your contributions under the same license as the original.

No additional restrictions — You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.

UBC MDS syllabi, lecture notes, labs & lecture videos

Course resources from the UBC Professional Master of Data Science program.

Past and present core MDS teaching team members

- Tomas Beuzen
- Vincenzo Coia
- Giulio Valentino Dalla Riva
- Mike Gelbart
- Varada Kolhatkar
- Rodolfo Lourenzutti
- Firas Moosvi
- Joel Ostblom
- Alexi Rodríguez-Arelis
- Tiffany Timbers

UBC MDS syllabi, lecture notes, labs & lecture videos

- Course resources from a Master of Data Science program: github.com/UBC-MDS/public
- Program uses OER's from others, as well as creates its own
- Syllabi and course notes created via Jupyter, R Markdown & hosted on GitHub
- Videos of some courses hosted on YouTube

UBC MDS public teaching materials

This repository contains publicly available versions of some of the teaching materials from the UBC Master of Data Science program. See below for the currently available material. The links in the left column lead to course-specific repositories.

Course Number	Course Title	syllabus	lecture notes	labs	lecture videos
DSCI 511	Programming for Data Science	✓	✓		✓
DSCI 521	Computing Platforms for Data Science	✓	✓	✓	
DSCI 551	Descriptive Statistics and Probability for Data Science	✓	✓		
DSCI 542	Communication and Argumentation	✓			
DSCI 523	Programming for Data Manipulation	✓			
DSCI 512	Algorithms and Data Structures	✓	✓	✓	✓
DSCI 531	Data Visualization I	✓	✓	✓	
DSCI 552	Statistical Inference and Computation I	✓			
DSCI 522	Data Science Workflows	✓	✓	✓	
DSCI 513	Databases and Data Retrieval	✓			
DSCI 561	Regression I	✓			
DSCI 571	Supervised Learning I	✓			
DSCI 572	Supervised Learning II	✓	✓	✓	
DSCI 563	Unsupervised Learning	✓			
DSCI 573	Feature and Model Selection	✓			
DSCI 532	Data Visualization II	✓			
DSCI 562	Regression II	✓	✓	✓	
DSCI 524	Collaborative Software Development	✓	✓	✓	
DSCI 553	Statistical Inference and Computation II	✓			
DSCI 574	Spatial and Temporal Models	✓			
DSCI 541	Privacy, Ethics, and Security	✓			
DSCI 525	Web and Cloud Computing	✓			
DSCI 554	Experimentation and Causal Inference	✓			
DSCI 575	Advanced Machine Learning	✓			
DSCI 591	Capstone project	✓	n/a	n/a	

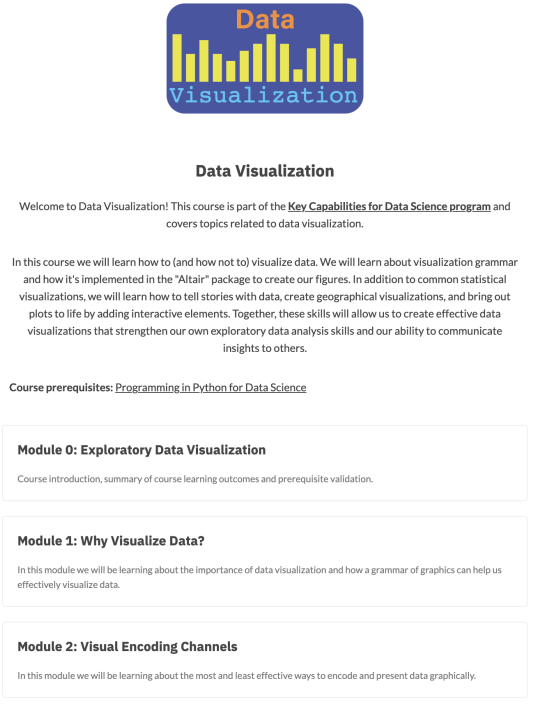
Course instructors choose to share and distribute their resources using one of the available Creative Commons licenses.

Key Capabilities in Data Science online courses

Suite of online courses, and a certificate program, aimed at mid-career learners in British Columbia, Canada, but open to anyone.

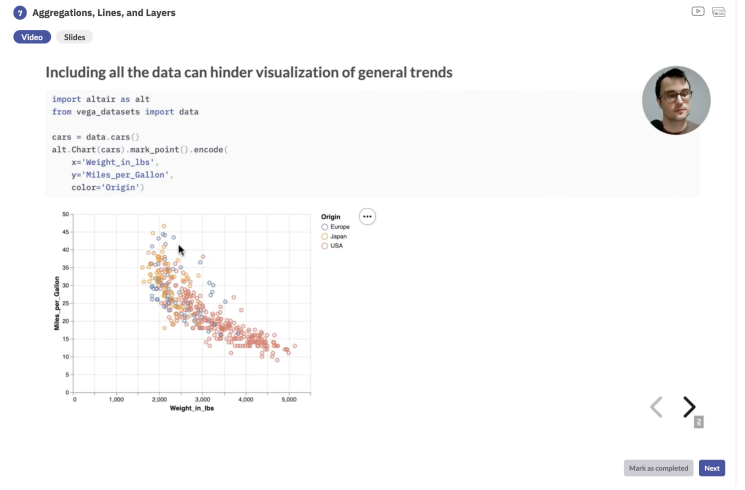
UBC certificate program: https://extendedlearning.ubc.ca/programs/key-capabilities-data-science

- Course notes, videos and knowledge checks are openly available.
- Paid version of the courses includes a course facilitator, office hours, graded assignments and quizzes.



Key Capabilities in Data Science online courses

Open content created via Ines Montani's open course framework, whose front-end is powered by Gatsby and Reveal.js and back-end code execution is powered by BinderHub.



Key Capabilities in Data Science online courses

Course open content	Course source
Programming in Python for Data Science	github.com/UBC-MDS/programming-in-python-for-data-science
Data Visualization	github.com/UBC-MDS/exploratory-data-viz
Introduction to Machine Learning	github.com/UBC-MDS/introduction-machine-learning

Lessons learned from building, sharing and maintaining data science OER's

- don't do it alone
- don't be afraid to share
- use a license
- maintenance is needed to keep material current
- OER's which need interactive computation require money or infrastructure to share

Impact of OER's

Attribution

Resources

Thanks