# NEUROBIO 735: Quantitative Neurobiology

Spring, 2024 3 credit hours (Pass/Fail) Instruction mode: In-person

Through lectures and hands-on problem solving, this course will provide students with a working, practicable background in coding in Python, and theoretical and computational neuroscience. The material will be oriented strongly towards the needs of working neurobiologists, and will require considerable independent work.

# Meeting Dates and Location:

January 5 – April 27 Tuesday and Thursday 3:00-4:30 PM 301 Bryan Building

# Course directors:

John Pearson	john.pearson@duke.edu	101H Bryan Research
Dmitry Velmeshev	dmitry.velmeshev@duke.edu	301E Bryan Research

# Teaching Assistants:

Ziyi Gong	ziyi.gong@duke.edu
Marija Rudzite	andra.rudzite@duke.edu

Office hours by appointment.

# Module Overview

The goal of this module is to give you a sampler of techniques and ideas in quantitative neurobiology, which we consider as encompassing computation, data analysis, modeling, and theory. The course is divided into three main sections (with a final week for project presentations):

- 1. **Introduction to programming:** There are several good options here, but we'll be using Python. More on details below, but *we will not assume prior programming experience.*
- 2. **Analyzing neural data:** The goal here is to get you comfortable using programming to explore, visualize, analyze, and model several types of data generated by neuroscience experiments.

## 3. Bioinformatics approaches to single-cell genomics data analy-

**sis:** Here, we will introduce the main types of single-cell genomics data and will cover key algorithms for their analysis. This module will focus on practical aspects of the analysis, with hands-on workshops.

## Logistics

#### Getting together

Class will be held **in person** in Bryan 301. A few lectures (indicated below) will take place over Zoom.

#### Computing

This year, we will encourage you to use Google Colab for your assignments. This has the advantage of standardizing our Python environment and allowing you to easily share assignments for grading. If you prefer to use a local machine, that's fine; please just discuss with us in advance, since we can't necessarily provide support to debug your setup. Colab uses a variant of Jupyter Notebooks, which are covered in the *Python Data Science Handbook*, but we will also cover the basics of navigating this in our first class.

#### Outside of class

In-class assignments are supposed to be that. After class is done, we will post solutions so you can compare your answers. Outside of class, you are responsible for doing any of the recommended and assigned reading.

You are also responsible for checking the class website. All class materials will be posted there, as well as changes and corrections to homework assignments.

## Getting help

Please make use of the TAs and their office hours. We, your instructors, are also glad to help. If something is confusing with the assignments, the fault is probably ours, and you're probably not alone. If you alert us early, we can probably fix it.

## Assignments and Grading

For the data analysis portion of the class, we will have both in-class work and assignments. You are allowed to work collaboratively on homework, but your write-ups must be done independently. Please also note everyone you worked with when turning in your assignments.

Use of automated coding tools like ChatGPT, Copilot, etc. is *allowed* but must be acknowledged in your solutions. We would like to know which tool and what prompt(s) you used, if any, to generate your solution.

Solutions should be submitted as saved Jupyter Notebooks on Colab. We'll tell you how to name and share these.

**Solutions are due before class on Tuesdays (3pm EST).** We can work with you if something unexpected occurs, but *we need to know in advance*. Please help us help you. We can't release solutions until we have everyone's assignments turned in.

Grading is on a pass/fail basis.

#### Thread I: Beginning Python

For the first several weeks of class, we'll be offering a crash course in basic Python covering *A Whirlwind Tour of Python*, transitioning to the *Python Data Science Handbook*. This will be basic and focused on students who have limited programming background. **This is purely optional.** Those of you who are already comfortable with programming do not need to attend, though you will be responsible for the material. We will also be working with the TAs to set up additional help during this period for those who would like it.

Date	Торіс	Reading	
1/11	Housekeeping, access-	WWTP Chs. 1–6	
	ing computing, advanced		
	Googling		
1/16	What can Python do?	WWTP Chs. 7-8	
1/18	Data structures	WWTP Chs. 9-12	
1/23	Patterns, iteration, duck	PDSH Ch. 1	
	typing		
1/25	NumPy and arrays	PDSH Ch. 2	
1/30	Dataframes	PDSH Ch. 3	
2/1	Plotting	PDSH Ch. 4	

# Thread II: Data Analysis (Systems)

This second thread of the course will cover five weeks and will focus on analyzing real neuroscience data sets.

John will not be lecturing. At least, not much. Most of what we'll

cover isn't really learned effectively that way, so we'll use our class time to complete programming and data analysis exercises that build on the basic Python knowledge you gained by reading *A Whirlwind Tour of Python*.

Each week, we'll do two sessions of in-class assignments, for which you'll be encouraged to work with a partner. The weeks are organized around both data and programming themes, and the inclass assignments often build on one another. After class is done for the day, we'll post links to solutions. Typically, we'll be walking you through an example analysis, with the goal of setting you up for the homework.

Date	Topic
2/6	Point process data I
2/8	Point process data II
2/13	Tabular data I
2/15	Tabular data II
2/20	Image data I
2/22	Image data II
2/27	Debugging I
2/29	Debugging II
3/5	Fitting basic models
3/7	Fitting simulated models

## Thread III: Data Analysis (Cellular & Molecular)

The third phase of the course will cover bioinformatics approaches to single-cell genomics data analysis. It will include an introduction to single-cell genomics methods, basics of R, as well as a tutorial of Seurat, a commonly used R package for single-cell data analysis. The tentative schedule is as follows:

Date	Торіс	Location
3/19	Single-cell genomics meth-	in person
	ods	
3/21	Basics of R	in person
3/26	Seurat: clustering	in person
3/28	Seurat: multiome data	in person
	analysis	
4/2	SCENIC+: gene regulatory	in person
	networks	
4/4	Monocle 3: trajectory analy-	in person
	sis	
4/9	Differential expression	in person
	analysis and visualization	
4/11	Basics of Python for bioin-	in person
	formatics	
4/16	Scanpy – Single-Cell Analy-	in person
	sis in Python	