Semester and Year: Spring 2017

Stony Brook University, Department of Physics and Astronomy
PHY 390, Data Analysis for Physics and Astronomy with Python
Course Instructor: Prof. Joanna Kiryluk, email: Joanna.Kiryluk@stonybrook.edu, office: Physics C109.
Course URL: http://skipper.physics.sunysb.edu/~joanna/Lectures/PHY390/

COURSE DESCRIPTION:
This course is an introduction to statistical data analysis with modern techniques, including the Python programming language on Windows computers. It will review concepts and methods to characterize experimental data such as averages, variances, standard deviations, errors and error propagation and will discuss probability distributions, confidence intervals, hypothesis testing, chi-squared minimization and straight line fitting. This course will also introduce students with no prior experience in programming to Python, with emphasis on practical data-centric applications. It will require extensive use of computers outside the classroom. The aim of this hands-on course is to prepare Physics and Astronomy majors for experimental laboratory work and research.

Credits: 3
SBC: TECH

Part A: Introduction to Data Analysis

1. Introduction: what is a measurement, random and systematic uncertainties
2. Data characteristics: distribution, mean and variance
3. Graphic representation of data: histograms, plots, linear and logarithmic scales
4. Statistics: binominal, Poisson and Gaussian probability distributions
5. Central Limit Theorem
6. The meaning of sigma
7. Partial differentiation, propagation of small uncertainties
8. Covariance and correlation
9. Least squares method
10. Combining results of different experiments, weighted averages
11. Straight line fit
12. Parameter and distribution testing and comparing results:
    test 3 sigma, chi-squared test, p-values, confidence levels

Part B: Python Programming for Data Analysis:

1. Python from scratch:
    a. Installation and setup
    b. IPython: An Interactive Computing and Development Environment
    c. Variables, basic math, types of data, input, print formatting and strings
    d. Decisions, loops, lists, functions, objects, modules
    e. Pandas, data structures
    f. Data files: input and output, file formats
    g. Data wrangling: Clean, transform, merge, reshape
    h. Plotting and Visualization

2. Data analysis modules
    a. NumPy Basics
    b. SciPy Basics

3. Data analysis report
    a. Latex

COURSE LEARNING OBJECTIVES:
Include course objectives. If this course is approved to satisfy D.E.C. and/or the SBC, the objectives **must** address how the course outcomes relate to the specified D.E.C or SBC category. See the DEC descriptions in the Bulletin. Include a brief description of the opportunities this course would provide for students to acquire the knowledge or skills necessary to achieve the course learning outcome(s)

The student learning objective is to understand technology by gaining a hands-on experience in data analysis with modern computer-based tools, such as the Python programming language, with the goal to prepare Physics and Astronomy majors for experimental laboratory work and research. Students will become proficient at working with data in Python by learning the fundamentals of the language to write computer programs to solve data analysis problems. Specifically they will learn to solve data analysis problems using quantitative statistical methods and will learn how to obtain, and present results quantitatively and graphically. This hands-on course is ideal for freshman and sophomore Physics and Astronomy majors new to Python.

COURSE REQUIREMENTS:

Attendance and Make Up Policy
-- Instructors may enforce their own course attendance policy (the university has no standard policy except the policy regarding absence due to university sanctioned events … (e.g., athletic games; Stony Brook in Albany)

Attendance is mandatory.

-- Policy for late work and tardiness

Homework: No late work will be accepted. Any (serious!) excuses (medical or otherwise) are to be documented and discussed with the instructor in a timely manner. Homework must be turned in by the date and time on the assignment, which will typically be scheduled during a lecture period. If you cannot make the lecture, you can bring your assignment to the instructor's office before the due date/time.

Description and schedule of Required Readings and/or Assignments.

Text books:
1. "A Practical Guide to Data Analysis for Physical Science Students -1st edition", L. Lyons, Cambridge University Press, 1991.
2. "An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements -2nd Edition", John R. Taylor, University Science Books, 1996.
3. "Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython -1st Edition", W. McKinney, O'Reilly Media, 2012.

Homework: homework problems will be posted weekly.

Exams
Number and Description of Exams (include dates in meeting schedule section below).

There will be two midterm exams, one on data analysis methods (Part A), and one on Python programming (Part B). Midterm exams will be given during regular lecture hours.

There will be one (final) take home project. Students will analyze a data set provided to them by using data analysis techniques, writing a Python computing program and writing a report.

GRADING:
Describe the components of the course, and indicate how each component of the course will be factored into the final course grade; often expressed in %
Indicate alphanumeric breakdown of grades or explicit description of the grading technique (i.e., description of the "curve," or A= 95-100, A- = 90-95, or whatever the instructor defines)

Final grades will be determined by weighting the various portions of the course as follows:
- 40% two midterm exams (20% each)
- 30% homework
- 30% final "take home project": analyze a provided data sample and write a report.

Grades:
- Weighted average: 90 % - 100 % grade A
- Weighted average: 85 % - 90 % grade A-
- Weighted average: 80 % - 85 % grade B+
- Weighted average: 70 % - 80 % grade B
- Weighted average: 65 % - 70 % grade B-
- Weighted average: 60 % - 65% grade C+
- Weighted average: 50 % - 60 % grade C
- Weighted average: 45 % - 50 % grade C-
- Weighted average: 40 % - 45 % grade D+
- Weighted average: 35 % - 40 % grade D
- Weighted average < 35% grade F

All grades will be accessible via the blackboard for this course.

MEETING SCHEDULE
The class will meet twice a week. There will be one 80-minute long session on Data Analysis (Part A, TUE 8.30am-9.50am, S-235) every week, and one 80-minute long session on Python programming (Part B, THU 8.30am-9.50am, S-235) every week.

CLASS PROTOCOL
No cell phones or other "smart" devices with phone or WiFi capability will be permitted in the exams.

The University Senate Undergraduate and Graduate Councils have authorized that the following required statements appear in all teaching syllabi (graduate and undergraduate courses) on the Stony Brook Campus. See also http://www.stonybrook.edu/provost/facultyinfo/Syllabus%20Statement.doc

DISABILITY SUPPORT SERVICES (DSS) STATEMENT (must be the following language)
If you have a physical, psychological, medical or learning disability that may impact your course work, please contact Disability Support Services, ECC (Educational Communications Center) Building, room128, (631) 632-6748. They will determine with you what accommodations, if any, are necessary and appropriate. All information and documentation is confidential.

[*In addition, this statement on emergency evacuation is often included, but not required:*
Students who require assistance during emergency evacuation are encouraged to discuss their needs with their professors and Disability Support Services. For procedures and information go to the following website: http://www.stonybrook.edu/ehs/fire/disabilities]

ACADEMIC INTEGRITY STATEMENT (must be the following language as approved by the undergrad council):
Each student must pursue his or her academic goals honestly and be personally accountable for all submitted work. Representing another person's work as your own is always wrong. Faculty are required to report any suspected instances of academic dishonesty to the Academic Judiciary. Faculty in the Health Sciences Center (School of Health Technology & Management, Nursing, Social Welfare, Dental Medicine) and School of Medicine are required to follow their school-specific procedures. For more comprehensive information on academic integrity, including categories of academic dishonesty, please refer to the academic judiciary website at http://www.stonybrook.edu/uaa/academicjudiciary/

CRITICAL INCIDENT MANAGEMENT (must be the following language as approved by the undergrad council):
Stony Brook University expects students to respect the rights, privileges, and property of other people. Faculty are required to report to the Office of Judicial Affairs any disruptive behavior that interrupts their ability to teach, compromises the safety of the learning environment, or inhibits students' ability to learn. Faculty in the HSC Schools and the School of Medicine are required to follow their school-specific procedures.