Inside Twitter: Data Analysis for Journalists

1/6/2019

Jour 405v, Sec. 7, Spring 2019

School of Journalism and Strategic Media

University of Arkansas

Class Time and Location: Monday-Wednesday, 9:40 a.m.-10:55 a.m., Kimpel 146

Instructor: Rob Wells, Ph.D.
Office: Kimpel 128
Office phone: 479-575-6305
Office hours: Mondays 11:30 a.m-1:30 p.m. or by appointment
E-mail: rswells@uark.edu

Course Goal: Students will learn the latest data journalism techniques that drive modern newsrooms and public relations / advertising offices. The class will extract and analyze Twitter data with the goal of producing an interactive multimedia presentation.

Course Description: This course will teach students how to code in R and how this powerful tool is used in modern news reporting. Quality reporting in newsrooms requires a solid foundation of data analysis. The data skills taught in this class are in high demand in newsrooms and corporations.

Learning Outcomes:

```
Proficiency in R and R Studio
Best practices in data project management
Create basic data visualizations for web publication
Manage GitHub into a data journalism workflow
Identify the limitations, strengths and weaknesses of datasets
Determine how data management by governments may impact the public
Understand how data analysis and reporting can advance journalism
Storytelling techniques that can help broad audiences understand data
```

Required Text:

Machlis, Sharon. Practical R for Mass Communications and Journalism. Chapman & Hall/CRC The R Series. 2018. ISBN 9781138726918 https://www.amazon.com/gp/search?keywords=9781138726918

http://www.machlis.com/R4Journalists/

Supplementary Texts:

Verification Handbook http://verificationhandbook.com/downloads/verification.handbook.pdf

Select readings posted course website

Cohen, Sarah Numbers in the Newsroom: Using Math and Statistics in News.Any edition. Columbia, Mo.: Investigative Reporters & Editors Inc., 2014. http://store.ire.org/collections/books/products/numbers-in-the-newsroom-using-math-and-statistics-in-news-second-edition

Meyer, Philip. The New Precision Journalism. Rowman & Littlefield, 2002. Free version: https://www.unc.edu/~pmeyer/book/Chapter1.htm.

Follow these websites:

ProPublica
The Upshot (The New York Times)
FiveThirtyEight
Vox
Reddit Rlanguage

Grading:

Assignments: 65 percent
Homework/Quizzes: 20 percent
Class Participation: 15 percent

Plagiarism or fabrication will result in your dismissal from class with an F for the course and a recommendation you be dismissed from the college.

Your work will be marked on the following scale:

A+:	100 - 98
A:	97 - 93
A-:	92 - 90
B+:	89 - 88
B:	87 - 83
B-:	82 - 80
C+:	79 - 78
C:	77 - 73
C-:	72 - 70
D+:	69 - 68
D:	67 - 63
D-:	62 - 60
F:	Below 60

A - The work is of professional quality (for journalism "professional" track students) or high academic quality (for others). It reflects a depth of research, clarity of writing, and a complete grasp of the main concepts presented in the class. B - The work is good but needs editing or is flawed in one of the categories mentioned above. C – The work is weak, needs major editing or reflects an average understanding of key concepts presented in class. D - Work fails to meet requirements and needs a complete rewrite. F – Unacceptable.

Libel:
Any story that includes libelous material will result in an F (55 percent) Examples would be if you describe someone as a murderer in your story before he or she is convicted, or if you mistype the name of a convicted murder and thereby implicate someone not guilty of the crime.

Attendance:
Class attendance figures in your class participation grade. An excused absence requires notification by e-mail before the start of class. Be prepared to submit documentation to validate your absence, especially if it is for an extended period of time. Students who miss more than six classes may have their final grade reduced by a full letter grade.

Class Communications:
Email: I will email individual students on occasion about important issues. I expect a timely reply, which is in the same day. It is your responsibility to check your email account.

I respond to email quickly, usually within an hour. I stop responding to student email at 9 p.m.

I use e-mail, the course website and Blackboard to communicate with students. Course website: I will post readings, announcements and grades on the class website. You will post some homework and project

materials on this site. Blackboard: Grades will be posted on Blackboard. You will post some assignments on Blackboard. It is your responsibility to check your email and Blackboard announcements.

Classroom Etiquette:
We will be working in a computer lab. Show respect for your colleagues and instructor by refraining from personal computer use during class. You are being rude to your instructor and distracting to your classmates when you engage in computer activities unrelated to class. Anyone misusing classroom computers for personal matters will receive a zero for class participation that day and may be asked to leave the class if the behavior persists.

Academic Honesty:
Please refer to http://provost.uark.edu/245.php for the academic integrity policy.

Class Weather Policy:
If the university is closed, there will be no class. If I need to cancel class, for whatever reason, I will do my best to notify you by e-mail and notify the journalism office: 479-575-3601.

CEA, Center for Education Access:
If you are a student with special needs, contact me personally **and as soon as possible**. The CEA is at 479-575-3104. I will accommodate students who require assistance.

Emergency Preparedness Plan:
The university has a new emergency plan; review it at http://emergency.uark.edu/

About the Instructor:
Rob Wells is an assistant professor of journalism and has been teaching at the University of Arkansas since the Fall 2016 semester. He earned his doctorate in philosophy in Journalism Studies at the University of Maryland's Philip Merrill College of Journalism. His academic research is in business journalism and history, along with data journalism and technology. He is the author of a forthcoming book about the problems and future of business journalism, to be published by the University of Illinois Press in 2019. As an adjunct instructor, he taught reporting classes at the Merrill College between 2010-2016. He was a 2012 Reynolds Visiting Professor at the University of South Carolina, Columbia, a program sponsored by the Donald W. Reynolds National Center for Business Journalism.

Wells is the former deputy bureau chief for Dow Jones Newswires/Wall Street Journal in Washington, D.C., where he oversaw 22 reporters who covered real-time business, economics and financial news in the nation's capital. Prior to this, he was a business reporter for Dow Jones, Bloomberg News and The Associated Press. He holds a master's degree in liberal studies from St. John's College in Annapolis, where he studied philosophy, literature, history and political science.

Schedule of Instruction
Please see the course outline: https://profrobwells.github.io/Data-Analysis-Class-Jour-405v-5003/Spring_2019_Adv_Reporting_Jan_4_2019.html

Assignments: An assignment uploaded late (after 11:59 p.m. on the designated day, according to Blackboard) will be reduced by one grade, and will be reduced a full grade for every subsequent day. Students with excused absences should contact me immediately about making up missed assignments. The final assignment represents the final examination; there is no separate final examination.

1: Static Graphic - Managing Data in R. Due Feb 6

Students will use R Studio to gather, analyze and visualize FBI Uniform Crime data for Arkansas. Import and clean data from the web Convert string data to numeric data Perform percentage change calculation in new column Visualize findings using ggplot

2. Visualization of Twitter data. Due Feb 27.

Students will produce publication-ready graphics from Twitter data. Results will be posted on GitHub. Data dictionary required

3: Interactive Map. Due March 27

Students will use R Studio to build interactive maps of data in Arkansas. Data dictionary required

4. Final Project: Interactive Data Visualization. Due April 29.

Students will use R Studio to extract a corpus of Twitter data, manage that data in R and produce a basic interactive visualization using a sentiment analysis.