AST 5765: Advanced Astronomical Data Analysis

Fall 2016 Syllabus
Prof. Joseph Harrington

1. Course Vitals

Room:	HPA1 246	Lecture:	TR 10:00 – 11:50
Grading:	ABCDF w/ +-	Credits:	3(3,1)
Dates:	23 August – 1 December 2016	Final:	Project due 7 Dec 13:00
Class URL:	http://planets.ucf.edu/academics/ast5765/		
Class directory:	thneed.physics.ucf.edu:/home/ast5765		
Textbooks:	Howell, S. B. 2006. *Handbook of CCD Astronomy,* 2nd Ed. Cambridge, ISBN-13: 978-0521617628. Bevington, P. R., and D. K. Robinson 2002. *Data Reduction and Error Analysis for the Physical Sciences*, 3rd Ed. McGraw Hill, ISBN-13: 978-0072472271.		
Resources (short assignments in Press):	Press, W. H., *et al.* 2007. *Numerical Recipes: The Art of Scientific Computing,* 3rd Ed. Cambridge, ISBN-13: 978-0521880688. Lantangen, H. P., 2008. *Python Scripting for Computational Science*, 3rd Ed. Springer, ISBN-13: 978-3540739159.		
Prerequisites:	MAC 2313 (calculus) Ability to write simple computer programs An upper-level course in astronomy or planetary science, or CI		

Job:	Instructor	Teaching Assistant
Name:	Joseph Harrington	Ryan Challener
Office:	PSB 441	PSB 316
Cell phone:	(321) 696-9914	(814) 720-1304
Email:	jh@physics.ucf.edu	rchallen@knights.ucf.edu
Office hours:	Fri 12:00-1:00 pm in PSB 441	Mon 10:30-11:30, Fri 2-3 in PSB 355

Contacting: If computer or other emergency that prevents you from working, call Ryan, then Joe. Else, email or text. Email preferred, but text may get a quicker response. Please state name and class in texts.

2. Objectives

Those who successfully complete this course will be able to:

1. Understand basic statistics and error analysis as used in the physical sciences,
2. Extract physical measurements and error estimates from raw data,
3. Find, educate themselves about, and select appropriate numerical analysis methods,
4. Fit a theoretical model to the measurements,
5. Draw scientifically-valid conclusions from the measurements,
6. Manage and carry out online work with large amounts of data, and
7. Present scientific results.

3. Approach

We cover the following topics in roughly this order:

1. Computers, programming, online management.
2. Introductory statistics and modeling.
3. Array detectors and corrections, image analysis.
4. Photometry and Spectroscopy.
5. Fitting.
6. Project.

In addition, throughout the course, we will read about and discuss a series of advanced topics in computational data analysis.

Data are information, and computers are information manipulators. You should therefore embrace the power of computation and become as expert in computing as in your chosen field. This class approaches the computer as a tool to master and command rather than a black box to fear. Almost all assignments are online. We provide an online environment for student use, but students must arrange their own access to it.

4. Format and Grading

Lecture attendance is mandatory. We encourage you to **bring a laptop to class** and to be an active learner by doing the demos and exercises along with the instructor; don't just watch. Preparation for and participation in class discussion counts toward the final grade. The homework assignments are due at the beginning of class on the due date. **No late work will be accepted**, because we discuss the assignments in class, so PLAN AHEAD. Reading should be done **before** the class indicated.

We will use peer evaluation (not grading) of homework assignments. Your participation in the peer evaluation counts toward your grade.

We use the Python language. No Python experience is required, but students will need to become functional in Python within the first two weeks of the course, so **general programming ability is required** for success in this course. Students who have taken the course without programming experience have struggled a great deal to succeed.

Evaluation weighting for final grade:	
Homework	40%
Discussion participation	5%
Peer evaluation participation	5%
Quizzes	10%
Project results	20%
Project execution	10%
Project paper	10%

This is the graduate version of this class. It meets concurrently with the undergraduate version, but has additional problems on homework assignments, a more challenging project, and additional readings assigned during the semester, such as optimal photometric and spectral extraction, spectrum and time-series convolutions, interpolation methods, and wavelet analysis. Undergraduates with programming and data analysis experience may register for the graduate version, with permission of the instructor.

To encourage group participation, grades will not be curved. It is possible for everyone to get an A. It is also possible for everyone to fail (but I hope not!). All reasonable questions regarding grading are welcome, but pure negotiation is not.

5. Academic Honesty, Sharing, and Information Sources

We will follow the letter and spirit of the UCF Golden Rule. Research in astronomy and physics relies on taking advantage of resources developed elsewhere: software libraries, descriptions of methods, etc. *Unless we state otherwise*, please use such external sources in your work. However, there are several conditions:

1. All math, code, and text answers must be your original work. You may (and should) discuss the relevant general topics with each other, but you may neither give nor receive specific help on nor share assigned work. Sharing code, even to the extent of making it visible on-screen or reading it from someone else's screen, is not allowed, regardless of whether the other party gives permission. This includes help from others not in the class.
2. For coding problems, the portion of the answer relevant to the problem must be your original work. For example, if the question asks you to subtract two images, you must write the code to do the subtraction but you may use third-party code to read the images from files.
3. You must have legal permission to use an external source (assumed if publicly posted).
4. You **MUST** give credit to all external sources **on a problem-by-problem basis**. Credits must include the name of the item, a sentence fragment describing it if it is not obvious from the name, its author(s), year of authorship, and location (e.g., the name, volume, and pages of a journal article, or the URL of a software package distributed online).
5. As with any scientific research project, you alone are responsible for the output: if you download a package that claims to do something and it has a bug that gives the wrong answer, the answer is wrong and you will be marked accordingly.
6. Work you did prior to the start of the course may not be handed in for grade (talk to the instructor for exceptions).

6. Working Effectively

There will be approximately weekly homework assignments and project work. It is critical that you do the homework and readings by the beginning of class on the due date, as we discuss answers in class. Your personal understanding is what counts in the discussions, and discussions count toward your final grade. Since answers will be discussed in class, **no late homework will be accepted**.

The **homework solutions** are a critical teaching element. **You must read them**, including all comments in code, each week. They contain discussion of how and why to do important things, knowledge you will need for later assignments.

Not all information needed to do the assignments is covered in lecture or the texts. You must find sources for and read about programming and numerical methods on your own. The Recommended Textbooks are a good place to start, as is the web.

Compared to most physics courses, this course is heavy on skills, methods, and experience. These are taught with practice on real data in the homework assignments. Unlike the "example" datasets of other courses, real data has real problems with it that you will have to overcome. You should budget significant time each week to work on your homework and project. Assignments will depend heavily on prior work done in the class, so skipping work is not very useful: you'll be doing the work anyway in order to do later assignments, so it makes sense to do it in time to get credit for it. Remember that debugging can take a long time, so start your assignments early! While time spent on the class varies a great deal according to students' prior programming experience, you should expect to spend an average of 6-10 hours per week outside of class on this course, and more if this is your first exposure to computing “for real”.

7. Homework Format

Hand in all homework electronically, unless otherwise stated, including prose and math.

Almost all assignments are programs. Put the answers to the assigned problems in a single, **executable**, cleanly-coded and commented, Python source code file named, for example, `hw3_jh.py`. This is your “main homework file.” It should contain nothing but your homework answers; it is not a log file. If you write functions, whether or not they are specifically assigned, put them in separate, non-executable Python source files with reasonable names (e.g, `gaussian.py`).

In comments at the top of the main homework file, indicate your name, the course number, the assignment name (e.g., `HW3`), and the date. Put the problem numbers in comments as well. Be sure to comment out any answers that are not commands, such as justifications or explanatory notes.

Use plain ASCII text files wherever possible, and certainly for all source code and data tables. Other allowed formats: PDF (for writing), FITS (for data), PNG (for images of plots), JPEG (for scans or other photographic images). You can convert MSWord to PDF by loading into `ooffice` and clicking “export PDF”; check the output before handing it in. Math (only) may be handwritten on paper, scanned, and handed in as a PDF or JPEG, though a typeset PDF answer is preferable. **Grammar, spelling, and complete sentences** count for grade, even in math (remember that “=” is a verb). **Math** problems must show your logic and calculations. **Box or circle** final math answers.

If the problem asks for a **plot or image display**, include commands for output to a file as part of the program, and commands to put the plot on screen in comments. Include the files it makes in your directory when you hand it in (see below). Plots should have titles and sensible axis labels, including units. Put each item in a separate file. The filenames should follow the format: `hw3_jh_prob2_plot1.png`. *Only if requested,* put ASCII output to the screen (like tables) in files named like `hw8_jh_prob2_table1`. ASCII tables should have titles and column headers that distinguish them from one another and that make sense to the reader. You may hand-edit headers onto tables written by the computer.

Handing in homework: Make a directory called `~/handin/hw3_jh` on the course computer, substituting the right assignment number and your username. **Before class**, put the files you wish to hand in in that directory. We will copy all the assignment directories at the beginning of class automatically. **No late homework will be accepted,** so be sure you **actually *save* your files** before class starts! Do not email your homework. Also, make sure all necessary plot and program files (such as functions you are asked to write) are in the directory. **We grade by running your program from the shell** after copying just the `.py` files to a separate directory. If it produces errors or fails to produce the assigned output (plot files, etc.), the assignment is incomplete. The best way to ensure a complete assignment is to do this yourself and check that everything works and that it produces all the files it needs to.

8. Project and Advanced Readings

In October, each student will start a final project based on real data. You will apply the methods learned in the course to produce a measurement and reach a scientific conclusion. Three components of the project together contribute 40% of your final grade: a paper, which will follow the format of the *Astrophysical Journal*, what your coded analysis routines produce (results), and how well your analysis routines are coded and documented (execution).

In addition to the lecture schedule below, there will be readings in numerical analysis from Press *et al*. and other sources. Topics may include the following list, with additional readings determined in part by class interest and the needs of students:

- Optimal photometric extraction (handout)
- Optimal spectrum extraction (Horne 1986, *PASP* **98**, 609-617)
- Monte Carlo error analysis (Press *et al*.)
- Robust estimation (Press *et al*.)
- Lomb-Scargle periodogram (Press *et al*.)
- Wavelet analysis (Torrence and Compo 1998, *BAMS* **79**, 61-78)

9. Departmental Policies

Missed Work Policy

It is Physics Department policy that making up missed work will only be permitted for University-sanctioned activities and bona fide medical or family reasons. Authentic justifying documentation must be provided in every case (in advance for University-sanctioned activities). At the discretion of the instructor, the make-up may take any reasonable and appropriate form including, but not limited to, the following: a replacement exam, replacing the missed work with the same score as a later exam, allowing a “dropped” exam, replacing the missed work with the homework or quiz average.

NOTE: Those unable to attend class in person on a particular day may, by arrangement with the instructor, attend via the internet. This is intended mainly to handle mild flu cases and should not be used when it is physically possible for the student to attend class.

Disabilities Policy

The University of Central Florida is committed to providing reasonable accommodations for all persons with disabilities. This syllabus is available in alternate formats upon request. Students with disabilities who need accommodations in this course must contact the professor at the beginning of the semester to discuss needed accommodations. No accommodations will be provided until the student has met with the professor to request accommodations. Students who need accommodations must be registered with Student Disability Services, Student Resource Center Room 132, phone (407) 823-2371, TTY/TDD only phone (407) 823-2116, before requesting accommodations from the professor.

Establishing Academic Activity For Financial Aid

All instructors/faculty are required to document students' academic activity at the beginning of each course. In order to document that you began this course, please be present in class, where attendance will be taken, and/or complete the first homework assignment, by the end of the first week of classes or as soon as possible after adding the course. Failure to do so may result in a delay in the disbursement of your financial aid.

10. Disclaimer

This syllabus is a guideline, not a contract. The instructor may alter it at any time.

11. Schedule

This is an approximate schedule; pace adjusts to the needs of the class. Homework and reading are due as assigned on the homework handouts, not here. The extensive narratives in the homework solutions and demos are also required reading. Quizzes may happen at any time.

Lecture	Date		Topic	Reading	Assignment
			Tools and Theory		
1	23 Aug	T	Introduction	Handout	
2	25 Aug	R	Learning Linux and Python	Handout	1 (Unix)
3	30 Aug	T	Shell, Python	SciPy docs	
4	1 Sep	R	Python Data and Programming	Handout	
5	6 Sep	T	Python Arrays and Functions	Matplotlib tut., Pydatatut 1-2, readings on handout	2 (SciPy)
6	8 Sep	R	Python Functions, Coding, FITS		
			Photometry		
7	13 Sep	T	Measurement, Prob., and Stats	Bev. Ch 1,2	3 (programming)
8	15 Sep	R	Probability Distributions	Handout, Pydatatut 3-4	
9	20 Sep	T	Fitting and Astron. Measurements	Bev. Ch 3,4	4 (stats)
10	22 Sep	R	CCD Arrays, Array Corrections		
11	27 Sep	T	IR Arrays, Sky, and Corrections	Bev. Ch 6	5 (fitting)
12	29 Sep	R	Combining Frames	How. Ch 1,2, Press 14.1-14.3, 14.8, Pan-Starrs	
13	4 Oct	T	PSF and Photometry	Press 15.6, 15.8	6 (S/N, 2D Gaussian)
14	6 Oct	R	Bad Pixels	How. Ch 3-4.5	
			Spectroscopy		
15	11 Oct	T	Introduction to Spectroscopy		7 (setup, dark)
16	13 Oct	R	Spectrograph Optics	How. Ch 4.6-5	
17	18 Oct	T	Presenting Research (DPS)	How. Ch 6	8 (sky)
18	20 Oct	R	Error Analysis Review (DPS)	How. Ap C	
19	25 Oct	T	Spectrum Calibration	Horne paper	9 (flat field)
20	27 Oct	R	Spectroscopy Applications, Project		
			Project		
21	1 Nov	T	Interpolation		10 (centering)
22	3 Nov	R	Convolution and Correlation		
23	8 Nov	T	Bad Pixels in Practice		11 (photometry)
24	10 Nov	R	General Model Fitting		
25	15 Nov	T	Model Parameter Error Estimation		12 (S/N, project uncertainties)
26	17 Nov	R	Model Parameter Error Estimation		
27	22 Nov	T	Detector Physics Review		(Project)
	24 Nov	R	Thanksgiving day, **no class**		
28	29 Nov	T	Optimal Extraction		(Project)
29	1 Dec	R	Project help or review session		
	6 Dec	T	Project help session 10-12:50	Exam period	(Project)
	7 Dec	W	(in exam week)	(no meeting)	Final Project Due 1:00 pm
	13 Dec	T	Project discussion 11:30 – 1:30	(optional)	