

JOUR 407 Data Journalism

Spring 2016

T-Th 3:30 - 4:50 p.m.

ANDN 114

Instructor: Matt Waite

Email: mwaite3@unl.edu

Twitter: @mattwaite

Phones: (402) 802-5202 cell, (402) 472-5840 office

Office: 244 Andersen

Course description:

Every day, more of our lives is being stored in a database somewhere. With that explosion of data, journalists now more than ever need the skills to analyze and understand data to then produce the stories hidden in the information. In this class, we'll use brainpower and software to look at raw data -- not summarized and already reported information -- to do investigative reporting. We're going to get our hands dirty with spreadsheets, databases, maps and some basic stats. And we're going to do journalism. So buckle up and hold on.

Prerequisites:

JOUR202. JOUR302/304 or 370 is strongly encouraged.

Course goals:

- Understand the basics of data and data journalism, including the history of the practice
- Master the use of data in journalistic storytelling
- Master spreadsheets and basic data analysis for storytelling
- Use public records laws and understand your rights as a citizen and a journalist

- Gain exposure to advanced tools of data journalism, like GIS and statistics, as well as advanced methods of gathering data.

Required Materials:

- A functioning laptop computer that you must bring with you to class every time and ~800 MB of free hard drive space.
- The administrative privileges to install software on your computer.
- [An up-to-date version of Firefox](#). Free!
- A Google Account. Free!
- [The Data Journalism Handbook](#) Free!
- A copy of [Numbers in the Newsroom](#). \$10.
- A copy of Data Literacy: A Users's Guide by David Herzog. Available in the bookstore.
- Online materials and class handouts, as needed.
- A sense of humor.

Grading:

The grading will be based on the stories you produce, the work you put into them and your participation in class.

The bulk of the graded work in this class is as follows:

Assignment	Percentage of your grade
Four enterprise stories	60%
Data negotiation	20%
Assignments	10%
Data diaries	5%
Class participation/attendance	5%

Enterprise stories: You will be required to pitch and execute four data journalism stories on your own during the semester. The stories must include original

analysis of data you have obtained, a graphic or visualization of that data and a story worth publishing on NewsNetNebraska (or elsewhere if possible). Each story will be line edited and produced for online at an appointed time outside of class.

Data negotiation: During the semester, you will identify a database held by a government agency that you need for a story and go get it. You are negotiating for public data as a journalist, you may not promise to not use the records.

Downloading data from the Internet does not fulfill the requirements of this exercise. You will be assigned a deadline for your dataset, and on your deadline day, you will give a 5-min. lightning talk about your experience along with a written report. We will discuss details in class.

Data diaries: Throughout your analysis of data, you need to keep a running diary of what you have done -- what actions you took, commands you ran, thinking behind what you are doing. Yes, it will seem odd, but think of it like writing future you a note explaining how to do this again.

Class participation and attendance: You are bad at college if you don't know what this means. Get involved in class, speak up, ask questions, challenge ideas and you'll get the points. Sit on your hands, do nothing, skip class and you will not like what happens.

Notes on attendance:

Yes, we all get sick. Yes, things happen. I don't want you to be sick in my class any more than you want to be sick. You've got no fewer than four ways to get ahold of me, including my cell number. If you are going to miss class, tell me before class. We'll work it out. But you have to tell me before class for me to help you.

This said: this class builds each class onto the next one. Miss a class and you are behind. We're going to be covering a lot of new material in this class. Miss one at your own peril.

Policies

Here's the short version.

You cheat, you fail, no exceptions.

If I'm doing something that's keeping you from learning, tell me. Tell the Dean. Tell someone, because that's not cool. I won't tolerate it from myself and you shouldn't either.

Now the longer versions.

ACEJMC Competencies

After this class, you should be able to:

- Understand and apply the principles and laws of freedom of speech and press for the country in which the institution that invites ACEJMC is located, as well as receive instruction in and understand the range of systems of freedom of expression around the world, including the right to dissent, to monitor and criticize power, and to assemble and petition for redress of grievances;
- Demonstrate an understanding of the history and role of professionals and institutions in shaping communications;
- Demonstrate an understanding of gender, race ethnicity, sexual orientation and, as appropriate, other forms of diversity in domestic society in relation to mass communications;
- Demonstrate an understanding of professional ethical principles and work ethically in pursuit of truth, accuracy, fairness and diversity;
- Think critically, creatively and independently;
- Conduct research and evaluate information by methods appropriate to the communications professions in which they work;
- Write correctly and clearly in forms and styles appropriate for the communications professions, audiences and purposes they serve;
- Critically evaluate their own work and that of others for accuracy and fairness, clarity, appropriate style and grammatical correctness;
- Apply basic numerical and statistical concepts;
- Apply tools and technologies appropriate for the communications professions in which they work.

Academic integrity:

Every student must adhere to the policy on academic integrity set forth in the UNL Student Code of Conduct as outlined in the UNL Bulletin. Students who plagiarize may receive a failing grade on an assignment or for an entire course and may be reported to the Student Judicial Review Board. The work a student submits in a class must be the student's own work and must be work completed for that particular class and assignment. Students wishing to build on an old project or work on a similar project in two classes must discuss this with both professors. Academic dishonesty includes:

- handing in another's work or part of another's work as your own.
- turning in one of your old papers (including something you wrote in high school) for a current class.
- turning in the same or similar paper for two different classes,
- using notes or other study aids or otherwise obtaining another's answers for a quiz or an examination.

Anything and everything you include in your papers that comes from another source must be attributed with proper citation. That includes ideas and opinions. Plagiarism consists of using phrases, sentences or paragraphs from any source and republishing them without alteration or attribution. The sources include, but are not limited to, books, magazines, newspapers, television or radio reports, Web sites and other students' papers.

Students with disabilities:

Students with disabilities are encouraged to contact the instructor for a confidential discussion of their individual needs for academic accommodation. It is the policy of the University of Nebraska-Lincoln to provide flexible and individualized accommodation to students with documented disabilities that may affect their ability to fully participate in course activities or meet course requirements. To receive accommodation services, students must be registered with the Services for Students with Disabilities (SSD) office, 132 Canfield Administration, 472-3787 voice or TTY.

Diversity:

The College of Journalism and Mass Communications values diversity, in the broadest sense of the word – gender, age, race, ethnicity, nationality, income, religion, education, geographic, physical and mental ability or disability, sexual orientation. We recognize that understanding and incorporating diversity in the curriculum enables us to prepare our students for careers as professional communicators in a global society. As communicators, we understand that journalism, advertising and other forms of strategic communication must reflect society in order to be effective and reliable. We fail as journalists if we are not accurate in our written, spoken and visual reports; including diverse voices and perspectives improves our accuracy and truthfulness. In advertising, we cannot succeed if we do not understand the value of or know how to create advertising that reflects a diverse society and, thus, appeals to broader audiences.

Class Schedule:

A word of warning: This class is pretty fluid. I will move things up and back, depending on how well you're getting things. If things change, I will update the syllabus on Blackboard and I will update you.

Jan. 12, 2016**In class:**

- Introductions
- Syllabus
- What is data journalism?

Homework:

- Read: [Chapter 1 of the Data Journalism Handbook](#)
- Read: Read Meyer Chapter 1 (on Blackboard under Course Documents) and realize it was written in 2002.
- Read: Herzog Chapter 1.
- Read: The Myth of the Machine by Michael Berens in Nerds and Words on

Blackboard. If you have time, read a few others (like Steve Doig) and realize that document was written in 1999.

There will be a reading quiz. Alert me immediately if the bookstore doesn't have Herzog.

Jan 14, 2016

In class:

- Public Records: Your rights as a citizen and a journalist.
- The public records request assignment.
- More examples of data stories; how to start looking for them.

Homework:

- Read: Herzog chapters 2, 3 and 4.
- Read: [Sun Sentinel: Speeding Cops](#)
- Review [the records retention schedule](#) for an agency you're interested in. Find a database. Make sure it's not online. Be ready to pitch it as your semester long quest next class.

Jan. 19, 2016

In class:

- Records quest pitches
- The foundations of analysis: The general questions you ask.
- Summary statistics, central tendency, and more.

[Lecture slides](#)

Homework:

- Read: Herzog Chapter 8
- Read: Numbers in the Newsroom, Chapters 1 and 2.

Jan. 21, 2016

In class:

- Math quiz
- Basic Spreadsheets: rows, columns, cells, importing, sorting, filtering
- Applied analysis basics: calculating the formula for percent change

[Lecture slides](#)**Homework:**

- Using tax data from the Nebraska Department of Revenue, calculate the following:
- The change in Federal Adjusted Gross Income between 2000 and 2013 for every county in Nebraska.
- The average the number of exemptions per return in 2000 and 2013 for every county in Nebraska.
- The change in average Nebraska Net Taxable Income per return between 2000 and 2013 for every county in Nebraska.
- In a paragraph, explain what you did. How did you arrive at the answers you got? What steps did you have to take? And answer this question: If you had to explain how you got those answers to an editor, how comfortable are you that you could do that, and that you'd get the same answers?

Jan. 26, 2016**In class:**

- Basic Spreadsheets II - mean, median, min, max, ranking
- Does anyone see the problem here?
- Reliability, replicability, transparency.

[Lecture slides](#)**Homework:**

- [Install Anaconda](#).
- Read [Reducing barriers between programmers and non-programmers in the](#)

[newsroom](#)

Jan. 28, 2016

In class:

- Changing how we approach data journalism
- Working in the environment
- An in-class example of working with a dataset: sorting, means and medians

[Lecture Jupyter Notebook](#)

Homework:

- [Complete the assignment at the end of this walkthrough of Agate.](#)

Feb. 2, 2016

In class:

Snow Day: Class cancelled.

Homework:

Feb. 4, 2016

In class:

- More working in Jupyter: Group by and aggregates in Agate

Homework:

- More working with NU salaries: Group by, counting, averages and medians by job titles.
- Develop a story idea for your first data story. You will have to find a story, find some data and use Jupyter to analyze that data for your story. Be ready to pitch stories on Tuesday.
- Install [Open Refine](#).

Feb. 9, 2016

In class:

- Row-wise vs column-wise calculations in Agate
- Percent change calculations.

Homework:

[Complete the Percent Change walkthrough and the assignment at the end.](#)

Feb. 11, 2016

In class:

- Intro to data cleaning
- Intro to Open Refine

Homework:

- [Complete the data normalization walkthrough and assignment](#) and [here's a little help with it.](#)
- Install [Github Desktop](#)

Feb. 16, 2016

In class:

- Realizing Phil Meyer's dream of using all of the scientific method, especially the open and replicable parts.
- Examples
- Intro to GitHub

Homework:

- [Complete the GitHub assignment at the end of the walkthrough.](#)

Feb. 18, 2016

In class:

- Sanity checks and data smells
- Data smells in Agate

Homework:

[Complete the data smells assignment and the end of the walkthrough.](#)

Feb. 23, 2016

In class:

- Intro to joins

Homework:

Complete the join assignment

Feb. 25, 2016

In class:

Line edits of your first story. Sign up for a meeting time. No class meeting.

Homework:

Prepare a story pitch

March 1, 2016

In class:

- Story pitches
- Intro to data visualization as a reporting tool
- Data visualization in Jupyter

Homework:

Basic data visualization assignment TBD.

March 3, 2016

In class:

More data viz in Jupyter

Homework:

March 8, 2016

In class:

More data viz in Jupyter

Homework:

March 10, 2016

NICAR: No Class

March 15, 2016

In class:

- Line edits, sign up for a time, no class meeting.

Homework:

March 17, 2016

In class:

- Line edits, sign up for a time, no class meeting.

Homework:

Prepare your final story pitch.

March 22, 2016

Spring break: No class

March 24, 2016

Spring break: No class

March 29, 2016

In class:

- GIS in journalism -- stories with maps
- Intro to GIS
- When is a map not a map

Homework:

Prepare your story pitch GIS assignment

March 31, 2016

In class:

- Story pitches.
- More GIS

Homework:

GIS assignment TBD Reading Assignment TBD

April 5, 2016

Matt in Italy. Class TBD

April 7, 2016

Matt in Italy. Class TBD

April 12, 2016

In class:

- Writing with numbers
- Truth and epistemic justification

[Lecture slides](#)

Homework:

None

April 14, 2016

In class:

- Working with PDFs

Homework:

PDF assignment TBD

April 19, 2016

In class:

- Writing thousands of stories with data in seconds
- Story bots
- Ethical implications

Homework:

Prepare your open records request lightning talk and paper.

April 21, 2016

In class:

The Open Records Quest Lightning Talks;

Homework:

Final Projects are due on April 22 at 5 p.m.

April 26, 2016

In class:

Line edits, sign up for a time, no class meeting.

April 28, 2016

In class:

- The future of data journalism;
- How to get a job with what you've learned;