SDS 236: Data Journalism

About the Course

The course

The course meets Tuesdays and Thursdays from 10:30 am - 11:50 am in Bass 002 (the new psychology/SDS lab in the basement).

Instructors

- Amelia McNamara (`amcnamara@smith.edu`, McConnell 211), office hours 2-3 pm Monday, 1-2 pm Thursday and Friday.

Description

Data journalism is the practice of telling stories with data. This course will focus on journalistic practices, interviewing data as a source, and interpreting results in context. We will discuss the importance of audience in a journalistic context, and will focus on statistical ideas of variation and bias. The course will include hands-on work with data, using appropriate computational tools such as R, Python, and data APIs. In addition, we will explore the use of visualization and storytelling tools such as Tableau, plot.ly, and D3. No prior experience with programming or journalism is required.

Prerequisite: Enforced prerequisite: An introductory statistics course, including SDS 220, MTH 220, SDS 201, PSY 201, ECO 220, GOV 203, AP Statistics, or permission of the instructor.

Textbook

Numbers in the newsroom: Using math and statistics in news Sarah Cohen, 2nd edition. The textbook is available in two formats:

- electronic edition
- paperback edition from the Smith bookstore

We will also be reading many articles, both examples of data journalism and readings specifically about the practice of data journalism, but these will be made available via links on moodle.

Other materials

You will need a notebook to use for notetaking while interviewing. Many reporters prefer notebooks with spiral binding at the top of the page like this.

You will also want some method of recording audio, again for interviewing. If you have a smartphone, it probably has an app for this. If not (or you want higher-quality audio), the Center for Media Production lends audio recorders to students.

We'll be using Google Drive to organize our documents, and Slack for communication.

This course will explore many computing technologies, including R and RStudio (either the server version on the web, or a desktop version).

This class is supported by DataCamp, the most intuitive learning platform for data science. Learn R, Python and SQL the way you learn best through a combination of short expert videos and hands-on-the-keyboard exercises. Take over 100+ courses by expert instructors on topics such as importing data, data visualization or machine learning and learn faster through immediate and personalised feedback on every exercise."

Depending on the stories we pursue, we will explore Tableau, plot.ly, d3, OpenRefine, Tabula, and more. All software should be available for free, and instructions for installation will be provided.

If you have financial constraints you find are limiting your ability to be successful in this course, please see me. There are resources at the college for laptop loans and textbook purchase, and I will do everything I can to help you meet your needs.

Classes

Classes meet Tuesday and Thursday. I expect you to attend class. Your participation is an important part of the learning process. If you cannot attend a particular class I would appreciate the courtesy of advanced notice and an explanation for your absence. Class participation and attendance contribute 20% to your final grade.

I hope it goes without saying, but while the class is in session, you should not use your computer or cell phone for personal email, web browsing, Facebook, or any activity that's not related to the class. Please try to bring a laptop to class (we'll often need to have sufficient numbers of both to allow students to work in pairs).

Policies

Collaboration

Much of this course will operate on a collaborative basis, and you are often expected to work together with a partner or in small groups to complete assignments. However, **every word that you write must be your own**. Copying and pasting sentences, paragraphs, or blocks of code from another student is not acceptable and will receive no credit. All students, staff and faculty are bound by the Smith College Honor Code, which Smith has had since 1944.

Academic Honor Code Statement

> Smith College expects all students to be honest and committed to the principles of academic and intellectual integrity in their preparation and submission of course work and examinations. Students and faculty at Smith are part of an academic community defined by its commitment to scholarship, which depends on scrupulous and attentive acknowledgement of all sources of information, and honest and respectful use of college resources.

Cases of dishonesty, plagiarism, etc., will be reported to the Academic Honor Board.

Assignments

1. Data diary [15%]: To get you in the practice of writing regularly, you will keep a Data Diary on Google Drive. Diaries will be checked on Sunday evenings, and most weeks you are expected to have four entries. Entries can be short (three sentences) but should be written throughout the week rather than all at once.
2. Short writing [35%]: There will be regular short writing assignments, generally due on Monday evening.

3. Large reporting projects [30%]: Three times throughout the semester, you will create a more significant piece of long-form journalism.
4. Participation [20%]: Our classes will be highly interactive. The remainder of your grade will be based on your participation in our in-class discussions. This category includes our regular news briefing.
5. Extra Credit [?]: Extra credit may be offered for attending an out-of-class lecture (as will be announced) and writing a short review of it.

Reporting

We will be exploring the concept of reporting throughout this course. You will be producing regular written assignments, following the quick pace of a standard newsroom and news cycle.

In general, assignments will include the following components:

- A pitch. In the first half of the course, I will provide the pitch to you. As the course progresses, you will be expected to research and pitch your own stories to me.
- Data research. This may include: searching on the web for data, reaching out to sources, submitting Freedom of Information Act (FOIA) requests.
- Initial writing. Each piece of reporting you produce must go through a series of editing stages. Your initial draft should be produced quickly, and include the basic structure of a data story (headline, byline, lede, explanation, and additional information).
- Editing. This will primarily be done by someone serving in an "editor" role (sometimes me, sometimes your fellow students). Editing will include copyediting, proofreading, and reading for clarity.
- Finished product. Your finished product should be print-ready.

For assignments that include visualization, the components will be largely the same, with "visualization" substituted for "writing." Visualizations still need to be pitched, go through a draft process, and be edited for clarity and accuracy.

Grading

When grading your written work, we are looking for writing this both technically correct and clearly explained. Assessment will take into account the intended audience, the venue for publication, the word count and other reporting briefs, accuracy of reporting, and grammatical clarity.

Your ability to communicate results, which may be technical in nature, to your audience, which is likely to be non-technical, is critical to your success as a data analyst. The assignments in this class will place an emphasis on the clarity of your writing.

Resources

Course Website

The course website and Moodle page will be regularly updated with lecture handouts, project information, assignments, and other course resources. Announcements will primarily be made via Slack. It is your responsibility to check these resources regularly.

Tentative Schedule

The following is a brief outline of the course. Please refer to the complete day-to-day schedule for more detailed information.

Week	Reading	Topic
1	Ch. 1	Introduction to Data Journalism

Week	Reading	Topic
2	Ch. 2	Fractions, rates, percents, and per capita
3	Ch. 2	Averages, central tendency
4	Ch. 3	Data visualization
5	Ch. 3	Data visualization
6	Ch. 4	Standard stories
7	Ch. 5	Working with survey data
8	Ch. 6	Mistakes in the news
9		Mapping
10		Reporting on algorithms and AI
11		Timelines
12		Interactivity
13		Deep-dive reporting