JOU 4930 Advanced Data Journalism

Spring 2023 | Class 25101, section NT1O | Periods 6-8 Mondays | Weimer 2056 | 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 4:00 to 5:00 pm and Thursdays 1:00 to 2:30 pm

In-person office hours are first-come, first-served. If those hours conflict with your schedule, contact me and we will find a time that works.

PART 1: ABOUT THE COURSE

Course Description

The course blends journalism and data science to equip you to be a full-time data journalist. Learn to program in R to replace spreadsheets and databases with reproducible data analysis, and to clearly communicate results to peers and lay audiences.

Prerequisites

Required: Completion of either JOU 3305 Data Journalism or JOU 4930 Sports Data Journalism (fall). Recommended: Prior or concurrent enrollment in JOU 3121 DataViz and Mapping (spring). No prior coding experience necessary.

Course Objectives

To be hirable 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.
- Apply statistical significance testing to common journalism situations.
- Analyze census data to originate or supplement news coverage.
- Create point and choropleth maps to detect meaningful patterns.
- Collaborate on data projects using GitHub and version control.
- Communicate data analysis for internal and external use via Quarto.

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	Emphases	Assignment
1	Jan 9	Univariate	*Focus*: Analysis with 1 variable. *Skills*: tidyverse: filter, group, summarize, arrange, mutate; ggplot aes, geom_col. Quartile analysis. *Charts*: bar, column.	H1: tba Due: Fri Jan. 13
2	Jan 16		MLK Jr Day; no class	
3	Jan 23	Bivariate	*Focus*: Analysis with 2 variables. *Skills*: confidence levels, anchor y-axis at zero. C*harts*: scatterplot with trend lines and error bars.	H2: tba Due: Fri Jan. 27
4	Jan 30	Wrangle	*Focus*: Importing, exporting, formats, dates. *Skills*: Sheets and Excel import, export, convert data types, handling dates, standard deviation.	H3: tba Due: Fri Feb. 3
5	Feb 6	Assemble	*Focus*: Joins, text search, normalization. *Skills*: inner join, left join, right join, anti-join, combine CSVs by row or col, string detect functions, handling NA, converting text case, collapsing spaces.	H4: tba Due: Fri Feb 10
6	Feb 13	Format	*Focus*: Cleaning, tidy data, string parsing. *Skills*: cleaning variable names, messy data, numbers; string parsing; convert tidy formats; slice. *Charts*: grouped column, stacked column, filled stacked column.	H5: tba Due: Fri Feb 17
7	Feb 20	Refine Charts	*Focus*: Refining charts. *Skills*: scales, axis labels, color usage by variable type, chart annotations and titles, chart themes. *Charts*: scale to 100%.	H6: tba Due: Fri Feb 24
8	Feb 27	Project	*Focus*: How to complete a project. *Skills*: collapse factors, split axis labels, value and % labels, text labels, string substitution, color highlights, if_else.	H7: tba Due: Fri Mar 3
9	Mar 6	Sharing	*Focus*: GitHub, version control, Quarto. *Skills*: GitHub for sharing, collaboration, and backup; web publishing with Quarto.	Midterm Due: Fri Mar 10
10	Mar 13		Spring break	
11	Mar 20	Statistics	*Focus*: Applied statistics for journalism. *Skills*: chi-square, t-test, correlation, simple linear regression, multiple linear regression. *Charts*: boxplot.	H8: tba Due: Fri Mar 24
12	Mar 27	Census	*Focus*: Analyzing Census Bureau data. *Skills*: API key, census tables and geographies, string remove. *Charts*: histogram, visual margin of error, beeswarm.	H9: tba Due: Fri Mar 31
13	Apr 3	IPUMS	*Focus*: Importing and analyzing IPUMS micro census data *Skills*: IPUMS data.	H10: tba Due: Fri Apr 7
14	Apr 10	Geospatial	*Focus*: Geospatial analysis. *Skills*: Census maps and geometries, Shapefiles, GeoJSON, geocoding, CRS, Jenks natural breaks. *Maps*: point, choropleth.	H11: tba Due: Fri Apr 14
15	Apr 17	Presentation	Classroom presentation of Quarto project to get peer feedback before final submission.	Project Due: Fri Apr 21
16	Apr 24	No class	Exam 2 (at home)	Exam 2 Due: Fri Apr 28

PART 2: ASSIGNMENTS AND GRADING

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

Grade Allocation
Homework (10 highest count; lowest drops) 50%
Project (due 11:59 pm Friday, April 21) 25%
Midterm (due 11:59 pm Friday, March 10) 10%
Exam 2 (due 11:59 pm Friday, April 28) 15%

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.

Better yet, you have an incentive to get it done early. If you complete the homework within 48 hours (before 4 pm Wednesday), you get 10 bonus points.

Project (20%)
Analyze data and create a public Quarto Pub website, as a link you could share with an employer. This is like a job audition. An A project is so good an employer would hire you.

Any journalistic-style topic is acceptable. The best projects (1) are driven by a meaningful question, and (2) analyze at least 1 meaningful, current, and substantial data set. Double-dipping with another class (such as environmental journalism) is allowed, provided you clear it with me beforehand, to ensure the submissions are distinct.

Submit to the Canvas assignment a .zip file containing your Quarto script and a data folder unless you obtain data online through the Quarto script. Rubric:

Data Analysis (40%)
- Driven by a meaningful question
- Relies on at least 1 meaningful, current, and substantial data set
- Created for an audience of less data-savvy peers in your organization
- Identifies 1 or more meaningful stories that peers could pursue
- Complete yet concise; no meaningful gaps yet quality over quantity

- Data are normalized as necessary to avoid misleading conclusions
- Calculations (percent, rate, mean, median, etc.) are appropriate for the data
- Statistical analysis is not necessary, but if performed, appropriate for the data

Table and Visuals (40%)
- At least 1 substantial, searchable table for peers using DT package
- At least 3 visual elements (charts/maps) that illuminate meaningful patterns
- Charts and maps follow best practice for readability, clarity, and precision

Quarto Code (20%)
- Transparent: Ample comments explain code for a non-R user
- Repeatable: Code can be reused with few or no changes if data are updated
- Embedded hyperlinks to sources used as appropriate

Midterm (15%) & Exam 2 (15%)
Like a job tryout. You will be given data and a list of requirements to produce with code. You will have at least 1 week to complete. The test is open book but must be solely your work. The work product is an R script. Exam 2 is cumulative. An A exam is one that is:
- *Complete*. It fulfills all requirements specified.
- *Self-explanatory*. Create graphs or tables with sufficient information so that anyone could look at them and understand what they show.
- *Replicable*. Anyone who knows R should be able to understand and repeat the steps you took to test that your results are fair and accurate. Create an outline in the R script and use comments liberally. I will run replication tests.
- *Repeatable*. You will be given a data set that is updated regularly. Assuming the data structure (variables, data types, etc.) remains unchanged, your script should be re-useable a year from now.

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.

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 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:

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 diversity 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 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 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](#).