

COSC 480 - Topics in Computer Science

Data Science for Social Good

Spring 2017

Syllabus

How to use data for the greater good.

Just the Facts

Course Number: COSC 480

Title: Topics in Computer Science - Data Science for Social Good

Semester: Spring 2017

Meeting Time: MWF 1:20 - 2:30 pm

Locale: Schaefer 165

Instructor: Alan Jamieson

Office: Schaefer 154

Office Hours: W 4:30-6, TR 4-5:30

Email: acjamieson@smcm.edu

Google Messenger: acjamieson@smcm.edu

Slack: via Alan's COSC480 Group

Online Office Hours: Most evenings and weekends

Recommended Textbooks: *Grus, Data Science from Scratch, O'Reilly*

O'Neil and Schutt, Doing Data Science: Straight Talk from the Frontline, O'Reilly

Website: <http://ripark.github.io/s17/cosc480>

Catalog Description: This course is a rigorous study of an important field in computer science. Examples: data security; bioinformatics; natural language processing; compilers. A detailed course description will be available before registration. The course may be repeated for credit where the topic is not repetitive. Prerequisite: COSC 201 or permission of the instructor.

Overview: Data science is one of the hottest subfields of computer science and mathematics over the past few years. The ability of a programmer to take large sets of data and glean some useful nugget for predictive analysis and decision-making has been sought after by the largest and smallest companies. In this course, we will do a full introduction to data science, focusing on the algorithmic techniques required in Python and R. We will also do readings and projects on how data science can be used to drive societal change.

Purpose: Data science and "Big Data" blend programming, statistics, and, sometimes, social sciences in order to find connections in chaos and make predictions for the future. It can be used for great things, or it can be used to tear societies apart. The goal of this course is to give you the tools to be able to solve these great problems, and to introduce the ethical questions behind many of the decisions we make as computer scientists.

Grade Distribution:

Labs/quizzes/homework - 20%

Miniprojects (2) - 15% each

Major Project - 40%

Major Project Presentation - 10%

You will be expected to participate in class by asking questions and answering questions posed by myself and those in class. Rather than a drab lecture, the class sessions will be run in a discussion style environment. Activity and debate are highly encouraged. In addition to this, there will be coding exercises and every student will be expected to submit at least one solution throughout the semester. Failure to do so will result in a grade penalty.

Learning Objectives: At the completion of COSC 480, students will be able to:

explain the aspects of the topic of the course

implement problem work in the topic of the course
design solutions to large problems in the topic of the course

Final Information: There will not be a final exam in this class, but teams will be presenting their final projects during the final period Thursday, May 4 at 10am (not 9) in Schaefer 165. Failure to attend your presentation will result in an F for the course.

Assignments: There will be three projects in this course, two smaller and one major. The major project will be a half-semester team project. Each team will take on some data science problem related to a larger societal problem, procure the data sets, and then present the results during the final period. Your group will also prepare a position paper arguing your results to be sent to appropriate officials in the hopes of instigating change. In addition, there will be a smattering of in-class and homework programming assignments.

Blackboard Use: I will be utilizing Blackboard primarily for your grades in this course. Course materials will be provided on the course website.

GitHub Use: Major project files and some other sources will be maintained on GitHub. Any questions, concerns or objections should be noted during the first week of classes. You will be given a primer on how to use GitHub during the first part of the semester.

Policies

Cell Phones: Please, turn off or turn to silent any cell phones prior to getting to class. If they go off in class they are distraction not only to myself, but to everyone else in the class as well. Habitual offenders will be excused from the class with a 0 for any quizzes and class participation for that day.

Computer Use: Computer use in this lab is for academic use only. If you bring a laptop with you to this class I expect you to be only using it for purposes related to this class. The same goes for the computers in this lab.

Attendance and Tardiness: Attendance is highly recommended. Missing a class not only causes you to miss the information disseminated in that lecture, but can cause you to miss important information in regards to assignments and the potential of receiving a 0 for a quiz that day. I start class promptly on the hour and expect the students to be in class at that time. If you have circumstances that can prevent you from being in class on time, please let me know as soon as possible. Habitual offenders will be excused from the class with a 0 for any quizzes and class participation for that day.

Exams and Quizzes: There will be no exams in this class. Every class has the potential of having a quiz to reinforce the ideas from the lecture the previous class. These will not be announced ahead of time. They will be 1-3 question quizzes that can be easily done in 15 minutes either at the start or the end of the class period.

Assignments: Assignments and other outside of class work should be done on an individual basis unless otherwise specified in the description of the assignment. Assignments and other outside of class work will not be taken late except under extraordinary and documented circumstances.

Extra Credit: I will not be offering any extra credit opportunities in this class.

Communication: The simplest way to get in touch with me is by coming by my office during my office hours or contacting me via email. The easiest way to get in touch with me "after hours" is to send me an email. I habitually check my St. Mary's email account all hours of the day. If you come by my office and the door is open, feel free to stop in to chat. The open door indicates that I'm not working on anything that has to keep my undivided attention at that time so do not feel that you are interrupting me or anything like that. I do make appointments if you have a certain time that you'd like to meet with me. If it fits in my schedule (meaning I'm not teaching class during that

time) I will be happy to meet with you.

Academic Honesty: Academic misconduct policies are covered in the Student Code and Student Rights and Responsibilities, Article III. Pay close attention to the definitions of academic misconduct noted in Section 1. This can be found in the Student Handbook.

Disability: If you have any kind of disability that can affect your performance in this class, please let me know privately through email or stopping by my office.

Schedule: The schedule for the class will be posted to the class website. The schedule is subject to change (multiple times).

Closing: The most important thing in any of my classes is that you are learning and expanding your horizons. If you are having any undue difficulty with your work as it pertains to this class, please contact me as soon as possible. Always remember that professors succeed when you don't need us any longer. I want you to be bouncing ideas off of each other throughout the class and it is my hope that by the end of the semester that you are driving the class session rather than me.