Bootcamp Info Sheet

Instructor

Name: *Ryan Ellison*

Bio: *My name is Ryan Ellison. I am currently a PhD-ABT at Ohio University in Neuroscience, with research specialties in Computational Neuroscience and Neurophysiology. My dissertation research focuses on the dynamics of slow biophysical processes in component neurons of a biological neural network (i.e., a CPG neural circuit that underlies rhythmic motor movement) and how those slow-evolving processes alter whole-cell properties resulting in stereotypical network output. My multi-disciplinary education and research afforded me the opportunity to join NASA, and then, the University Space Research Association, where I conducted space-based computational neuroscience research. I had the pleasure of working on the Artemis Mission as part of the Cross-Cutting Computational Modeling team for the Human Research Project, tackling a brain-behavior problem that results due to space-environment stressors. The overarching aim of this work was to help identify potential neural countermeasure targets, with the goal being to contribute to astronaut safety and mission success during long-term spaceflight. In all these endeavors, artificial intelligence/machine learning/data science have been integral in my research, whether tools utilized as part of a larger domain-specific project or the fundamental aspect of a project itself.*

Bootcamp Details

Bootcamp Title: *Introduction to R and Data Wrangling*

Number of Days: *4 days*

Hours per Day: *3 hours per day (11-2 eastern)*

Type of Instruction: *lecture with knowledge checks and hands on exercises*

Description: *This course introduces learners to the fundamentals of the R programming language. By the end of this course, learners will identify how data scientists use R, recognize basic data types and data structures, and perform basic calculations using foundational base R, load a dataset into their environment, working with functions that allow data to be selected, filtered, summarized, rearranged, and otherwise transformed according to analyst-vetted best practices.*

Target Audience: *Students with some experience working with data.*

Technologies: *R and R Studio. Packages that need to be installed:*

tidyverse

dplyr

nycflights13

box

Prerequisites: *While there are no prerequisites for this course, it is best suited to students with some experience working with data.*

Student References: *Class slides, class code, exercise files*

Bootcamp Syllabus

Day 1

- Discuss how programming is used across industries and define core functions of data scientist
- Identify stages of the data science control cycle
- Perform basic calculations and work with variables in R
- Identify good coding practices for clarity and reproducibility
- Distinguish data types

Day 2

- Identify basic data structures in R
- Read and write data in RStudio
- Load and clear variables in the R environment

Day 3

- Load and evaluate the dataset
- Address missing values in data

- Demonstrate installing a package and loading a library
- Define the six functions that provide verbs for the language of data manipulation, from the package dplyr
- Apply the filter function and arrange function

Day 4

- Select specific variables using the select command
- Derive new variables from the existing variables using the mutate and transmute commands
- Summarize columns using the summarise and group by functions
- Understand tidy data and its advantages
- Transform messy data to tidy data using tidyr package
- Manipulate columns by using the separate and unite functions