ASTRONOMY 3190: Intro to Research
Spring 2020
10:30-11:20 MWF Dale Hall Tower Rm 0105

Instructor:	Prof. Karen M. Leighly
Department:	Physics and Astronomy
Office:	243 Nielsen Hall
Phone:	325-7045
Voice Mail:	325-7045
Email:	leighly@ou.edu
Office Hours:	TBD
TA / Grader:	Daniel Vagie
Learning Assistant	Collin Dabbieri

TEXT: No text required; however, you will need to bring a laptop or suitable substitute to class every day.

Learning Management System: Canvas – canvas.ou.edu – access notes, HW assignments, submit HW, and grades.

Course Description: Prerequisite – Freshman physics courses (PHYS 1205, PHYS 1215) or equivalent

Suggested Preparation: Sophomore or junior course level in Physics and Astronomy courses. No astronomy preparation will be assumed, so suitable for Physics and Engineering Physics majors.

Why Study Data Analytics?

Astrophysics is the application of physics to astronomical objects. We learn about astronomical objects by collecting data from them. Unlike many branches of physical science, astronomy is purely observational; while we can choose a time to observe an object, and the instrument to observe it with, we cannot interact with the object to make it do different things. Thus, astrophysicists are at the mercy of the data collection.

To appreciate the impact of this limitation, imagine that you want to determine the physical properties of a quasar during the early universe. An atom in some gas in that quasar may emit a photon. In most cases, we will not detect that photon directly. The energy of the photon may be changed. It could be scattered and reprocessed, or suffer a doppler shift (physical reasons) but the energy could be changed by the detector, or the detector may not capture that photon because it is not sensitive to photons of that energy. Assuming it is detected, for statistical considerations, almost nothing can be learned from a single photon (in most cases); a few photons (at least) must be measured to assign uncertainties. After collecting a bunch of data (data wrangling), you need to look at it (exploratory data analysis and visualization). Then, you may need to describe the data mathematically (modeling), and there is generally no equation or software package that will extract the physical properties from your data; you will need to be able to write your own (computation). Science doesn't happen in a vacuum, so the final step is usually communicating the results to your peers.

It is also appropriate to appreciate that data analysis is not limited to astronomy or even the physical sciences. Data is becoming increasingly important in the modern world, and it is critical that STEM students have data analytics skills. Some projections predict 2.7 million new data science positions by 2020, and a recent poll showed that 69% of employers expect candidates with data science and analytics skills to get preferences for jobs (Data Science for Undergraduates Opportunities and Options, The National Academies Press).

This class is structured to help you understand the "data life cycle", i.e., process of data analysis, as well as gain skills necessary to perform the individual steps. For the purpose of this class, the data life cycle includes the following:

> **Data Wrangling:** A catch-all category that includes data mining and exploratory data analysis.
>
> **Visualization:** Data visualization is the graphic representation of data.
>
> **Statistical Thinking:** Statistical thinking is, roughly speaking, the recognition of the fact that all data is influenced by statistics.
>
> **Modeling:** In the context of this class, this category includes the fitting of physical or empirical models to data, and how to evaluate the goodness of fit, but it also includes construction of theoretical models.
>
> **Computational Thinking:** Computational thinking involves expressing problems and their solutions in ways that a computer could execute.
>
> **Communication Skills**: Scientific results must be communicated to peers and the public. Technical writing is rather different than the type of writing used in composition classes.

Science Content of the Class

We analyze data to solve astronomical problems. Therefore, the data analytics skills addressed in this class will be embedded in a particular scientific problem: the evolution of galaxies. Specifically, we will use data, computations, and simulations to explore various aspects of this problem. The topics and data analytics concepts that we will look at and use follow:

Topic	Data Analytics Concepts
Properties of Stars	Computation Model Fitting Visualization
The Distribution of Stars	Statistical Thinking Data Wrangling
Colors of Galaxies	Statistical thinking Computation
Expansion of the Universe	Model fitting Statistical thinking
Galaxy Evolution	Data wrangling Visualization Statistical thinking
Masses of Galaxies I	Data Wrangling Computation
Masses of Galaxies II	Data Wrangling Computation Visualization
Black Hole Masses I	Computation Statistical Thinking Visualization
Black Hole Masses II	Model fitting Statistical thinking
Galaxy Dynamics Simplified	Computation
Galaxy Dynamics and Star Formation	Data Wrangling Statistical thinking
Cosmological N-body Simulations	Computation Visualization
The Subgrid	Computation Modeling

Learning Goals:

Through the above exercises and related assignments, homework, and project, the students will attain the following benchmarks in the data life cycle categories:

Data Wrangling: Students will be able to access astronomical data from on-line catalogs, either directly or using simple SQL queries. Students will be able to use plots and statistical methods to compare data sets.

Visualization: Students will be able to construct tables, plots, and other graphics, both for exploratory data analysis and for publication.

Statistical Thinking: Students will understand the role of uncertainties in model fitting. Students will demonstrate an understanding of the role of distributions in producing astronomical phenomena.

Modeling: Students will be able to describe least-squares fitting of data with a model, and will be able to assess the quality of a model fit. They will gain facility using Sherpa.

Computational Thinking: Students will be able to write python functions to do computations. They will be able to make sensible choices for axes. Specific expertise in convolution and solving equations of motion.

Communication Skills: Students will be able to construct a properly documented python notebook. They will be able to employ latex to express equations in markdown.

Assessed Activities:

Assessment in this class will consist activities in class, activities completed outside of class before the next class meeting, quizzes, homework, and a project, as follows:

Activity	Fraction of grade
Assignments and in class-activities	25%
Quizzes	25%
Homework	25%
Project	25%

In-class Activities and Assignments. This class is all about development of skills, so there will be many in-class activities, and many of them will be handed in and assessed. There will also be numerous short activities assigned at the end of a class, and due at the beginning of the next class. Both of these activities will help you learn because will need to be engaged with the material continuously (rather than, e.g., waiting until you have a homework set due to try to understand what is going on.). In order to compensate for missed classes due to illness or other missed classes, we will drop the lowest 10% of this category.

Quizzes. We will have short (15-20 minute) quizzes approximately 10 times during the semester. These will be on one or two recently-discussed topics, and generally will have an essay or short answer format. In order to compensate for missed classes due to illness or other missed classes, we will drop the lowest 10% of this category.

Homework: While the assignments and quizzes will focus on one or two concepts, the homework will be designed to use several concepts to solve a problem. Homework will be assigned approximately every two weeks. Late homework will incur a 10% penalty per day. Exceptions and extensions will be made under certain circumstances (e.g., illness).

Project: The project is the capstone for the course. Typically, a project will allow you to explore to a greater depth any of the concepts or science that we cover in the course. A list of projects with guidelines will be distributed after the midpoint in the course, and the last two weeks of the course will be set aside to work on the project. Students are allowed to work in groups of two or three, but they may also work alone. Note that the amount of work in each project must be proportional to the number of students in a group. A project writeup will be required, and the projects will be presented during finals week.

Computation: Python is the programming language that we will use for this class. To make things easier, we will use http://www.sciserver.org, which provides an on-line interface to Jupyter notebooks and SDSS data. Typically you will construct your notebook on the server, and then download to hand in via Canvas.

Collaboration: Students are permitted to collaborate on all aspects of this course except the quizzes. However, each student will be required to hand in his or her own work. In particular, written discussions and code must be your own words. Substantial similarities between assignments or homework sets will be flagged. The first instance will incur a warning. The second instance will be sent to over the Academic Integrities.

Class participation: Attendance is required for this class. Illness or other absence may be excused with appropriate documentation. On a related topic, while I encourage questions and discussion in class, conversations among students, cell phone activities, and other annoying behavior will not be tolerated.

Grading: I will use a fixed grading scale in this class, but will retain the option to curve down. For example, if you earn 90% of the points, you will automatically earn an A, but at the end of class, it may turn out that people with between 88-90% also earn an A.

Grade	Automatic	Criterion
A (or S+)	90-100%	Excellent mastery of all material
B (or S)	80-90%	Good mastery of most material, excellent mastery of some.
C (or S-)	70-80%	Most material mastered, with some significant lacks.
F	<50%	Fewer than 50% points earned

Disabilities: Any student in this course who has a disability that may prevent him or her from demonstrating his or her abilities should contact me personally as soon as possible so that we can discuss accommodations necessary to ensure full participation and facilitate your educational experience.

Religious holidays: It is the policy of the University to excuse absences of students that result from religious observances and to provide without penalty for the rescheduling of examinations and additional required classwork that may fall on religious holidays. Students that plan to observe a religious holiday should notify me ahead of time (at least one week; preferably at the beginning of the semester) so that appropriate rescheduling can be made.

Cheating: Students are required to understand and adhere to standards of scholarship from the time they are admitted to the university.
Please read information about academic misconduct at OU at http://integrity.ou.edu.

Adjustments for Pregnancy/Childbirth Related Issues

Should you need modifications or adjustments to your course requirements because of documented pregnancy-related or childbirth-related issues, please contact me as soon as possible to discuss. Generally, modifications will be made where medically necessary and similar in scope to accommodations based on temporary disability. Please see www.ou.edu/content/eoo/faqs/pregnancy-faqs.html for commonly asked questions.

Title IX Resources

For any concerns regarding gender-based discrimination, sexual harassment, sexual misconduct, stalking, or intimate partner violence, the University offers a variety of resources, including advocates on-call 24.7, counseling services, mutual no contact orders, scheduling adjustments and disciplinary sanctions against the perpetrator. Please contact the Sexual Misconduct Office 405-325-2215 (8-5, M-F) or OU Advocates 405-615-0013 (24.7) to learn more or to report an incident.