

**Knowledge in the Age of Big Data**  
**HCOL086 Spring 2019**

**Instructor:** Prof. Sara Helms Cahan, Department of Biology

**Contact info:**

Office: Marsh Life Sciences 120B (inside the Biology main office and to the left)

E-mail: [scahan@uvm.edu](mailto:scahan@uvm.edu)

Phone: 656-2962

Office hours: MT 11:00am-noon

**Overview:** In the digital age, we have the capacity to generate, store, and analyze essentially limitless amounts of information about our physical, biological, and social environments. Collectively, this storehouse of information is referred to as Big Data: information far larger and/or more complex than our minds can easily comprehend in its entirety. The advent of Big Data has been alternatively hailed as a tool to solve our most vexing problems, and as a false prophet that deceives us as much as (or more than) it enlightens us. In this course, we will explore what it means to take a data-driven approach to problems, and how such an approach fits into the larger human quest for knowledge, wisdom, and understanding. What does it mean to “collect data”? Do data represent objective truth? Should they replace or supersede other ways of knowing? What kind of questions can they answer? How is meaning created from a bunch of numbers? Can data be misused, and if so, are they of any objective value at all? What should we, as consumers of information, trust?

We will begin with the first-year read, *Between the World and Me* by Ta-Nehisi Coates, to help us identify how data collection fits in to the main approaches to finding knowledge you were introduced to in the Fall semester. This will lead us to our central question: how can data be used to see and understand processes beyond what we individually experience? We will use two books to help us develop a robust data-analysis framework to address questions in a variety of fields, including environmental science, social science, biology, sports, and marketing: *Truth and Truthiness* by Howard Wainer, which will set up the main frameworks by which we use data to answer questions, and *Big Data* by Viktor Mayer-Schönberger and Kenneth Cukier, which considers what changes, improves, and becomes problematic when we scale up to extremely large datasets. Finally, we will consider the ethical implications of living in an age of ever-increasing volumes of data, and complex algorithms that analyze and use them.

**Learning Objectives:**

By the end of this course, you should be able to:

- 1) Be able to articulate the elements of the scientific method and the role of data in evaluating evidence
- 2) Be able to recognize, interpret, and create summary data tables and figures
- 3) Be able to integrate data-based evidence into an argument in multiple formats (public-interest writing, technical report, oral presentation, research poster).

**Texts:** *Between the World and Me* by Ta-Nehisi Coates

*Truth or Truthiness* by Howard Wainer

*Big Data* by Viktor Mayer-Schönberger and Kenneth Cukier

We will also be reading a number of articles from the popular press and from academic journals, which will be provided as links or pdfs on the course Blackboard site.

**Attendance:** Attendance is mandatory. You get up to one session you can miss with a reasonable excuse, but after that, you are going to lose points! This is true for the plenary as well as our class sessions. I go to all the plenaries, and I expect to see you and check off that you have attended.

**Assignments:** My philosophy in this course is that the best way to learn is through practice, which translates into a lot of regular reading, writing and problem-solving. There will be a short prompt for each reading assignment posted as a channel on Slack, to which everyone is expected to contribute at least one thoughtful posting, two 2-3 page written assignments associated with different topics, and a variety of graded in-class and at-home activities. A longer (6-8 page) individual research paper will be due as a first draft roughly 2/3 of the way through the semester and a final draft at the end of the semester to accompany the group project (detailed below).

**Group project:** All sections of HCOL086 complete group projects designed to give you practice in building an effective and evidence-based argument through teamwork presented through different modes of communication. Your task in our class is to apply data *to a problem of your choice*, either by studying results from published sources or collecting them yourselves. Each group will be composed of 3-4 students. Groups will need to make decisions about the overall story they want to tell, what to quantify and how, and the appropriate way to visualize and analyze the data, along with an evaluation of their implications, value, and limitations. In addition to writing an individual research paper on your own piece of your group's topic, your group will be responsible for producing 1) a 10-minute Powerpoint presentation that you will present to another HCOL086 section, and 2) a research poster that will be presented to the entire HCOL086 community (faculty and students) in a public poster session near the end of the semester. We will be spending a lot of class time on the group project, and many resources will be provided by HCOL and the library to assist you with these tasks along the way. Overall, the group poster and presentation come to 25% of the total points for the semester, graded on a combination of individual contribution (15%) and group performance (10%).

**Exams:** There will be a mid-term and a final exam designed to assess your ability to apply the topics and skills you have learned to a novel situation. Both exams will be take-home, and due one week after they are assigned. The fact that these are take-home exams does NOT mean you will not need to take notes on what we do in class! I will be looking for evidence that you can apply the terminology, concepts, and skills we have covered in class to a novel problem.

**Grades:** The values below are approximate, as our schedule will adjust depending on how long activities take and where our discussions lead. In the end, your grade will be calculated as a percentage of the total points that end up being assigned, and everything will be posted on the Blackboard grade center.

Class participation:	20 pts
Reading questions:	50 pts total
Writing assignments:	30 pts total
Activities/assignments:	50 pts total
Research paper:	50 pts (30 for first draft, +20 for final)
Group poster:	40 pts
Group presentation:	50 pts
Project participation:	10 pts
Mid-term exam:	50 pts
Final exam:	50 pts
<b>Total</b>	<b>400 pts</b>

**Policies on Academic Honesty:** University policies are in the Cat's Tale Student Handbook (<http://www.uvm.edu/~dosa/handbook/?Page=Academic.html>). Plagiarism of ANY sort - e.g., copying from original references, texts, or websites - will NOT be tolerated. Acts of plagiarism or cheating constitute grounds for failing the course and university disciplinary action.

**Religious Observances:** If you have to miss any important events (exams, assignment deadlines) due to religious observance, be sure to tell me at least A WEEK in advance so I can make alternative arrangements for you.