Syllabus

Jeff Brabec and Montana Kay Lara

9/27/2021

# NGP-CMB Data-Science Seminar Syllabus

## Course Description and Motivations

This course is designed to introduce graduate students in the cellular, molecular, and neurosciences to basic data analysis and visualization with R. Taught in a seminar style with real-time coding prompts, we will move through the Tidyverse packages, overview statistical analysis, and use some more advanced packages for bioinformatics.

The monthly seminar will be accompanied by office hours where students can share their issues with one another, receive help from the instructors, or ask for help designing and conceptualizing an analysis. The course will run year-round with breaks during the major holidays. Each semester will have a project that students will complete to reinforce the skills and principles they are learning in the course.

Skills in data science are highly sought after in academia and beyond as computational analyses become the norm in basic science. This course is designed to be a continuous resource throughout a student’s time in their respective program. With the introduction of novel analysis techniques, community resources, and an accessible introduction into data science, students will have a stronger computational skill set to tackle their future scientific and career goals.

## Learning Outcomes

By the end of the academic year you will be able to:

1. Read-in, clean, and visualize data using R and the [tidyverse](https://www.tidyverse.org) family of functions
2. Design reproducible statistical analyses in R for a range of data-sets and data-types
3. Write simple functions to operationalize repetitive analyses
4. Troubleshoot specific coding questions or formulate detailed questions to ask in an online forum
5. Read, comprehend, and critically analyze computational methods

## Instructors

### Co-Directors:

Jeff Brabec

Montana Kay Lara

### Selected Topics for Future Guest Lecturers:

* Statistical Analysis
* Linear Models
* RNA-seq
* Machine Learning

## How to Succeed in This Course

There is no grade for this course. It is designed for students to maximize their time learning skills that are relevant to their research or career goals. Does your lab need to do an RNA-seq analysis and you need the quick nuts and bolts of what to do? Great, show up for those sessions and some office hours. If you want to build a broad data science tool-box to answer a multitude of questions and do a range of analyses then come to each session, ask questions, and engage with the broader UVM data science community. Learning a programming language on your own can be [hard](https://rfortherestofus.com/2019/09/my-r-journey-thomas-mock/) and we want to make it as painless as possible by providing you with the foundational skills you’ll need to succeed. This course will also foster a tight-knit community that can provide support when your code isn’t working, you’ve hit a roadblock, or need to employ new tools for a novel analysis.

## Required Materials

* You
* Computer
* RStudio and R (if unable to download on your computer for any reason we can set you up with an RStudio online account)
* Internet Connection

## Course Logistics and Schedule

Each seminar session will follow the same format. We will open with the presentation and live-coding examples of a new package/tool/principle lasting between 30-45 minutes. We will then spend the next hour addressing problems from the class. For the first few sessions, we will ask for students to submit issues to talk about, and as we progress, we will relax the structure to accommodate more discussion and coding with real data.

The Co-Directors will hold office hours every two weeks. This is to encourage students to work on their projects in between seminars and maintain active engagement with the material. The first semester schedule will occur as follows:

|  |  |  |  |
| --- | --- | --- | --- |
| Date | Session Type | Session Title | Instructor |
| 2021-10-01 | Seminar | Getting Started with ggplot | Jeff and Montana |
| 2021-10-05 | Office Hours | Office Hours 1 | Jeff and Montana |
| 2021-10-19 | Office Hours | Office Hours 2 | Jeff and Montana |
| 2021-11-05 | Seminar | Wrangling Data with dplyr | Jeff and Montana |
| 2021-11-09 | Office Hours | Office Hours 3 | Jeff and Montana |
| 2021-11-23 | Office Hours | Office Hours 4 | Jeff and Montana |
| 2021-12-01 | Seminar | Advent of Code/Making the Most of Online Resources | Jeff and Montana |
| 2021-12-07 | Office Hours | Office Hours 5 | Jeff and Montana |
| 2021-12-14 | Office Hours | Office Hours 6 | Jeff and Montana |