

JOU 4306 Advanced Data Journalism

Spring 2024 | Class 21644, section ADV1 | Periods 7-9 Mondays | Weimer 2050 | 3 credits



Associate Professor Norman P. Lewis, Ph.D.

Office: 3052 Weimer Hall

Contact: Email nplewis@ufl.edu or message through Canvas

Office hours: Mondays 12:00p to 1:30p and Thursdays 1:00p to 2:30p (in person)

Office hours are in-person (in my office) and drop-in (no appointment).

If those hours conflict with your schedule, contact me and we will find a time that works.

PART 1: ABOUT THE COURSE

Course Description

Blend journalism and data science to become a full-time data journalist. Program in R to replace spreadsheets and databases with reproducible data analysis. Clearly communicate results accurately to internal (newsroom journalists) and external audiences.

Prerequisites

Completion of either fall data course: JOU 3305 Data Journalism or JOU 4318 Sports Data Journalism. No prior coding experience expected.

Course Objectives

To be hireable as a full-time data journalist (the in-house expert) who can:

- Program in R to analyze and communicate data for common news topics.
- Acquire and parse data from standard file formats and APIs.
- Analyze census data to originate or supplement news coverage.
- Apply the Tidyverse package to common journalistic analyses.
- Conduct basic statistical analysis using R's built-in tools.
- Collaborate on data projects using GitHub and version control.
- Communicate data analysis for internal and external use on websites.

Required Materials

- Basic laptop, either Mac or Windows.
- No textbook. No materials and supplies fee.

IRE Membership



For \$25 (student rate), [join](#) Investigative Reporters and Editors (IRE), which includes NICAR for data journalists. Get [access](#) to 30,000 ideas, stories, tip sheets, data – and job prospects. Perfect organization for data journalists.

Tentative Schedule (subject to change)

	Date	Theme	Skills	Assignment
1	Jan 13	Univariate	Single-Variable Analysis Tidyverse essentials and ggplot charts. Quartile analysis.	H1: Univariate Due: Fri Jan 17
2	Jan 20		No class: MLK Jr Day	
3	Jan 27	Bivariate	Analysis with 2 variables Confidence levels. Anchor y-axis at zero. Scatterplot with trend lines and error bars.	H2: Bivariate Due: Fri Jan 31
4	Feb 3	Wrangle	Import, Format, Dates Import from Socrata websites, CSV, TSV. Convert data types. Date math and parsing. Chart scales.	H3: Wrangle Due: Fri Feb 7
5	Feb 10	Assemble	Joins, Text Search, Normalize Joins. Combine CSVs by row or col. Strings: Detect, convert case, handle NA, collapse space.	H4: Assemble Due: Fri Feb 14
6	Feb 17	Format	Cleaning, String Parsing Clean variable names, data, and numbers. Strings: locate, parse, separate, combine, remove, replace.	H5: Format Due: Fri Feb 21
7	Feb 24	Refine	Refining Chart Appearance, Tidy Data Color usage. Annotations. Themes. Convert to/from tidy formats. Stacked & grouped column charts.	H6: Refine Due: Fri Feb 28
8	Mar 3	Uni_Pay	How to Complete a Project How to interview data. Charts: annotations, value labels, color highlight. If_else.	H7: Fire Due: Fri Mar 7
9	Mar 10	RPubs	Publishing Project Online via RPubs Use markdown within R Studio to publish a project, with tables and charts, on a public website.	Exam 1 (at home) Due: Fri Mar 14 Plan due Mar 14
10	Mar 17		No class; spring break	
11	Mar 24	(no class)	No class so you can work on your project. H8 due Friday, March 28. Revised project plan due March 28.	H8: SOMETHING Due: Fri March 28 Revise due Mar 28
12	Mar 31	Census	Obtain and Analyze Census Bureau Data API key. Census tables, geographies. String remove. Histogram. Visual margin of error. Beeswarm.	H9: Census Due: Fri April 4
13	Apr 7*	Geospatial	Geospatial Analysis Census maps and geometries. Point analysis. Geo data. Projections. Geocoding. Simple Features.	H10g: Geo Due: Fri April 11
14	Apr 14*	Sports	Sports Data via R Packages R packages to get play-by-play data in pro sports: MLB, NFL, European football (soccer)	H10s: Sports Due: Fri April 18
15	Apr 21	Project Demo	Project Demonstration Show your completed (not a draft) project in R Pubs and Markdown to receive (and give) feedback	Data Project Due: Wed April 23
			Exam 2	Exam 2 (at home) Due: Wed Apr 30

* Your choice: Attend either April 7 or April 14 class, or both. Do either H10g or H10s, or both.

PART 2: ASSIGNMENTS AND GRADING

Attendance and Deadlines

Because this is a hands-on course, weekly attendance is expected. Deadlines are firm.

Grade Allocation

Weekly Homework 50% Due 11:59 pm Fridays. Best 9 scores count.
 Project 30% Due 11:59 pm Wednesday, April 23.
 Exam 1 10% Due 11:59 pm Friday, March 14.
 Exam 2 10% Due 11:59 pm Wednesday, April 30.

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-	72-70%	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](#) details how GPA is computed.

Homework (50%)

Homework is assigned after each class and is due before 11:59 pm on Fridays. The low score drops. Because you have 4 days to complete each assignment, no extensions will be allowed unless you have an exemption covered by UF policy for the entire time period.



Bonus: Each homework submitted before 5 pm Wednesday earns 10 extra points. Why the bonus? Completing work within 2 days improves learning and retention.

Extra Credit

The only extra credit is the weekly opportunity to earn a full letter grade by turning in each week’s homework early (above). No other option is available.

Exam 1 (10%) and Exam 2 (10%)

- Like the quizzes, with more extensive data sets.
- Each is a take-at-home exam. Each is open-book, but only-you.
- Each has a full week to complete. No extensions are possible unless documented and covered by UF policy for the entire week (and not the last minute).
- Exam 1: Due before 11:59 pm Friday, March 14.
- Exam 2: Due before 11:59 pm Wednesday, April 30.

Project (30%)

- For grading, see project rubric below.
- See project deadlines below for progress dates (each with 1-point bonus)

Project Rubric

Assignment: Create web page of your original data analysis, for colleagues and public.

Work product: (1) URL of your R Pubs page, (2) R script used to analyze data, and (3) the data.

Grading standard: An A project would get you an interview with a potential employer.

Category	Criteria	Pts
Data (40%)	1.1 Topic. Is the project driven by an explicitly stated question that is substantial and meaningful?	5
	1.2 Data. Does the project use data from the original source that is current, reliable, and substantial enough to require analysis?	5
	1.3 Analysis. Has the analysis been conducted in a way that substantially adds value by finding fresh and meaningful insights or patterns?	10
	1.4 Normalize. Have the data been normalized as necessary to compensate for uneven measures such as time, population, or inflation?	10
	1.5 Calculations. Have calculations (percentage, rate, ratio, mean, median, etc.) been made, and were they appropriate for the data?	10
Web Page (40%)	2.1 Clarity. Is the page organized for clarity, allowing a colleague unfamiliar with the project to quickly discern its scope and meaning?	5
	2.2 Table. Does the page offer at least one meaningful, searchable table that resulted from the project’s analysis?	5
	2.3 Visuals. Does the page offer at least 3 charts or maps that illuminate meaningful patterns, supplemented by readable text and annotations?	10
	2.4 Mechanics. Is the text clear, accurate, and mechanically correct in spelling and grammar?	5
	2.5 Suggestions. Does the page conclude with at least 2 suggested stories or analytics that would be meaningful and interesting to the audience or boss?	5
	2.6 Complete. Is the story complete, without glaring omissions, and yet concise, without padding or superfluous elements?	5
	2.7 References. Are all sources for text and data referenced with embedded hyperlinks?	5
Code (20%)	3.1 Accurate. Is the code used to complete the assigned tasks and analysis correct both technically and in substance?	5
	3.1 Transparent. Does the R Studio document offer ample # comments that explain the code for a novice proofreading the analysis?	5
	3.2 Repeatable. Can the code be reused with minimal changes if/when the data are updated?	5
	3.3 Efficient. Is the code concise in commands and structure, using just enough to complete the task?	5

Project Deadlines

Bonus percentage points are available per the details below. If you complete all 3, a project with a score of 85 would be raised to 88.

Assignment	Description	Bonus	Due Date
1. Project Plan	<p>Upload a Word doc completing these 10 items:</p> <ol style="list-style-type: none"> Question: The specific, non-binary, timely question driving my project is: Data: The dataset(s) I will analyze is/are: Source: The data are from (hyperlink to source): Variance: This dataset is likely to show variance or patterns because: Analysis: The analysis I will conduct in R to add value to this data involves (specify): Table: The table I will produce online will be approximately (number) of observations and show these variables (specify): Chart 1: Likely will be (specify headline, data to be shown, and type of chart): Chart 2: Likely will be (specify headline, data to be shown, and type of chart): Chart 3: Likely will be (specify headline, data to be shown, and type of chart): Story: The project is likely to enable others in my organization to develop their own story or analyses such as these two: <p>1% awarded <u>only</u> if all 10 completed <i>thoroughly</i>.</p>	1 pct point	Fri, Mar 14
2. Revised Project Plan	<p>Submit a revised plan in response to feedback. Complete the same 10 elements above.</p> <p>1% awarded <u>only</u> if all 10 completed <i>thoroughly</i>.</p>	1 pct point	Fri, Mar 28
3. Demo Data Project	<p>Demo completed (not draft) project in class.</p> <p>1% awarded <u>only</u> if story is <i>completed</i> and shown to the class on RPubS and RMarkdown.</p>	1 pct point	In class Mon Apr 21

1. I invite you to discuss your **idea** with me informally before March 14.
2. On March 14, a well-defined and researched project **plan** is due.
3. For the record, an idea is not a plan. I need to see a specific, well-defined **plan**.

PART 3: OTHER IMPORTANT DETAILS

Academic Integrity

UF students are bound by 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 work submitted for credit by students at the University of Florida, 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, you 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.

In-Class Recording

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

Artificial Intelligence

Use of AI apps or assistants (Chat GPT, Gemini, Copilot, etc.) is allowed *only* if you:

1. Record the source and prompts you used and include them with your assignment.
2. Detail what changes you made to improve or repair errant coding.

Failure to do either will result a 50% reduction in points.

I do not recommend using AI in this course because it can be a crutch that inhibits learning. Success in a coding job depends on your ability to write code creatively, accurately, and efficiently. If all coding can be done via AI, no employer needs to hire you. Better to develop a skill set that makes you essential and employable by delivering more than Chat GPT can.

Students with Disabilities

Students with disabilities who experience learning barriers and would like to request academic accommodations should connect with the Disability Resource Center by visiting the [get-started page](#). 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 multiple dimensions. My goal is to foster an atmosphere in which we better understand and value our commonalities and differences within the United States and among national cultures. To that end:

1. Please let me know if you find any material in the course violates that expectation.
2. 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.

3. If you have any concerns involving content in this course that you feel uncomfortable discussing with me, I encourage you to contact Professor Joanna Hernandez, CJC director of inclusion and diversity, at jhernandez@jou.ufl.edu.

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](#). 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 [Bluera](#). Summaries of course evaluation results are [publicly available](#).

Health and Wellness

- **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](#).