JOU 3305: Data Journalism

Fall 2025 | 9:35 a.m. to 12:25 p.m. Thursdays | Weimer 2056

Lecturer Gavin Off
Phone: 352-392-6399
Email: gavin.off@ufl.edu
Office hours: Wednesdays and Thursdays in room 1200A. No appointment or any day via Zoom

MY EXPERIENCE:

I landed my first newspaper job in 2002. It was with a small paper in southwest Florida. I covered traffic accidents, local governments, police and everything in between. But by 2006, I was ready to move on and applied for a reporter position at the Louisville Courier-Journal in Kentucky. Two weeks later, I called and asked the editor if he received my resume. He said he got my resume and 200 others. Needless to say, I did not get the job.

Geez, I thought. What could I do to separate myself from other future job applicants?

Later that year, I enrolled at the University of Missouri's master's program. Missouri is home to Investigative Reporters and Editors and the National Association for Computer-Assisted Reporting – two nonprofits. That's where I learned data journalism.

I've worked in three newsrooms since graduating from Missouri in 2008. All three had mass layoffs. But my name was never called. Why not? I'm convinced it's because I know data journalism. I had a skill that no one else in the newsroom had and that brought value to the paper, which – at least as far as I'm concerned – saved my job.

COURSE DESCRIPTION

The world is immersed in data: Police calls, hunting accidents, bridge inspections, car wrecks and student test scores. Even names of neighborhood pets are often stored in a database somewhere. And good news – the information is public, open for reporters to analyze and understand. But how does someone analyze thousands, sometimes millions of records quickly? That's what you'll learn. This hands-on course teaches students how to use spreadsheets, and database managers to write data-centered stories, ranging from local features to in-depth investigations. It is open to non-majors.

COURSE OBJECTIVES

You'll be asked to learn Excel functions and some coding in R, a database manager. The ultimate goal: To become better journalists.

LEARNING OBJECTIVES

In this class you are reporters. The objective is to become better ones. To do that, you'll be asked to:
- Understand what data are publicly available
- Learn how to request data by using public records laws
- Scrape data from PDFs and websites without writing complex code
- Master cleaning and querying data by using spreadsheets and database managers

- Discover how to use data to guide reporting
- Find out how to weave data into storytelling
- Produce an accurate, fair, engaging publishable story based on data

OTHER (IMPORTANT) NOTES

- It is up to you to stay on pace, learn the material and do the work. I'll help every way I can, but you need to put in the effort.
- Pay attention to Canvas. This is a must.
- These classes build on each other. If you do not understand a concept, let me know. I want to catch you up to speed before you get too far behind.
- The more you use the skills that you learn in this class, the better data journalists you'll become. Practice, practice, practice.

BOOKS/MATERIALS

- Google Sheets.
- Excel. This is part of the Microsoft Office package available to students for free.
- R and Rstudio. We'll download it in class using these links: https://cran.r-project.org/ and https://posit.co/downloads/
- Recommended: "Data Literacy" by David Herzog. ISBN-10 1483333469
- Recommended: An 8-gb flash drive to store files.

GRADING

Your grade will be determined on the following scale:

A	93 – 100	B-	80 – 82.9	D+	67 – 69.9
A-	90 – 92.9	C+	77 – 77.9	D	63 – 66.9
B+	87 – 89.9	C	73 – 76.9	D-	60 – 62.9
B	83 – 86.9	C-	70 – 72.9		

A – Superior. This work shows a superior understanding of the concepts, research, and analysis required by the assignment with few, if any, data-related errors. All answers must be supported by work that is clear, accurate and reproducible.
B – Above Average. This work shows a very good understanding of the concepts, research, and analysis required by the assignment with minimal data-related errors. Answers must be supported by work that is clear, accurate and reproducible.
C – Average. This work demonstrates a basic understanding of the concepts but is deficient in the research and analysis required by the assignment and includes many data-related errors. Some of the work needed to produce the answers is unclear, inaccurate or not reproducible.
D – Below Average. This work does not demonstrate an understanding of the concepts, is seriously deficient in the research and analysis required by the assignment, and includes excessive data-related errors. Little or no work is shown.
F – Deficient. This work is deficient on almost all counts.

EARNING THE GRADE YOU WANT

- Homework – 35%
- Data journalism quiz – 15%
- Tests – 25%
- Final story – 25%

HOMEWORK ASSIGNMENTS

Most weeks you'll have some sort of data journalism task to complete. These will be short, focused assignments that cover the material we recently learned.

There will also be weeks in which you will file story memos that show your findings until that point. I will edit and grade these, and my edits are to be incorporated into your next memo. The point of these memos is simple: To make sure you're on track to finish a final story that is worthy of being published.

DATA JOURNALISM QUIZ

Data journalism quiz:

This quiz will test you on some of the data journalism basics, including – but not limited to:

- What is data journalism
- Why is data journalism important
- What it takes to write a data journalism story
- How to find data
- How to request data

TESTS

Excel test:

This will open-notes test your skills on importing, cleaning and analyzing data in Sheets. You'll be given raw data to clean and analyze using the skills learned in class. You'll also be given a document of possible Excel functions that you may be tested on.

R test:

This will test your skills on cleaning and querying data using the computer language R. You'll be provided the raw data. Like the Excel test, this, too, is open notes. You'll also be given a worksheet of possible R queries that you may be tested on.

FINAL STORY

This is a data-centered story that you – or a group of you – complete specifically for this class. The goal is to write a publishable story by the end of the semester. I will likely choose the dataset for you but please offer suggestions if there is good, available data on a topic that you're passionate about.

Keep in mind that this is a news story. It is meant to inform, educate or shine light on an issue.

To make sure you're on pace to complete the story, we'll have periodic check-ins, where you'll update me – and sometimes the entire class – on where your data analysis and reporting stands. These check-ins are to make sure that you're working on the story throughout the semester. This is not an assignment that you can complete in the final few weeks.

To do well, your story needs to include all the necessary components of a strong data-centered story. These components include: findings from the data, interviews with experts and people effected, terrific writing, concise storytelling and possibly additional documents.

CLASS POLICIES

Attendance is imperative. It is not part of your grade, but these classes build on each other. What you learn today will help you understand the skills I'll teach in the next class.

We all are expected to attend each scheduled class and arrive on time prepared to participate. This includes coming to class with a computer and the necessary software loaded onto your machine.

Missing class does not change a deadline. All assignments are due when specified regardless of whether you expect to be in class, unless I've approved an alternate deadline in advance.
If there are special circumstances that prevent you from making arrangements for work missed because of an absence, please contact me or have a family member contact me as soon as possible, and I will do my best to work with you. I may require documentation.

This is a hands-on class that requires your participation. Part of that participation means asking questions, especially if you do not understand or cannot replicate a data query. Please speak up. If you have questions about a code that we're writing, chances are other students have the same questions. I need to know so I can help.

Behavior policies:
- Do not have conversation in class. This is disruptive.
- Do not arrive late or leave early. This is disruptive.
- Silence and keep your personal devices out of site. They are disruptive.
- Think twice about bringing food to class.

SUBMISSION OF ASSIGNMENTS

All written assignments must be submitted on the day they are due according to the guidelines outlined above and per the assignment description. Typically, assignments will be due at 11:59 p.m. Mondays. Please note that I will try to give you feedback as quickly as possible, but I too can get busy as the semester progresses.

All grades will be posted on Canvas.

You may turn in an assignment late once without penalty. After that, I will not accept late assignments.

ACADEMIC INTEGRITY

UF students must adhere to The Honor Pledge which states, "We, the members of the University of Florida community, pledge to hold ourselves and our peers to the highest standards of honor and integrity by abiding by the Honor Code."

On all student work at UF, the following pledge is either required or implied: "On my honor, I have neither given nor received unauthorized aid in doing this assignment." The Honor Code (https://sccr.dso.ufl.edu/process/student-conduct-code) specifies a number of behaviors that are in violation of this code and the possible sanctions. Also, students are obligated to report any condition that facilitates academic misconduct to appropriate personnel. If you have any questions or concerns, please consult with me. Violations can result in a failing grade for the course and referral to the dean of students.

STUDENTS WITH DISABILITIES

Students with disabilities who experience learning barriers and would like to request academic accommodation should connect with the Disability Resource Center. It is important to share any accommodation letter with me and discuss access needs as early as possible in the semester.

DIVERSITY

The College of Journalism and Communications embraces an intellectual community enriched and enhanced by diversity along several dimensions, including race, ethnicity and national origins, gender and gender identity, sexuality, class, and religion. Each course is expected to help foster an understanding of the diversity of peoples and cultures and of the significance and impact of mass communication in a global society. To that end:

- Please let me know if you find any material in the course violates that expectation.
- Please alert me if you have a name or preferred pronouns that differ from the class roll information, which is my only source of information about you.
- If you have any concerns involving diversity in this course that you feel uncomfortable discussing with me, I encourage you to contact the department chair.

COURSE EVALUATIONS

Students are expected to provide professional and respectful feedback on the quality of instruction in this course by completing course evaluations online via GatorEvals. Guidance on how to give feedback in a professional and respectful manner is available at **https://gatorevals.aa.ufl.edu/students/.** Students will be notified when the evaluation period opens, and can complete evaluations through the email they receive from GatorEvals, in their Canvas course menu under GatorEvals, or via **https://ufl.bluera.com/ufl/**. Summaries of course evaluation results are available to students at **https://gatorevals.aa.ufl.edu/public-results/**.

IN-CLASS RECORDING

The official UF policy regarding in-class recording, to comply with a 2021 Florida law can be found here.

HEALTH AND WELNESS

- U Matter, We Care: If you or someone you know is in distress, please contact umatter@ufl.edu, 352-392-1575, or visit U Matter, We Care website to refer or report a concern and a team member will reach out to the student in distress.
- Counseling and Wellness Center: Visit the Counseling and Wellness Center website or call 352-392-1575 for information on crisis services as well as non-crisis services.
- Student Health Care Center: Call 352-392-1161 for 24/7 information to help you find the care you need or visit the Student Health Care Center website.
- University Police Department: Visit UF Police Department website or call 352-392-1111 (or 911 for emergencies).
- UF Health Shands Emergency Room / Trauma Center: For immediate medical care call 352-733-0111 or go to the emergency room at 1515 SW Archer Road, Gainesville. Visit the UF Health Emergency Room and Trauma Center website.

ACADEMIC RESOURCES

- E-learning technical support: Contact the UF Computing Help Desk website, or phone 24/7 at 352-392-4357, or email helpdesk@ufl.edu.
- Career Connections Center: Career assistance and counseling services. Visit the

website. Reitz Union Suite 1300, 352-392-1601.
- Library Support: Various ways to receive assistance with respect to using the libraries or finding resources. Visit the website.
- Teaching Center: General study skills and tutoring. Visit the website. Broward Hall, 352-392-2010 or to make an appointment 352- 392-6420.
- On-Campus Student Complaints: Details are available through the Student Honor Code and Student Conduct Code, also known as the Orange Book.

COURSE SCHEDULE

(Please note: The course schedule may change depending how well students are grasping the techniques described in class. ***Pay attention to Canvas for any changes.****)*

Week 1 – 8/21: Intro to data journalism
- What is data journalism? Why is it important? And what stories can we tell with data?
- *HOMEWORK: Student introductions. Due date: 8/25.*

Week 2 – 8/28: Finding data and importing it into Excel
- Data is everywhere. What records are public and how do you get them?
- We'll find data on the web and learn how to request datasets.
- Then we'll import data into Excel and sort and filter it.
- *HOMEWORK: Find an example of data online. Turn in a link to the data, the name of the agency that collects the data and 100-150 words about the data. Due date: 9/1.*
- *HOMEWORK: Excel – sort, filter and find. Due date: 9/1.*

Week 3 – 9/4: Data journalism quiz AND cleaning data in Excel
- This quiz will test you on the basics of data journalism – the first two and a half classes.
- After the quiz, we'll use Excel to clean dirty data.
- UPPER, LOWER, PROPER, LEFT, MID, RIGHT, CONCATENATE and text to columns – learn what these mean and how they help clean dirty data.
- *HOMEWORK: Excel – cleaning dirty data. Due date: 9/8.*

Week 4 – 9/11: Doing math in Excel
- Journalists typically hate math, but Excel makes math easy.
- *HOMEWORK: Excel – math. Due date: 9/15*
- *HOMEWORK: Story memo No. 1 (More on this to come)*

Week 5 – 9/18: Dealing with dates and advanced formulas
- Dates in data can be crucial. Let's find out how to handle them.
- Even the "advanced" formulas – SUBSTITUTE, IF, COUNTIF and TRIM – can be easy.
- *HOMEWORK: Excel – formulas. Due date: 9/22*

Week 6 – 9/25: Creating pivot tables
- Pivot tables are a powerful tool for cleaning grouping and analyzing data. Learn how to quickly make sense of thousands of records by simply dragging and dropping column headers into the proper locations.
- HOMEWORK: Install Tabula
- *HOMEWORK: Excel – dates and pivot tables. Due date: 9/29*

Week 7 – 10/2: Scraping data from the web and PDFs

- Sometimes it's easier to "scrape" data than request it. ImportHTML and tools like Tabula and pdftoexcel can help.

Week 8 – 10/9: Excel test

- This open-notes test will cover everything we learned with Excel.
- After the test, we'll discuss the final story project
- *HOMEWORK: Install R and Rstudio. Due date: 10/20*
- *HOMEWORK: Story memo No. 2 (More on this to come)*

Week 9 – 10/16: Importing data and filter data in R

- Here, we'll continue our talk about the semester's final story, which could be done individually or in groups.
- Why do so many journalists love R? What is a package?
- Here, we'll import data into R and learn to filter it using FILTER, GREPL, STR_DETECT and ARRANGE.
- *HOMEWORK: R – filtering. Due date: 10/20*

Week 10 – 10/23: Grouping records in R

- R can aggregate records and perform mathematical calculations, allowing reporters to identify trends, patterns and anomalies that make up the backbone of a story.
- Let's add GROUP_BY, SUMMARIZE and some math.
- *HOMEWORK: R – grouping. Due date: 10/27*

Week 11 – 10/30: Joining records in R

- One of R's most powerful tools is making data frames talk to one another.
- How? By using INNER, LEFT, ANTI AND FULL joins.

Week 12 – 11/6: Creating charts in R

- R can do it all, including making charts with Ggplot.

Week 13 – 11/13: R test

- This test will cover everything we learned in R
- It's open notes
- *HOMEWORK: Story memo No. 3 (More on this to come)*

Week 14 – 11/20: Final story review

- Let's talk one more time about your story
- *HOMEWORK: Final story to be turned in by the day of the final exam. Due date: 12/8*