AST 5XXX: Data Analysis

Spring 2007 Syllabus
Joseph Harrington

1 Course Vitals

Room:	Physics Computing Classroom Trailer	Lecture:	TR 3:00 – 4:20
Grading:	ABCDF	Credits:	3
Dates:	8 January – 23 April 2007	Final:	project
Class URL:	http://physics.ucf.edu/~jh/ast/ast5xxx/ast5xxx.html		
Class directory:	/home/ast5xxx on all trailer machines		

Textbooks:

Howell, S. B. 2000. *Handbook of CCD Astronomy*. Cambridge, ISBN 0-521-64834-3.
Bevington, P. R., and D. K. Robinson 2003. *Data Reduction and Error Analysis for the Physical Sciences,* 3rd Ed. McGraw Hill, ISBN 0-07-247277-8.
Oliphant, T. 2005 *Guide to NumPy*. At `http://tramy.us`

Job:	Lecturer
Name:	Joseph Harrington
Office:	MAP 420
Phone:	(407) 823-3416
Email:	jh@physics.ucf.edu
Hours:	W 2–3

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. Fit a theoretical model to the measurements,
4. Draw scientifically-valid conclusions from the measurements,
5. Manage and carry out online work with large amounts of data, and
6. 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. Measurement extraction (example: stellar photometry).
5. Fitting.

6. Spectrographs and spectroscopy.
7. Project.

4 Format

The class meets twice weekly in the computer trailer. Lecture attendance is mandatory. The weekly homework assignments are due Thursday at the beginning of class. **No late work will be accepted**, so PLAN AHEAD. Reading should be done **before** the class indicated: preparation for and participation in class discussion counts toward the final grade.

We use the SciPy Language. No experience is required, but students will need to become functional in SciPy within the first few weeks of the course.

Evaluation weighting:	
Homework	40%
Discussions	10%
Project results	25%
Project execution	15%
Project paper	10%

5 Grading

To encourage co-operation and 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.

6 Academic Honesty

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 and text answers must be your original work. You may (and should) discuss the relevant general topics with each other, but you may not give specific help on or share assigned work.
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 may not use the work of other students in the class, even if they wrote it long ago.
4. You must have legal permission to use an external source (assumed if publicly posted).
5. 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).
6. 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.

7. Work you did prior to the start of the course may not be handed in for grade (talk to the instructor for exceptions).

7 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**.

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. You should budget significant time each week to work in the lab on your homework and project. You have priority access to the lab machines for the semester. 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.

8 Homework Answer Format

All homework questions are **electronic assignments**, unless otherwise stated, including prose answers. Math (only) may be handed in handwritten on paper, if necessary, but we prefer electronic formats. Grammar, spelling, and complete sentences count for grade, including in answers that involve math (remember that "=" is a verb). **Math** problems must show your logic and calculations. **Box or circle** final math answers.

Use plain ASCII text files wherever possible, and certainly for all program text. Other allowed formats, in preference order: FITS (for data), PDF, PS, TeX/LaTeX, Excel, MSWord (the latter two only if they work in `ooffice`). You can convert MSWord to PDF by loading into `ooffice` and clicking "export PDF".

Handing in homework: Make a directory called `~/handin/hw3-jh`, 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.

For problems that do not specify the creation of new files, hand in a file of commands named, for example, `hw3-jh.py`. This should be a cleanly-coded, presentable file of commands, not a log file with other notes (note the name difference between this and a log file). In comments indicate your name, AST 5xxx, the assignment name, the date, and the problem numbers. We will be running these as batch files, so be sure to comment out any answers that are not commands. Also make sure all necessary program files are in the directory.

If the problem asks for a plot or image display, include commands for **both** on-screen display and output to a PostScript file. Include the files it makes in your directory. 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.ps`. *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.

9 Project

In March, each student will choose 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. There are several projects to choose from, or you may request permission to analyze some other data for which an analysis does not yet exist. Three components of the project together contribute half 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).

10 Spring 2004 Schedule

Date		Topic	Reading	Assignment
		Tools and Theory		
Jan 9	T	Introduction, computer accounts	Lab doc	
Jan 11	R	Unix/SciPy basics	Handout	HW1 (Unix)
Jan 16	T	SciPy graphics, image I/O	SciPy docs	HW2 (SciPy)
Jan 18	R	SciPy Programming, FITS data	Handout	
Jan 23	T	Measurement, Probability	Bev. Ch 1,2	HW3 (programming)
Jan 25	R	Error Analysis	Bev. Ch 3,4	
		Optical/Infrared Photometry		
Jan 30	T	Fitting	Bev. Ch 6	HW4 (stats)
Feb 1	R	Array Detectors		
Feb 6	T	Detector Systematics	How. Ch 1,2	HW5 (fitting)
Feb 8	R	Infrared Arrays		
Feb 13	T	Sky and Flat Field Frames	How. Ch 3-4.5	HW6 (S/N, 2D Gaussian)
Feb 15	R	Flat Fields, FITS, Interpolation		
Feb 20	T	Finding & Fixing Bad Pixels	How. Ch 4.6-5	HW7 (dark & sky)
Feb 22	R	PSFs, Aperture Photometry		
Feb 27	T	PSF-Fitting Photometry		HW8 (flat field, bad pixels)
Mar 1	R	Atmospheric Absorption, Std. Stars	How. Ap C	
		Optical Spectroscopy		
Mar 6	T	Introduction to Spectroscopy		HW9 (photometry)
Mar 8	R	Spectrum Processing	How. Ch 6	
		Spring Break		
Mar 20	T	Wavelength Calibration		HW10 (spectrographs)
Mar 22	R	Line Profiles and Blends		
Mar 27	T	Interrogating Spectral Lines		HW11 (wavelength)
		Project		
Mar 29	R	Time Series Analysis	Handout	
Apr 3	T	[Another DA example from physics]	Handout	Project
Apr 5	R	["]		
Apr 10	T	Fourier Transforms	Handout	Project
Apr 12	R	Fast Fourier Transform		
Apr 17	T	FFT in Practice	Handout	Project
Apr 19	R	Project presentations		
Apr 26	R	(in exam period)		Final Project Due