

DS 120
Scripting for Data Sciences (1)
Fall 2015
Wednesday 1:00-2:15
202 IST
Schedule #627373

Instructor: Dr. Dan Coughlin

Contact: Email – use Angel

412-716-3586 (texting is fine, identify yourself with name and class)

Office Hours: 302 IST, Wednesday afternoons 2:30-4pm and by appointment

Prerequisite: None

Course Description:

This one-credit introductory course aims to teach the basic and practical skills in programming and data manipulation needed to complete hands-on assignments in DS 200 Introduction to Data Sciences. While some students may acquire scripting skills through self-study or as part of the assignments in DS 200, this course will provide a structured learning environment for acquiring basic proficiency in the fundamental tools that allow data manipulation, and the coding techniques needed to combine them to form a tool chain. The course is designed to be highly accessible (no pre-requisites) for students interested in the data sciences. The course aims to provide a complement rather than a succession to IST 140, which teaches basic programming concepts in Java.

Course Materials:

Book:

Automate the Boring Stuff with Python: Practical Programming for Total Beginners

OR

<https://automatetheboringstuff.com/>

Course Objectives:

Upon completion of this course, students will be able to:

1. Write and run basic routines in a scripting language
2. Manipulate files using standard tools on a command line

Course Evaluation:

1. Attendance (20%)
2. In-class quizzes (10%)
3. Hands-on scripting homework assignments (40%)
4. Reflection Paper (15%)
5. Final Project (15%)

Class Assignments:

Attendance and participation (200 points)

This course is highly participative and requires attendance and active engagement.

Your participation grade will be based on:

Attendance: 50 points

Contribution to in class activities: 150 points

*15 Classes: each class is worth 15 points (10 attendance, 5 participation) = 225 total points. Each class you miss will automatically cost 15 points against this portion of your grade. You are permitted to miss 2 classes and still receive full credit for attendance and participation, **if** you let me know prior to class that you will be absent.*

If you will be leaving class early, please let me know prior to class.

In-class quizzes (100 points)

There will be two short quizzes during the semester that are worth 50 points each.

Hands-on scripting homework assignments (300 points)

These assignments will build on the in class work we do. If you are interested we can learn Git and GitHub and post our assignments there.

Reflection Paper (150 points)

You will be asked to complete a reflection paper on the overall course. The purpose of this reflection paper is thinking critically about this course, and what you learned throughout the semester (500-750 words). For example:

- How does the course challenge you?
- Has the course changed your way of thinking?
- Does the reading, lecture, or experience leave you with any questions? Were these questions ones you had previously or ones you developed only after finishing?
- Did we fail to address any important issues? Could a certain fact or idea have dramatically changed the impact of your experience in this class?
- How do the concepts brought up in this course mesh with past experiences or classes? Do the ideas contradict or support each other?

Final Project (250 points)

- The final project will be processing a small spreadsheet or document (XML or JSON) of data for manipulation in Python. This project will build on the concepts we learn throughout the semester.

Grading Scale:

Grade	Points
A	93-100
A-	90-92
B+	87-89
B	83-86
B-	80-82
C+	77-79
C	70-76
D	60-69
F	<60

Mid-term Review:

I will send out a survey halfway through the semester to be filled out anonymously on what is going well and what is going not so well to try and improve for the second half. You do not have to wait for this survey to provide feedback. If you have suggestions or thoughts on class, my door (and email inbox) is open.

Class Policy:

Attendance: You are expected to be in class on time, throughout the entire class period, and be engaged. Absences for illness, or other unavoidable circumstances may be considered excused if prior notice is received regarding the absence – in short be proactive in letting me know you if are unable to attend class. If you are unable to attend class, you are still responsible for all information covered in class.

Posting assignments to Angel: Where noted in the syllabus, assignments should be uploaded to Angel by the beginning of class on the date due. Dropboxes will be provided for each assignment and will close (i.e., access will be turned off) at the beginning of class. If you experience issues with Angel that prevent you from uploading your assignment, you should email the assignment directly to me. This will serve to show that the assignment was completed on time.

Make up assignments: Unless otherwise specified, all assignments are due by the beginning of class on the day they are assigned. Assignments should be uploaded to the appropriate Dropbox on Angel or turned in at the beginning of the class (see Syllabus for individual instructions). Late assignments will be accepted for unusual or unforeseen circumstances beyond your control provided you inform me of the situation and the reason for the assignment being late before the due date of the assignment.

Email: Use Angel to email me. I will use Angel to email you. Don't pretend my email coming from Angel got marked as junk because that's mean.

Weekly Topics

Week 1 - August 26, 2015

- Intro

Week 2 - September 2, 2015

- Different shells
- Unix commands

Week 3 - September 9, 2015

- Aliases, profile
- Shell scripting
 - Strings and console output

Week 4 - September 16, 2015

- Vi
- Python Shell
- Data Types
- Variables and Scope

Week 5 - September 23, 2015

- **Assignment #1 Due**
- Comparison Operators
- Control Constructs

Week 6 - September 30, 2015

- Control Constructs
- **Quiz #1**

Week 7 - October 7, 2015

- Functions
- Function scope

Week 8 - October 14, 2015

- Lists, Tuples
- Dictionaries

Week 9 - October 21, 2015

- **Assignment #2 Due**
- Dictionaries
- Strings

Week 10 - October 28, 2015

- Regular Expressions
- Working with Files

Week 11 - November 4, 2015

- Assignment #3 Due
- Data Formats (JSON, XML)
- Data Types
 - Dicts
 - Lists
 - Tuples
 - Looping over them
 - Lists of lists
 - Lists of dicts
 - Dicts with lists
- Working with CSV

Week 12 - November 11, 2015

- **Git and GitHub**
- **Quiz #2**

Week 13 - November 18, 2015

- **Web Scraping**
- **Spreadsheets**
- Go over requirements for final project, Assignment #4 will be based on final project
- **Open** – Lab final project
- **Go over final paper**

Week 14 - November 25, 2015 - Thanksgiving Break

Week 15 - December 2, 2015

- **Open** – Lab final project

Week 16 - December 9, 2015

- **Open** – Lab final project

December 11, 2015 5pm

- Final paper due
- Final project due

Academic Integrity:

According to the Penn State Principles and University Code of Conduct: "Academic integrity is the pursuit of scholarly activity in an open, honest and responsible manner. Academic integrity is a basic guiding principle for all academic activity at The Pennsylvania State University, and all members of the University community

are expected to act in accordance with this principle. Consistent with this expectation, the University's Code of Conduct states that all students should act with personal integrity, respect other students' dignity, rights and property, and help create and maintain an environment in which all can succeed through the fruits of their efforts.

Academic integrity includes a commitment not to engage in or tolerate acts of falsification, misrepresentation or deception. Such acts of dishonesty violate the fundamental ethical principles of the University community and compromise the worth of work completed by others.”

Any violation of academic integrity will be investigated, and where warranted, punitive action will be taken. For every incident when a penalty of any kind is assessed, a report must be filed.

Affirmative Action & Sexual Harassment:

The Pennsylvania State University is committed to a policy that all persons shall have equal access to programs, facilities, admission, and employment without regard to personal characteristics not related to ability, performance, or qualifications as determined by University policy or by Commonwealth or Federal authorities. Penn State does not discriminate against any person because of age, ancestry, color, disability or handicap, national origin, race, religious creed, sex, sexual orientation, or veteran status.

Direct all inquiries to the Affirmative Action Office, 211 Willard Building.

Americans with Disabilities Act:

The College of Information Sciences and Technology welcomes persons with disabilities to all of its classes, programs, and events. If you need accommodations, or have questions about access to buildings where IST activities are held, please contact us in advance of your participation or visit. If you need assistance during a class, program, or event, please contact the member of our staff or faculty in charge. Access to IST courses should be arranged by contacting the Office of the Senior Associate Dean, 332G IST Building, (814) 865-4461.

An Invitation to Students with Learning Disabilities:

It is Penn State's policy to not discriminate against qualified students with documented disabilities in its educational programs. If you have a disability-related need for modifications in your testing or learning situation, your instructor should be notified during the first week of classes so that your needs can be accommodated. You will be asked to present documentation from the Office of Disability Services (located in 116 Boucke Building, 863-1807) that describes the nature of your disability and the recommended remedy. You may refer to the Nondiscrimination Policy in the *Student Guide to University Policies and Rules*.

Policy AD42 Statement on Nondiscrimination and Harassment:

The Pennsylvania State University is committed to the policy that all persons shall have equal access to programs, facilities, admission and employment without regard to personal characteristics not related to ability, performance, or qualifications as determined by University policy or by state or federal authorities. It is the policy of the University to maintain an academic and work environment free of discrimination, including harassment. The Pennsylvania State University prohibits discrimination and harassment against any person because of age, ancestry, color, disability or handicap, national origin, race, religious creed, sex, sexual orientation, gender identity or veteran status. Discrimination or harassment against faculty, staff or students will not be tolerated at The Pennsylvania State University. You may direct inquiries to the Office of Multicultural Affairs, 332 Information Sciences and Technology Building, University Park, PA 16802; Tel 814-865-0077 or to the Office of Affirmative Action, 328 Boucke Building, University Park, PA 16802-5901; Tel 814-865-4700/V, 814-863-1150/TTY.

For reference to the full policy: <http://guru.psu.edu/policies/AD42.html>