Course Director:
Garret D. Stuber, Ph.D.
Associate Professor
Departments of Psychiatry & Cell Biology and Physiology
Director, Neuroscience Curriculum

Instructor:
Vijay Namboodiri, Ph.D.
Postdoctoral Fellow, Stuber Lab

Teaching Assistant:
Randall Ung
MD/PhD student, Stuber Lab

Course Administrative Assistant:
Denise Kenney: denise_kenney@med.unc.edu

Meeting Times:
Fall Semester, 2017
Tuesdays and Thursdays 8:00am – 9:30
NRB room 3118

Course Short name and credits:
Neuroanalytics, 4 credits

Course Description:
The purpose of this course is to provide both practical and theoretical training in advanced data analysis approaches commonly used in neuroscience research. Over the past 10 years there has been a dramatic shift within the field from relatively simple data analysis approaches such as calculating means and standard errors of grouped data, to now performing complex analysis on higher dimensional datasets to uncover unappreciated features. The material in this course should be immediately useful to any student who is working with modern data collected in Neuroscience, from sequencing, electrophysiology, imaging, biochemistry, and behavior. The concepts in the course will be taught through programming in python. While understanding mathematical concepts behind analysis is important, we will largely focus on the big picture and try to illustrate concepts by emphasizing graphical representations of how datasets are treated with these approaches. Throughout the course, we will utilize real-world neuroscience data from a variety of sub-disciplines as examples, and also focus on teaching the implications and limitations of the approaches we cover. At the end of the course, students should have a solid foundation of scientific computing, which will prepare them to independently conduct analysis of their own data or prepare them for more advanced courses.

Prerequisites:
NEURO722/723 and some experience with basic statistics and data handling. No previous experience with programming is needed.
Limit on Enrollment and Priority for Registration:
In order to create a climate for active discussion and exploration of the material, enrollment will be capped to 15 students. All students are required to obtain permission from the Course Director to enroll. Priority is given to (1) students who have recently matriculated into the Neuroscience Curriculum, and to (2) graduate students from other departments or programs who are working in labs conducting neuroscience research.

Course Organization and Style:
Class periods will alternate between lecture style presentations and in-class programming examples. During the in-class demonstrations, students are expected to follow along on their own personal computers and interact with the instructors and other students to work through problems that arise.

Pre-class Requirements:
Python code academy tutorial
Installation of Anaconda and Python version 2.7 on your personal computer

Tentative Schedule:
Block 1: Introduction to Programming in Python (2.5 weeks)
- Installation and modules, basic file input/output
- Common data types (strings, integers, floating point, boolean logic)
- Common data objects (tuples, lists, variables, dictionaries, arrays)
- Basic syntax of Functions
- Loops and conditionals
- Array and matrix generation and manipulation (numpy, pandas)
- Plotting and data visualization (matplotlib, seaborn)

Block 2: Analysis and Statistics of Neural Data (5 weeks)
- Probability distributions and sampling
- Central limit theorem, bootstrapping, and randomization
- Basic parametric and non-parametric statistical tests
- Analysis pitfalls, and guidelines for multiple comparison corrections.
- Bayes Theorem and Bayesian Statistics

Block 3: Statistical Modeling and Machine Learning (5 weeks)
- General Linear Models
- Dimensionality Reduction (PCA/ICA)
- Data Clustering (k-means, more here)
- Data Classification (SVM, Naïve Bayes)

Block 4: Advanced Topic in Data Analysis (2.5 weeks)
- GUI development/interactive data analysis
- Cluster and Cloud Computing

Midterm projects will be assigned after block 2.
Student proposals for the final project will is due after block 3
Assigned Background Reading, Material, and Recommended Texts for Reference:
Text:
Automate the Boring Stuff with Python: Programming for Total Beginners by Al Sweigart, No Starch Press

Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython by Wes McKinney, O’Reilly Media, Incorporated. (recommended)

Python Data Science Handbook by Jake VanderPlas, O’Reilly Media, Incorporated. (recommended)

Additional background material will be provided in the form of video tutorials, and other related documents.

Assignments, Exams, and Grading:
Performance in the course is based on weekly homework assignments, a midterm project, a final project, and class attendance and participation. All work in the course will be submitted as coded iPython notebooks complete with comments explaining the analysis that is being performed. Datasets used for the analysis will be provided by the instructors, except for the case of the final project where students will be allowed to use their own data if they choose. Weekly assignments will focus on covering concepts that are discussed in class, while the midterm and final projects will focus on synthesizing material in order to build comprehensive data analysis pipelines.

Final projects for the course will be focused on developing data analysis pipelines that the students may be able to use in their own research. Students must first submit a short (1 page) proposal of their data analysis problem and what they intend to solve. Prior to the approval of the final project proposal, each student will need to meet with the instructors to discuss their objectives and for the instructors to inspect the dataset that the student intends to use for their project. The instructors will make suggestions of things to include in the final project in order to ensure each student will be utilizing approaches for statistical analysis and machine learning that are covered in the course.

Weekly homework assignments 40%
Midterm Project (assigned) 20%
Final Project (student proposed) 30%
Attendance and Class Participation 10%