FYOS 1001 - Data Science:
The Practice of Turning Numbers into Knowledge

Fall 2016

Time: 1 Hour. Monday 10:10-11:00am.
Place: Ecology Seminar Room (Room 117)

Instructor: John Drake, Ph.D.
Email: jdrake@uga.edu
Office: 133 Ecology Building

Overview
The contemporary world is flooded by data. The acquisition of data and use of analytics for decision-making are increasingly important in areas of society ranging from science and health care to government policy to business. Data science is the emerging discipline at the nexus of three primary disciplines – computer programming, machine learning, and statistics – that aims to make sense of these data. Students will engage data science through the lens of their instructor's research in ecology and epidemiology. Students will be introduced to the three primary disciplines that constitute data science and learn about applications of data science to research at UGA and across a range of careers. Students will think critically about evidence, numerical arguments, and the use of data in personal life; consider how the acquisition of data can lead to bias; and use Tufte's principles to identify intentionally misleading elements in graphical displays of data.

Objectives
The aim of this course is to understand the new interdisciplinary field of *data science*. As a First Year Odyssey Seminar, the course has three goals:

> **Goal 1:** Introduce first-year students to the importance of learning and academics so that we engage them in the academic culture of the University.
>
> **Goal 2**: Give first-year students an opportunity for meaningful dialogue with a faculty member to encourage positive, sustained student-faculty interactions.
>
> **Goal 3:** Introduce first-year students to the instruction, research, public service and international missions of the University and how they relate to teaching and learning in and outside the classroom so that we increase student understanding of and participation in the full mission of the University.

Format and assignments
Class sessions will consist of group discussions of course readings and a group project. These sessions are designed to cultivate habits of mind that facilitate critical thinking and to introduce techniques of oral discourse. Students are expected to have read the assigned reading and be prepared for discussion. Short writing assignments will be used as a platform for exploration and knowledge development, to hone written communication skills, and to provide a basis for sustained reflection. These will include a short library research project, an exercise in critical thinking, and an exercise in data analysis. Finally, the methods of data science will be introduced during a group project to be completed in class.

Attendance
Attendance is essential. If you miss a class, you are responsible for contacting fellow students for lecture notes. More than one unexcused absence will result in a lowering of your final score by one letter grade. An excused absence is constituted by the student notifying the instructor in advance and documenting reason of absence immediately upon return to the class.

Grade calculation
50% Participation and class discussions
40% Short writing assignments
10% Student sharing during final class session

Accommodations
Please contact the instructor if you require special accommodations due to learning disabilities, religious practices, physical or medical needs, or for any other reason. Please let the instructors know if you require special accommodation for the field trip.

General Notes
(i) All academic work must meet the standards contained in "A Culture of Honesty." Students are responsible for informing themselves about those standards before performing any academic work. The link to more detailed information about academic honesty can be found at: http://www.uga.edu/honesty/
(ii) The course syllabus is a general plan for the course; deviations announced to the class by the instructors may be necessary.

Schedule of Topics

Date	Topic	Reading	Assignments
08/15/16	What is data?		
08/22/16	Autobiography of a data scientist		
08/29/16	The best job of the 21st century	Davenport and Patil (2012)	
09/05/16	NO CLASS		
09/12/16	Computer-aided identification of disease reservoirs	Han (2015)	Writing assignment 1
09/19/16	Machine learning	Jordan & Mitchell (2015)	
09/26/16	Principles of graphical fairness	TBD	
10/03/16	What is big data?	McAfee & Brynjolfsson (2012)	
10/10/16	Data traps	Lazer et al. (2014)	Writing assignment 2
10/17/16	Model based prediction of Zika virus	Perkins et al. (2016)	
10/24/16	Doing data science 1		
10/31/16	Doing data science 2		
11/07/16	Doing data science 3		Writing assignment 3
11/14/16	Doing data science 4		
11/21/16	NO CLASS		
11/28/16	Data science summary		
12/05/16	Student sharing		