

JOU 4930, Data Journalism

Fall 2014 / Section 03C8 / Periods 3-5 on Mondays / Weimer G037 (basement news lab)



Associate Professor Norm Lewis, Ph.D.

Office: Weimer 3052

Office hours: Periods 5-6 (11:45 a.m. to 1:40 p.m.) Wednesday and Friday

Phone: 352-392-5137 (I respond within 24 hours Monday through Friday)

E-mail: nplewis@ufl.edu (I respond within 24 hours Monday through Friday)

About me: www.bikeprof.com

“Office hours” are times I dedicate to being in the office and are first-come, first-served. Otherwise, make an appointment or drop in.

ABOUT THE COURSE

Course Description

This course equips you to be hired as a data journalist.

The course is built around meeting the requirements in these two job ads posted in summer 2014:

- “The Sun Sentinel of South Florida is looking for a hungry and enterprising data journalist who lives to find and share the stories that data can tell.... The successful applicant will need strong data analysis skills, news judgment, and initiative. He or she must have advanced ability with Excel and Access, experience with mapping software and MySQL, and familiarity with web-scraping and data-visualization tools.”
- “New Mexico In Depth, a journalistic nonprofit committed to quality journalism, is seeking a data journalist to lead a team that builds and analyzes databases ... Requirements: Strong ability with Excel spreadsheets and cleaning methods; mastery of basic analysis, including Excel spreadsheets and relational databases; fluency in structured query language (SQL) is essential; knack for finding story ideas and trends within the data.”

Course Outcomes

Students will develop a “data frame of mind” by demonstrating proficiency in:

- Finding stories in data and finding data to report stories.
- Evaluating the strengths and weaknesses of data sources and methods.
- Acquiring data through public sources and by scraping websites and PDFs.
- Cleaning (fix mistakes in) data through Open Refine, Excel and Access.
- Analyzing data to find patterns and avoid erroneous conclusions.
- Writing SQL queries in Access to evaluate data and join databases.
- Joining and analyzing geographic data in ArcGIS.
- Creating basic data visualizations using Excel, Google Fusion and Tableau.
- Publishing a meaningful story based on acquiring and analyzing data.

Required Equipment

Bring your laptop. Any flavor of laptop is fine. We'll be using the [UF Apps](#) system so it won't matter whether you have a Mac or Windows device.

Required Software

None. All you need will be available to you free (paid by the student technology fee) via UF Apps or through public-domain software. You may benefit from Cometdocs and wish to use Tableau Desktop, which is more powerful than the free Tableau Public. See below for a sweet deal for both.

Required Textbooks



None. Instead, [join](#) IRE (Investigative Reporters and Editors). As a student, you get a one-year membership for only \$25. That gets you access to members-only tip sheets, story ideas and “how we did that” explainers. Membership allows you to join listservs so you can ask professionals to help you solve problems. And it gets you a free one-year license for the full version of Cometdocs (price: \$70) and Tableau Desktop (price: \$999). Pretty neat, eh?

Part of IRE is NICAR, the National Institute for Computer-Assisted Reporting. The 2015 NICAR conference is March 5-8 (that's during our spring break) in Atlanta. If you're serious about data journalism as a career, the conference is a great way to gain knowledge and get job leads.

ASSIGNMENTS AND GRADING

Grading Scale

	Percent		Percent		Percent		Percent
		B+	89-87%	C+	79-77%	D+	69-67%
A	100-93%	B	86-83%	C	76-73%	D	66-63%
A-	92-90%	B-	82-80%	C-	72-80%	D-	62-60%

Scores are rounded to the nearest whole point: 89.4 rounds down to 89 (B+) while 89.5 rounds up to 90 (A-). The UF grading policy can be found in the undergraduate catalog [online](#).

Grade Allocation

- Homework 40%
- Midterm exam 10%
- Final exam 15%
- Published data story 35%

Homework

You'll be given weekly assignments intended to reinforce the skills learned in class. Deadlines are inviolate, with no credit given for late work. However, you get 5 bonus points for turning work in 24 hours early. The two low scores will be dropped.

Midterm Exam

The midterm will cover material in the first half of the class. If you miss for any reason, a makeup exam is possible only if (a) the absence is covered by UF policy, (b) you call me at least 30 minutes before class begins, and (c) you promptly provide written documentation for the absence.

Final Exam

The final is cumulative. If you miss for any reason, a makeup exam is possible only if (a) the absence is covered by UF policy, (b) you call me at least 30 minutes before class begins, and (c) you promptly provide written documentation for the absence.

Published Data Story

During the semester, you'll find and develop a meaningful story involving data on any topic that interests you. The details are in the following rubric.

Category	Criteria
Topic (30%)	<ul style="list-style-type: none"> • Meaningful: Story answers a question the audience has • Compelling: Likely to be spread through social media or email • Timely: Relatively fresh; can't be old news • Scope: Covers who, what, when, where, why and how • Sourcing: At least 3 sources; full range of views; affected people interviewed
Data (40%)	<p>You will show me the steps you took to acquire, clean and analyze the data. The score depends on the quality of the data and the effort expended, such as:</p> <ul style="list-style-type: none"> • How the data was obtained (download vs. scraping) • Whether more than one set of data is used • How much cleaning was required and how it was done • Whether the analysis was just a simple sort or percentage or involved more elaborate techniques such as a pivot table (good) or a join (great!) • The accuracy and suitability of the data
Writing (20%)	<ul style="list-style-type: none"> • Length: Roughly 750 well-chosen words • Lead: 25 words max; either summarizes story or hooks reader • Concise: Gets right to the point; doesn't waste words • Readable: Smooth flow; active verbs; short sentences; descriptive • Clarity: Conveys meaning and impact with very few numbers – no more than one every other sentence, and no more than five numbers in the text
Visual (10%)	<ul style="list-style-type: none"> • Prepare at least one chart or map suitable for publication
<p>Deductions are made for inaccuracies (-50) or for ethical violations, such as quoting sources with whom you have a personal connection</p>	

Publication

This story must be published (not submitted) by Friday, Dec. 12. A good publication option is WUFT (website, radio and/or TV), which publishes or airs stories involving 18 counties in north

central Florida. Another professional news outlet or the Alligator is also OK. Supply evidence of publication (email to an online link) when your story is published.

If your selected news organization (say, WUFT online) has accepted your finished story, edited it and approved it for publication – but decides to hold it to publish the following week when news slows down – that will still count as “publication.” However, your story must be completely finished and completely approved by Friday, Dec. 12. Also, you must have the supervising editor or news manager let me see the story on Friday, Dec. 12 – because that’s the day I will be grading the stories. If the story isn’t publication-approved by then, two letter grades will be deducted from whatever score the story merits as of Dec. 12.

Remember: The early bird gets the worm. Waiting until deadline eve courts disaster.

Interim Steps

To keep you on track, meet (or exceed!) these interim deadlines:

- Oct. 20: Email me a one-paragraph summary of your story idea and identify the data source
- Nov. 17: Show me (office hours or appointment) how you have acquired, cleaned and analyzed your data, and describe your three sources you have or will interview
- Dec. 1: Come to class with four copies of your finished story so you can get feedback from your peers

Tentative Class Schedule (subject to change)

	Date	Theme	Topics
1	Aug. 25	Thinking Data	Data journalism defined; building a “data frame of mind”
	Sept. 1		Labor Day; no class
2	Sept. 8	About Data	Types of data; data pitfalls; data sources
3	Sept. 15	Spreadsheets	Excel basics
4	Sept. 22	Spreadsheets	Analytical tools in Excel; pivot tables
5	Sept. 29	Scraping	Acquiring data from websites and PDFs (Tableau, Cometdocs)
6	Oct. 6	Cleaning	Finding and repairing mistakes with Excel and Open Refine
7	Oct. 13	Midterm	Practice exercises; midterm
8	Oct. 20	Databases	SQL basics in Access
9	Oct. 27	Databases	Joining databases in Access
10	Nov. 3	Databases	Aggregate functions and commonly used queries in Access
11	Nov. 10	Mapping	Joining geographic-related data and evaluate results in ArcGIS
12	Nov. 17	Mapping	Display map data using Google Fusion
13	Nov. 24	Charts	Create online charts using Tableau
14	Dec. 1	Your story	Review drafts of stories with peers
15	Dec. 8	Final	Final exam

THE FINE PRINT

Attendance

Requirements for class attendance, make-up exams, assignments and other work are consistent with UF [policy](#). Exam make-ups are possible only if (a) the absence is covered by UF policy, (b) you call me at least 30 minutes before class begins, and (c) you promptly provide written documentation for the absence.

Academic Integrity

University of Florida students pledge to abide by an [honor code](#) that prohibits academic dishonesty such as fabrication, plagiarism and cheating. You have an affirmative obligation to understand what constitutes academic dishonesty. You also must report to appropriate personnel any condition that facilitates academic misconduct. If you have any questions or concerns, please contact me.

When I discover cheating, my default policy is to fail all involved for the entire course and report the details to the Dean of Students Office.

Students with Disabilities

Students with disabilities requesting accommodations should first register with the [Disability Resource Center](#) (352-392-8565) by providing appropriate documentation. Once registered, students will receive an accommodation letter, which must be presented to me when requesting accommodation. Such requests should be made as early in the semester as possible.

Help With Coping

The UF [Counseling and Wellness Center](#) is a terrific, free resource for any student who could use help managing stress or coping with life. The center, at 3190 Radio Road on campus, is open for appointments and emergency walk-ins from 8 a.m. to 5 p.m. Monday through Friday. To make an appointment or receive after-hours assistance, call 352-392-1575.

The UF Police can be reached at 392-1111 or, in an emergency, by dialing 911.

Course Evaluations

Students are expected to provide feedback on the quality of instruction in this course by completing online [evaluations](#). You will be notified by email when the evaluations are open, typically in the last two or three weeks of the semester. Summary [results](#) are available to you and the public.