JOU 4318 Sports Data Journalism

Fall 2024 | Class 23177, Section HW01 | Periods 6-8 Wednesdays | 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: In person, in office (no appointment): Wednesday 11:00am to 12:30pm
                  In person, in office (no appointment): Thursdays 1:45 to 3:15pm
                  Zoom: Make appointment (I turn on Zoom only by appointment)

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. We can schedule meetings via Zoom when both our schedules allow Monday through Friday.

PART 1: ABOUT THE COURSE

Course Description

This course helps you think empirically about sports journalism and create an original story by acquiring and analyzing sports data. It is a hands-on introduction to spreadsheets, databases, web scraping, and statistics. A working knowledge of journalism is required, but no prior data expertise necessary. Open to non-majors.

Course Outcomes

By the end of the semester, you should be able to:
- Think originally about sports data beyond team wins and individual scoring.
- Analyze sports data to find original, meaningful stories.
- Scrape data from website tables without customized coding.
- Clean (standardize) data so it can be categorized and analyzed accurately.
- Use statistics to detect meaningful patterns and differences in data.
- Contextualize numbers through normalization and fair comparisons.
- Produce a publishable sports data story suitable for a job application.

My Related Data Courses

- JOU 3121 Dataviz and Mapping (spring): online data viz & geospatial analysis
- JOU 4306 Advanced Data Journalism (spring): data coding using R
- JOU 3305 Data Journalism (fall): a general-topic version of this course

Required Materials

- Basic laptop, either Mac or Windows
- Spend $12 for an annual license for a Chrome browser extension for Table Capture
- Download Office 365 on your laptop free through UF Tech

Tentative Schedule (subject to change)

Wk	Date	Theme	Key Topics	After Class
1	Aug 28	Finances	*How does UF rank in athletic spending?* Excel: format, sort, filter, transpose, sum, mean, median, percent of whole, percent change	H1: Finances Due Sept. 1
2	Sept 4	Salaries	*Which position in baseball pays the best?* Excel: pivot table, histogram, Analysis ToolPak	H2: Salaries Due Sept. 8
3	Sept 11	Correlation	*Is skill in 2PT and 3PT shots correlated?* Excel: correlation, if-else, scatterplot	H3: Correlation Due Sept. 15
4	Sept 18	Equity	*How does UF rank in Title IX?* Excel: ratios, pct points. Other: complete project	H4: Equity Due Sept. 22
5	Sept 25	Scrape	*Where does UF get its athletes?* Web scraping via Google Sheets, Table Capture; Excel: import, text to columns, trim, absolute cell reference	H5: Scrape Due Sept. 29
6	Oct 2	Home	*Which sport has the biggest home advantage?* Excel: =ifs, extract date parts year, month, weekday	H6: Home Due Oct. 6 Project Plan Due Oct. 7
7	Oct 9	Parse	*Which WNBA position pays the best?* Excel: left, right, len, search, Flash Fill	H7: Parse Due Oct. 13
8	Oct 16	Midterm	(In class; open book)	
9	Oct 23	Project Plan Presentation	Present data project plan in class, per syllabus. Offer suggestions to help classmates.	H8: Esports Due Oct 27 Project Revise Due Oct. 28
10	Oct 30	Probability	*Is there such a thing as a "hot hand"?* Excel: exponents. Other: Expected value, probability	H9: Probability Due Nov 3
11	Nov 6	Individual Performance	*Who is the GOAT?* Excel: standard deviation, Z score, conditional formatting	H10: Persons Due Nov. 10
12	Nov 13	Team Performance	*What changes will most help your team win?* Excel: regression model	H11: Teams Due Nov. 17
13	Nov 20	Exam 2	(In class; open book)	H12: Hoops Due Nov. 24
	Nov 27		Thanksgiving No class	
14	Dec 4	Final Project Presentation	Present your completed data project in class. Offer suggestions to help classmates.	Data Project Due Dec. 8

PART 2: ASSIGNMENTS AND GRADING

Semester Grade Allocation

Category	Pct	Details
Homework	50%	Weekly. Best 10 of 12 count. Low 2 scores drop.
Midterm	15%	In physical classroom Oct. 16. 3 hours.
Exam 2	15%	In physical classroom Nov. 20. 3 hours.
Data Project	20%	See schedule below. Show final project in class Dec. 4. Due Dec. 8.

Deadlines are firm. If issues arise during the semester, contact me <u>before</u> deadline.

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.

Attendance Policy
Although no participation is not graded, regular class attendance is expected. This is a hands-on, learn-by-doing class.

Homework (50%)
Homework is open after each class. Each is due at 11:59 pm on Sunday. The 2 low scores drop and the highest 10 count. Each homework has 10 questions worth 10 points each.

Because you will 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 within 48 hours (before 6 pm on Friday) earns 10 bonus points. For the semester, that is 100 points, or a full letter grade.

Exams: Midterm (15%) and SQL Exam (15%)
Each exam is in person during the three-hour class. Each is an open-book, and open-Web (though must be your own work). Exams are like the homework assignments, only under deadline, as if you were in a job tryout.

A makeup exam is possible only if all three conditions are met:
1. The absence is covered by UF policy.
2. You contact me at least 30 minutes before class begins.
3. You promptly provide written documentation if appropriate.

Data Story (20%)
An original data story, solely by you, and unique to this class. Details in the next section.

DATA STORY

Grading: An A data story is ready for publication and likely to impress a potential employer.

Definition
A data-driven story is not just a story with a number. It meets three criteria:
1. Addresses an empirical, timely, and meaningful question.
2. Answers by original analysis of sufficient data requiring a spreadsheet.
3. Analysis summarized by a single data point in the lede.

Examples
1. Addresses an empirical, timely, and meaningful question.
   - **Yes**: How does UF athletic spending compare with peer institutions over the past five years?
     Good question. Also, enough data to compare fairly.
   - **No**: How do the amenities in UF's new $85 million football facility compare with those at schools like Clemson and Oregon?
     More qualitative than empirical. Insufficient data to form a fair comparison.

2. Answers by original analysis of sufficient data requiring a spreadsheet.
   - **Yes**: Quarterbacks who are effective scramblers are associated with offenses that rank among the league's best, an analysis of 5 years of data shows.
     Required spreadsheet analysis to calculate correlation coefficient.
   - **No**: The quarterbacks over the past 5 years who have the highest quarterback ratings are (player names).
     NFL generates these stats, so no spreadsheet analysis required beyond a sort.

3. Analysis summarized by a single data point in the lede.
   - **Yes**: Head football coach salaries in the SEC have grown faster than in any other conference, nearly doubling in the past 5 years.
     Single data point (rank: faster than any other conference). Also, fair comparison.
   - **No**: College football salaries keep increasing, as head coaches become more expensive and universities spend more on athletics.
     No implicit data point, and no insight.

How to Find a Data Story
1. Begin with a meaningful (not obvious) question.
2. Find data that can answer that question.
3. Analyze the data to find a meaningful result.

Three-Part Submission
1. Main Story (for the audience) | Word document.
2. Nerd Sidebar (for the audience) | Word document.
3. Data: spreadsheet(s) and data diary (for me) | Excel spreadsheet.

Data Project Grading Rubric (80 points total out of 100 for assignment)
The rubric presumes the story is factually accurate, without meaningful errors in calculation or interpretation. Any such errors result in deductions up to 100 points.

1. Main Story (65 points)

Word doc. About 300 to 400 words. Focus on clarity and concision. Lede is most important sentence.

Category	Criteria	Points
1. Topic	Driven by an (a) empirical, (b) timely, and (c) meaningful question for which the answer is not obvious.	10
2. Data	(a) Original to source and reliable. (b) Current. (c) Sufficient in quantity and (d) over enough time periods to fairly answer question. (e) Enough variance to produce a meaningful pattern.	10
3. Analysis	(a) Original analysis (b) requiring a spreadsheet to answer the question (c) beyond a sort or filter.	5
4. Measures	(a) Appropriate measures used such as average, rate, percent, or ratio, and (b) normalize if required.	5
5. Data Point	Analysis produced (a) an original, meaningful, and fair result (b) summarized by a single number, and (c) contextualized as appropriate.	5
6. Lede	(a) A single sentence of (b) 25 words max (c) focused on the pattern detected (d) with 1 number max, and (e) sufficient to summarize story.	5
7. Headline	(a) Engaging. (b) Matches lede but (c) does not repeat it. (d) 10 words max. (e) Includes present-tense verb.	5
8. Scope	Story is (a) publication-ready. It is both (b) concise and (c) complete, without glaring omissions.	5
9. Writing Mechanics	(a) Complete sentences that are (b) grammatically correct and (c) accessible to a general audience, (d) without cliches, jargon, or acronyms.	5
10. Writing Numbers	(a) One number max per sentence. (b) Five numbers max per story. (d) Prefer words when reasonable without losing precision.	5
11. Journalistic	(a) Tone is neutral, fair, and detached. (b) Write in the third-person. (c) Embed hyperlinks to the data and any sources cited.	5

2. Spreadsheet (15 points)

Spreadsheet with at least 2 sheets.

1. Spreadsheet Data	Include your data and your analysis. Can involve multiple sheets. I need enough detail to replicate your steps to test for accuracy.	10
2. Data Diary	In a sheet labeled "diary," detail the steps you took, as appropriate, to (1) clean or standardize, (2) calculate measures, (3) refine and re-calculate.	5

Data Project Milestones (20 points total out of 100 for assignment)

Assignment	Description	Points	Due Date
1. Project Plan Email	Email me with these four points: 1. The specific question you want to answer. 2. Source: The specific data will you use, with a hyperlink to the source or a copy of the data file. 3. Method: How you will clean and analyze the data. 4. Lede: The single, specific data point you hope to find, written as a lede sentence.	4	Monday, Oct. 7
2. Project Plan Presentation	Revise your story plan as appropriate. Show all 4 points to the class to obtain peer feedback and make refinements as appropriate. Time limit: 5 minutes.	4	In class Oct. 23
3. Project Plan Revision Email	Email me with a revised list of the four points above: 1. The specific question you want to answer. 2. Source: The specific data will you use, with a hyperlink to the source or a copy of the data file. 3. Method: How you will clean and analyze the data. 4. Lede: The single, specific data point you hope to find, written as a lede sentence.	4	Monday, Oct. 28
4. Project Presentation	Present completed (not a draft) data story in class, so you can get peer feedback and make refinements before submitting for a grade. Show class your (1) question, (2) headline, and (3) lede. Also, (4) briefly describe the data and analysis. Time limit: 5 minutes.	8	In class Dec. 4
5. Final Project due	Upload to assignment in Canvas.	See above	Sunday, Dec. 8

PART 3: OTHER VERY 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.

Any work you submit must be solely and entirely your own. Any work drawn from another unattributed source, whether by plagiarism, Chat GPT, Gemini, Copilot, or any other AI program, is an academic integrity violation and will result in a failing grade for the course.

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.

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](#).