syllabus-1360e

May 10, 2017

1 CSCI 1360E: Foundations for Informatics and Analytics

...or, more colloquially, **"Introduction to Data Science."**

1.1 Table of Contents

This page is, for all intents and purposes, the course syllabus. The primary course interactions will be through 1) the Slack channel, and 2) JupyterHub, both of which I'll address here.

1.2 Overview

Informatics, or "data science," are rapidly becoming essential skills for scientists across fields; in addition to field-specific specializations, researchers require knowledge of and experience with quantitative analytical techniques for extracting knowledge from raw data.

This course aims to provide an introduction to concepts in scientific programming and data science using the Python language. Students are given hands-on opportunities to learn techniques applicable to quantitative analyses across a broad range of fields. These core techniques involve formulating solutions in terms of their inputs and outputs (functional programming), repeated operations (loops), branching operations (conditionals), different methods of organizing data (data structures), how to implement an optimal problem-solving strategy (algorithm design), and methods for visualizing and interpreting results.

1.3 Prerequisites

The only hard prerequisite is **MATH 1113 Precalculus**.

This course assumes **no prior programming or statistics knowledge**. It is meant to be an introduction to these concepts in the larger context of data science and scientific programming through

the lens of the Python ecosystem. The course is targeted at undergraduate students across fields who, irrespective of their ultimate career goals, are interested in a foundational understanding of programming and quantitative data analytics.

1.4   Materials

As this course is online, there are few physical requirements in order to effectively participate.

- You'll need a computer with an internet connection. If you're reading this, you can most likely claim victory on this one.
- An up-to-date web browser. Any should fit the bill:
- Firefox (if you have enough memory)
- Chrome (if you have enough spare CPU cycles)
- Safari (if you're iRad enough)
- IE (if you hate life)
- Edge (if you merely strongly abhor life)
- Opera (because someone has to stand up for the little guy...?)
- Vivaldi (the two people who use it say it's pretty cool)
- Konqueror (...I honestly know nothing about this browser)
- Lynx. LOLOOL j/k, you'll need to be able to render, y'know, JavaScript and images.
- **There is/are no required textbook(s)**. There are a few recommended texts that I will steal borrow from over the course of the semester, and you're welcome to check them out or purchase them if you're interested, but any material from them that you'll need will be reproduced here in some form.

If you are so inclined to purchase a reference textbook, I would highly recommend the following (conveniently ordered from [in my humble opinion] most informative to...whatever the opposite of "most informative" is): - Grus, Joel. *Data Science from Scratch: First Principles with Python* (1st ed., 2015) ISBN-13: 978-1491901427. - McKinney, Wes. *Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython* (1st ed., 2012) ISBN-13: 978-1449319793. - Shaw, Zed. *Learn Python the Hard Way* (3rd ed., 2013) ISBN-13: 978-0321884916.

1.5   Layout

Given the online nature of this course, the primary vector through which you'll submit assignments is JupyterHub. This is where you'll submit assignments for grading.

Furthermore, this is the format of the lectures. Yes! No videos! Yay! These "lectures" will be released every **Monday, Wednesday, and Friday**, with assignments coming out every **Tuesday and Thursday**. More details below.

1.6   Grading

Yep, the part everyone pays attention to.

- Participation: **5%**
- Assignments: **70%**
- Midterm: **10%**
- Final: **15%**

There will be 10 relatively brief programming assignments, each worth 7% of your total grade. These are intended to give you hands-on experience with the concepts being taught in the class and to familiarize you with the Python language and its ecosystem. There will be midterm and final exams, and a participation component. The latter takes the form of asking and/or answering questions in the Slack channel, leading study discussions in the Slack channel, or participating in Slack office hours.

1.7 Course Resources and Online Portals

This is the crux of everything. If you read one part of the syllabus, let it be this one.

1.7.1 1: JupyterHub

This is the primary point of interaction for homework assignments and exams. Jupyter notebooks (like this one that you're reading!) will be posted here. For lectures, they'll be posted in the Slack channel. Here's the link:

http://jupyterhub.cs.uga.edu It's only accessible from on-campus, or with a campus VPN. Check out this EITS webpage on remote access if you need assistance.

1.7.2 2: Slack

This is the primary point of interaction for asking for / offering help. I will answer questions when I can, and will make myself avaialble at specific "office hours" during the week when I am guaranteed to be sitting in front of the channel, but otherwise I encourage everyone to help each other out, too!

`https://eds-uga-csci1360.slack.com/` I will send out invites to the Slack channel using your UGA email address.

1.8 Outline

Here is a rough outline of the progression of topics we'll cover in CSCI 1360E (subject to change without notice):

1.8.1 Week 1 (Mon, June 5 - Fri, June 9)

- **Monday**
- Lecture 1
- **Wednesday**
- Lecture 2
- **Thursday**
- Assignment 0 (A0) released
- **Friday**
- Lecture 3
- *Drop day*

1.8.2 Week 2 (Mon, June 12 - Fri, June 16)

- **Monday**
- Lecture 4
- **Tuesday**
- A0 due
- A1 released
- **Wednesday**
- Lecture 5
- **Thursday**
- A2 released
- **Friday**
- Lecture 6
- A1 due

1.8.3 Week 3 (Mon, June 19 - Fri, June 23)

- **Monday**
- Lecture 7
- **Tuesday**
- A2 due
- A3 released
- **Wednesday**
- Lecture 8
- **Thursday**
- A4 released
- **Friday**
- Lecture 9
- A3 due

1.8.4 Week 4 (Mon, June 26 - Fri, June 30)

- **Monday**
- Lecture 10
- **Tuesday**
- A4 due
- A5 released
- **Wednesday**
- Lecture 11
- **Friday**
- Midterm Review
- A5 due

1.8.5 Week 5 (Mon, July 3 - Fri, July 7)

- **Monday**
- **MIDTERM EXAM**
- **Tuesday**
- **HOLIDAY - NO LECTURE OR HOMEWORK**

- **Wednesday**
- Lecture 12
- **Thursday**
- A6 released
- **Friday**
- Lecture 13

1.8.6   Week 6 (Mon, July 10 - Fri, July 14)

- **Monday**
- Lecture 14
- **Tuesday**
- A6 due
- A7 released
- **Wednesday**
- Lecture 15
- **Thursday**
- A8 released
- **Friday**
- Lecture 16
- A7 due

1.8.7   Week 7 (Mon, July 17 - Fri, July 21)

- **Monday**
- Lecture 17
- **Tuesday**
- A8 due
- A9 released
- **Wednesday**
- Lecture 18
- **Thursday**
- A10 released
- **Friday**
- Lecture 19
- A9 due

1.8.8   Week 8 (Mon, July 24 - Thurs, July 27)

- **Monday**
- Lecture 20
- **Tuesday**
- A10 due
- **Wednesday**
- Lecture 21
- **Friday**
- Lecture 22

1.8.9 Week 9 (Fri, July 28 and Mon, July 31)

- **Fri, Mon**
- FINAL EXAMS WEEK
- More details to be announced later

1.9 Policies

Kinda boring, but absolutely necessary. Please be familiar with these points so you and I don't have to have unpleasant conversations.

1.9.1 Assignments

- Assignments are due by 11:59:59pm on the noted date. Assignments turned in after this deadline will lose 25/100 points for every subsequent 24 hour-period they are late.
- The presence or absence of any form of help or collaboration, whether given or received, must be explicitly stated and disclosed in full by all involved, on the first page of their assignment ("I did not give or receive any help on this assignment" or "I helped [person] with [specific task]."). Collaboration without full disclosure will be handled severely; except in usual extenuating circumstances, my policy is to fail the student(s) for the entire course.
- **DO NOT COPY CODE**. I cannot stress this enough. Coding is a lot like writing: everyone has their own style that is very recognizable. It's not difficult to tell when students share their code. Don't do it.

1.9.2 Exams

- Any material covered in lecture or homework assignments is considered fair game for both exams.
- Both exams will be cumulative. I may even be lazy / clever and copy-paste midterm questions into the final.
- The exact format of both exams will be variable; previous years' exams have included a mixture of multiple choice, matching, true/false, and hand-coding. However, this being the first online course in CSCI 1360, it is not yet clear what the ultimate formats will be.
- Time permitting, "practice" versions of the exams will be released. I will make every effort, but it will depend entirely on my schedule when the time comes; as such, do not assume it will happen.
- **Exams will be timed.** You can complete them within the time frame allotted in under the maximum time interval. As this is an online course, it's viable for multiple people to take the exam while co-located; please resist this option. Colloboration in **any** form during either exam will be grounds for immediate failure of the course.

The UGA Academic Honesty Policy is the final word on these matters. Lack of knowledge of these policies is not sufficient justification for violations. If in doubt, **ask me**.

1.10 Contact

Dr. Shannon Quinn: resident new professor, marathoner, and certified nerd (I have a degree to prove it).

Email: **squinn@cs.uga.edu** (include **1360E** in the subject line so it hits my email filter!)

Office: Boyd GSRC, Room 638A.
Office Phone: 2-4661
Website: `http://cs.uga.edu/~squinn`