Enhancing accessibility in data science education: lessons from a neuroscience undergraduate course

Theresa H. McKim, Ph.D.
Assistant Teaching Professor
University of Nevada, Reno (UNR)
07/19/2024

tmckim@unr.edu

Pre-session survey: 6 questions

- Link to google form (anonymous): https://forms.gle/rzPDtbgjZ4bGWPm36

- Check access to google colab: https://colab.research.google.com
  - Use chrome browser
  - Click ‘Sign in’ in top right corner

Pre-session google form QR code:

Materials are available from my github account:
https://github.com/tmckim/NTC_2024
- Publicly available so anyone can access

Objectives for today

discuss & demonstrate ways to effectively teach coding to undergraduates

address & discuss solutions to barriers of entry for both instructors and students

share open access resources/materials

Code along demo: ease and accessibility of code notebooks

Motivation (the **Why**): combining data science & neuroscience

Core Competencies in Neuroscience Training

Undergraduate Training
I. Conceptual Knowledge
II. Analytic & Scientific Thinking
Ability to **collect, analyze, and interpret** quantitative information
Abilities in scientific inquiry, such as hypothesis development, experimental design, and **data analysis & interpretation**
III. Rigorous & Responsible Conduct of Research (RCR)
IV. Communication Skills
V. Individual development & Professionalism

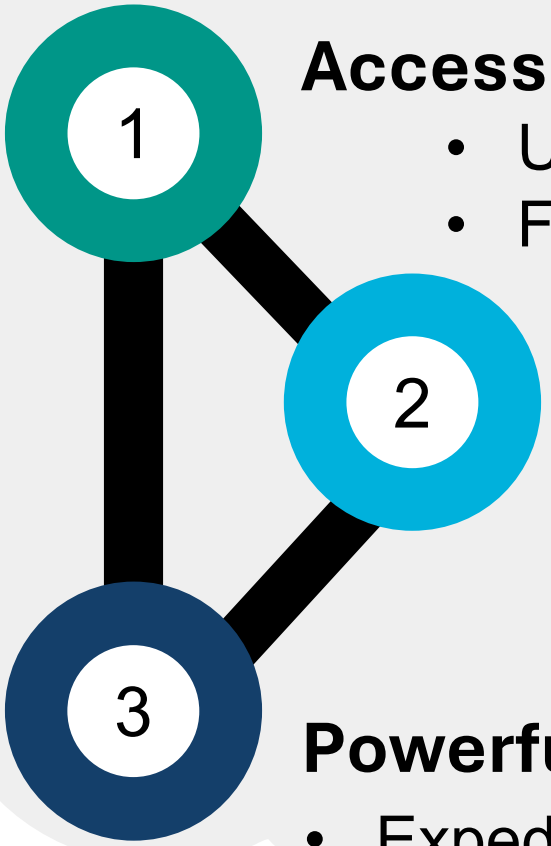
SfN Society for Neuroscience

Graduate Training
I. Conceptual Knowledge
II. Research Skill Development
Proficiency in experimental design and scientific tools and techniques, including **computer-based data handling**
Proficiency in **data analysis & interpretation**
III. Rigorous & Responsible Conduct of Research (RCR)
IV. Communication Skills
V. Career development & Professionalism

What students from a neurobiology lab course said

1 session
(4 hr lab)

Coding skills are:



Accessible

- Use your own computer/tablet
- Free resources and online coding environments

Collaborative

- Shareable and work together with others
- Use resources created by others (free packages)
- Forums and answers online

Powerful

- Expedite data processing & analysis
- Quick statistical analysis
- Visualization and prediction

Credit: NS 479 - Neurotechniques lab students, UNR SP24

The **How**: Including coding in the curriculum

Exposure to coding
- Existing course – short module
- Students edit code (see utility)
- Code could be optional (GUIs)

Discipline-based coding course(s)
- Class dedicated to coding in neuroscience context
- Design and write their own code
- May include novel research projects

Less independent coding
Lower instructor proficiency

More independent coding
Higher instructor proficiency

Integration of coding
- 1-2 hrs intro lessons (standalone modules)
- Edit code and begin to write their own
- Ask their own questions of data

Slide from: Juavinett, 2023 FUN workshop

The **What**: Resources for getting started

- Data8 at UC Berkeley: https://www.data8.org/zero-to-data-8/intro.html

PMC8040835 The Journal of Undergraduate Neuroscience Education (JUNE), Fall 2020, 19(1):A94-A104

ARTICLE

Learning How to Code While Analyzing an Open Access Electrophysiology Dataset

Ashley Juavinett
Division of Biological Sciences, Neurobiology Section, UC San Diego.

PMC9053425 The Journal of Undergraduate Neuroscience Education (JUNE), Fall 2021, 20(1):A100-A110

ARTICLE

Pandemic Teaching: Using the Allen Cell Types Database for Final Semester Projects in an Undergraduate Neurophysiology Lab Course

Yi-Yun Ho[1], Andrea Roeser[1], Gwenda Law[2], and Bruce R. Johnson[1]
[1]Department of Neurobiology and Behavior, Cornell University, Ithaca, NY 14853; [2]Department of Biomedical Engineering, Cornell University, Ithaca, NY 14853.

PMC10768818 The Journal of Undergraduate Neuroscience Education (JUNE),Spring 2023, 22(1):A66-A73

CASE STUDY

Mapping Human Neuronal Diversity in the Search for New Therapeutics: Using Real Human Neuron Data Sets to Build Student Quantitative Skills

Emma C. Milligan[1], Kaitlyn Casimo[2], Laurie Buchanan[3], Bryant Hutson[3], Sabrina Robertson[1]
[1]Psychology and Neuroscience Department, University of North Carolina Chapel Hill, Chapel Hill, NC 27599; [2]Allen Institute, Seattle, WA 98109; [3]Office of Institutional Research and Assessment, University of North Carolina Chapel Hill, Chapel Hill, NC 27599.

PMC8040850 The Journal of Undergraduate Neuroscience Education (JUNE), Fall 2020, 19(1): A105-A112

ARTICLE

Non-Disposable Assignments for Remote Neuroscience Laboratory Teaching Using Analysis of Human Data

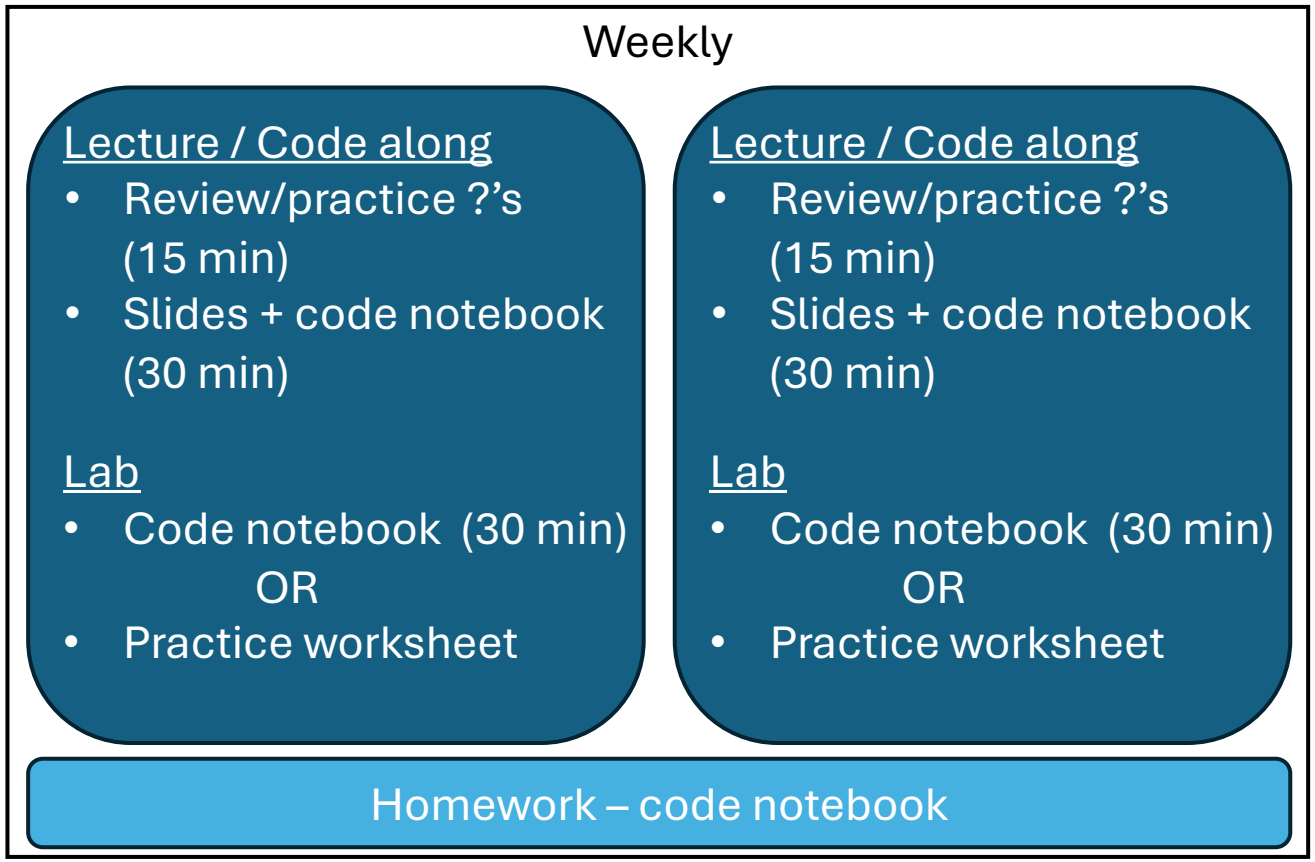
Sally B. Seraphin[1] and Shannon Stock[2]
[1]Neuroscience Program, Trinity College, 300 Summit Street, Hartford, CT 06106. [2]Mathematics and Computer Science, College of the Holy Cross, 1 College Street, Worcester, MA 01610.

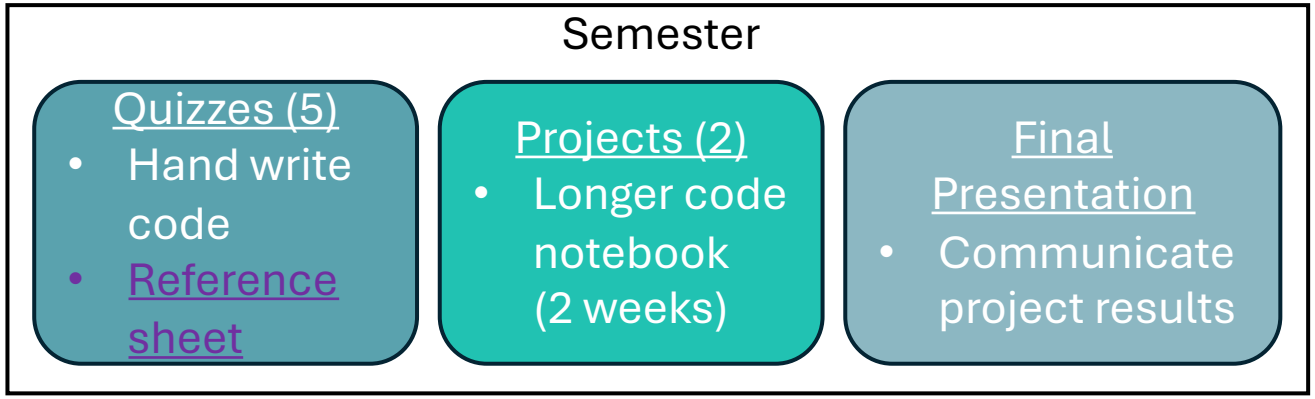
Course Structure

- 3 cr. / No pre-reqs
- 2x per week / 75 mins

Modules

1. Data organization & tables
2. Data visualization & methods to work with data
3. Sampling & models
4. Hypothesis testing & experimental design
5. Neuro methods, datasets, analysis & sci comm





Tech Stack: Accessible resources & materials

Code notebooks

Google colaboratory

- Requires google account

Python packages

Foundation:

- datascience
  - developed by UCB for teaching Data8
- NumPy
- matplotlib

Application:

- pandas
- seaborn

Other options for using Jupyter Notebooks

Jupyter Hub

- Distributions managed in the cloud (self or university)

Anaconda

- Download locally to computer

Barriers to bringing coding into neuroscience courses

- Ease of access to coding environments
- Instructor knowledge of coding
- Unknown need of ‘applied’ vs. ‘basic’ programming courses
- Student perceptions of coding

Addressing these requires:

Accessible, discipline-specific coding resources for all educators

Juavinett (2022) *Neuron* NeuroView
The next generation of neuroscientists needs to learn how to code, and we need new ways to teach them

Slide from: Juavinett, 2023 FUN workshop

Code notebook demo

Notebook File:

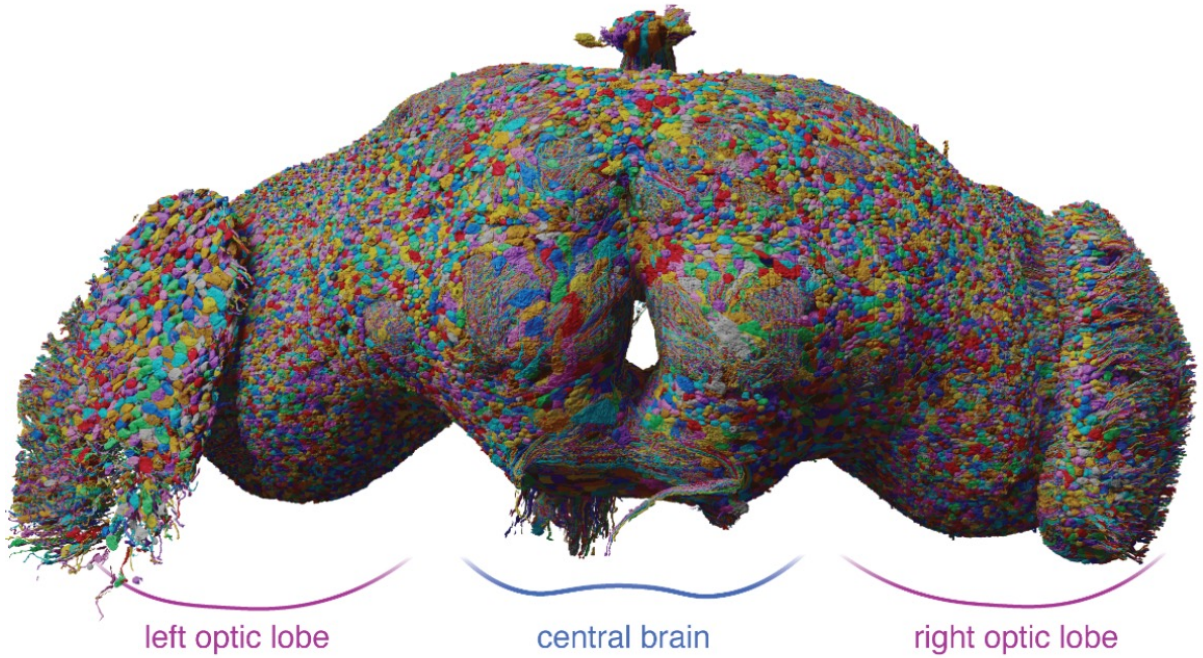
https://colab.research.google.com/github/tmckim/NTC_2024/blob/main/PlottingNeuralDataDemo_NTC2024.ipynb

All materials:

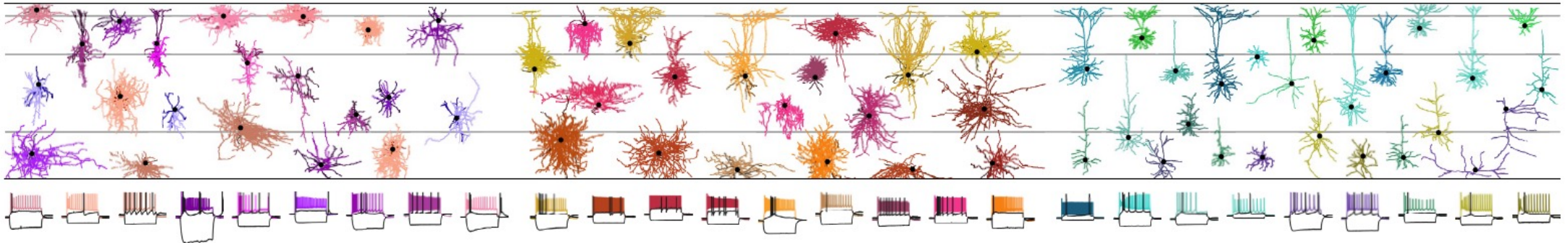
https://github.com/tmckim/NTC_2024

Coding skills will enable students to:

- Access & analyze large, open access datasets
- Perform statistics & visualization
  - Interpret & communicate results
- Apply knowledge toward computational modeling (neural networks, AI)
- Prepare for jobs, post-bacc research, grad school, etc.



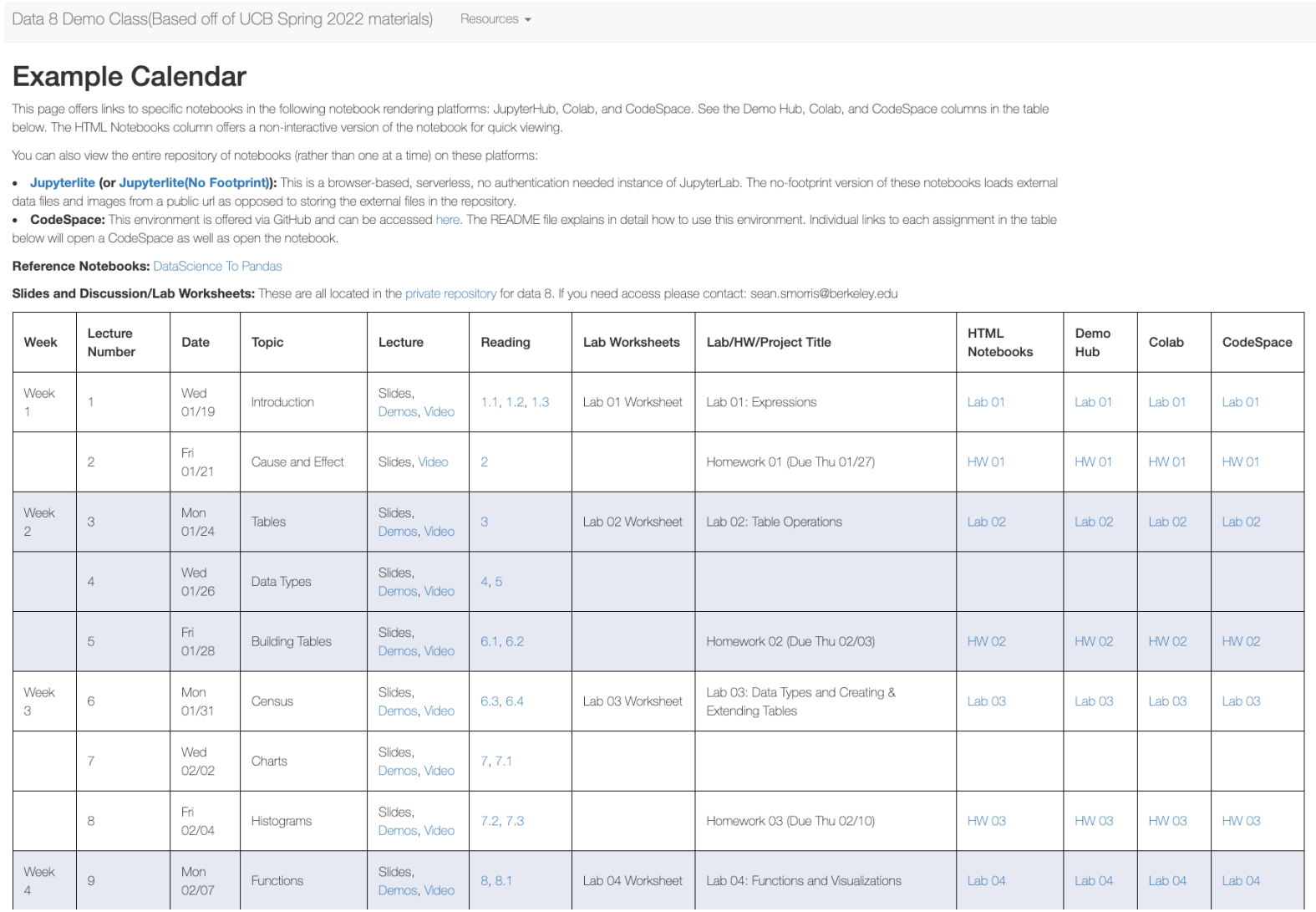
Dorkenwald et al. (2024 preprint)



Scala et al. (2021)

Resources from UCB Data8

- The Data8 Pedagogy Guide
- Data 8 Adoption Website
- Online Textbook - Computational and Inferential Thinking: The Foundations of Data Science
- Example course website: https://www.data8.org/materials-sp22/demo.html
- YouTube videos



Data 8 Demo Class(Based off of UCB Spring 2022 materials) Resources

Example Calendar

This page offers links to specific notebooks in the following notebook rendering platforms: JupyterHub, Colab, and CodeSpace. See the Demo Hub, Colab, and CodeSpace columns in the table below. The HTML Notebooks column offers a non-interactive version of the notebook for quick viewing.

You can also view the entire repository of notebooks (rather than one at a time) on these platforms:

- **Jupyterlite (or Jupyterlite(No Footprint)):** This is a browser-based, serverless, no authentication needed instance of JupyterLab. The no-footprint version of these notebooks loads external data files and images from a public url as opposed to storing the external files in the repository.
- **CodeSpace:** This environment is offered via GitHub and can be accessed here. The README file explains in detail how to use this environment. Individual links to each assignment in the table below will open a CodeSpace as well as open the notebook.

Reference Notebooks: DataScience To Pandas

Slides and Discussion/Lab Worksheets: These are all located in the private repository for data 8. If you need access please contact: sean.smorris@berkeley.edu

Week	Lecture Number	Date	Topic	Lecture	Reading	Lab Worksheets	Lab/HW/Project Title	HTML Notebooks	Demo Hub	Colab	CodeSpace
Week 1	1	Wed 01/19	Introduction	Slides, Demos, Video	1.1, 1.2, 1.3	Lab 01 Worksheet	Lab 01: Expressions	Lab 01	Lab 01	Lab 01	Lab 01
	2	Fri 01/21	Cause and Effect	Slides, Video	2		Homework 01 (Due Thu 01/27)	HW 01	HW 01	HW 01	HW 01
Week 2	3	Mon 01/24	Tables	Slides, Demos, Video	3	Lab 02 Worksheet	Lab 02: Table Operations	Lab 02	Lab 02	Lab 02	Lab 02
	4	Wed 01/26	Data Types	Slides, Demos, Video	4, 5						
	5	Fri 01/28	Building Tables	Slides, Demos, Video	6.1, 6.2		Homework 02 (Due Thu 02/03)	HW 02	HW 02	HW 02	HW 02
Week 3	6	Mon 01/31	Census	Slides, Demos, Video	6.3, 6.4	Lab 03 Worksheet	Lab 03: Data Types and Creating & Extending Tables	Lab 03	Lab 03	Lab 03	Lab 03
	7	Wed 02/02	Charts	Slides, Demos, Video	7, 7.1						
	8	Fri 02/04	Histograms	Slides, Demos, Video	7.2, 7.3		Homework 03 (Due Thu 02/10)	HW 03	HW 03	HW 03	HW 03
Week 4	9	Mon 02/07	Functions	Slides, Demos, Video	8, 8.1	Lab 04 Worksheet	Lab 04: Functions and Visualizations	Lab 04	Lab 04	Lab 04	Lab 04

Online resources: Finding datasets

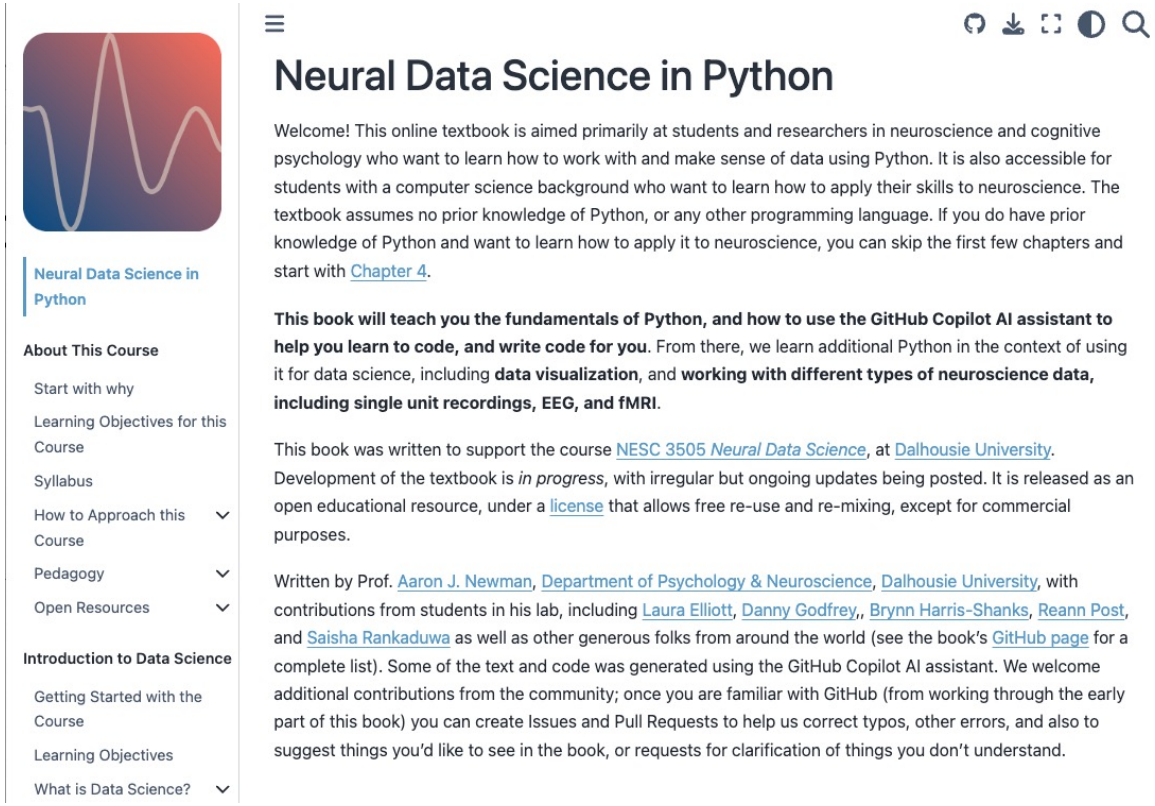
Tons of large, open access datasets, with some examples*:

- https://openneuro.org/
- https://crcns.org/data-sets
- https://www.humanconnectome.org/
- https://www.nitrc.org/
- https://osf.io/
- https://openlists.github.io/
- https://www.nwb.org/example-datasets/

* Not an exhaustive list

Resources: Online websites (notebooks w/code examples)

- https://mark-kramer.github.io/Case-Studies-Python/intro.html
- https://compneuro.neuromatch.io/projects/docs/datasets_overview.html
- https://nwb4edu.github.io/Data_Science_In_Python/SciPy.html
- https://pyforneuro.com/
- https://github.com/Columbia-Neuropythonistas/IntroPythonForNeuroscientists2023



https://neuraldatascience.io/intro.html

* Not an exhaustive list