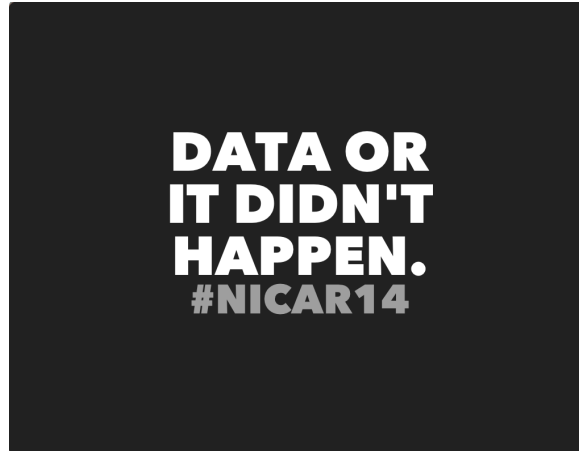


**JOMC 491.6: Data Driven Journalism**  
Spring 2014 Syllabus

Associate Professor Ryan Thornburg  
215 Carroll  
919-537-3681



**Welcome**

The economics of digital publishing technology have opened a floodgate of raw data, and journalism is just one of many fields that are being rapidly transformed by that floodgate. In law, medicine, business, politics, and ecology, people are using data to understand the social and natural worlds. They are also using data to tell stories.

To retain their position as brokers of trust and hubs of community conversation, reporters must understand how the people they cover are using data, and how journalists can use data to improve our shared understand of an increasingly complex world.

Your decision to take this course indicates that you are interested in learning the skills and concepts of data-driven reporting. My expectation is that you already have demonstrated clear news judgment and precise, brief storytelling either in or out of a classroom.

The class starts from the assumption that you've never or rarely used even a basic spreadsheet to aid either your reporting or storytelling. That's where the semester will begin. We will end just shy of an introduction to how computer programming and algorithms are using journalistic data to create new editorial products.

**The Goal of This Course**

Students who successfully complete this course will be able to acquire, organize, analyze and present data to a general news audience.

## What You Will Learn

The first few weeks of class will be dedicated to an introduction to basic statistics and numerical and mathematical literacy, as well as a look at professional data-driven journalism projects.

The bulk of the course will be spent on practical skills exercises using tools such as Excel, Access, Fusion Tables, Open Refine, Tableau and QGIS.

The culmination of the course will be a an explanatory or accountability news story and prototype of a working news data visualization or application.

## Required Readings

The best way to learn about the changing journalism environment is to keep a close eye on professionals working in the industry. The bulk of our reading will be contemporary articles and research about data driven reporting.

Books and articles you will need for this class:

- *IRE Tipsheets* available to IRE Student Members. Register at <https://www.ire.org/membersonly/join/register> for \$25.
- *Numbers in the Newsroom*. Sarah Cohen.  
<http://store.ire.org/collections/books/products/numbers-in-the-newsroom-using-math-and-statistics-in-news> (\$15 for IRE members)
- *Art of Access: Strategies for Acquiring Public Records*. David Cuillier and Charles N. Davis. Chapter 7: "Going Digital", including footnotes. (Chapter 7 is available via digital reserve. You may buy the complete book at Student Stores or online.)
- Other contemporary readings and tutorials available on Sakai.

## Recommended Readings

- *Precision Journalism*, Phil Meyer. The book that birthed computer assisted reporting, written by a former UNC journalism professor.
- <http://www.codecademy.com> - Good, free tutorials to help make your first step from prototype to professional product.
- *Mapping the News: Case Studies in GIS and Journalism*, David Herzog.
- Additional tutorials and links on Sakai

## Grades

### Attendance – 25%

Classes will be dominated by hands-on work and student presentations. The practice you enjoy during class will set you up for success on the homework assignments.

You may miss one class for any reason, as long as you meet with me and make up your work within one week. After that, each missed class will result in one full letter reduction of your attendance grade. You will also not be able to make up any work missed in subsequent classes.

### Midterm Project – 25%

Due: March 5.

You have three options for your midterm project. (Let me know by Feb. 5):

- *Assessing the state of open government data in N.C.* – You will devise a rubric for grading city and county government's on their adherence to the 8 Principles of Open Government Data. You will interview officials and then complete the report card for one city or county government, or one statewide cabinet-level agency. Finally, you will provide a brief analysis of barriers to increased government transparency, risks and opportunities, and a recommendation for steps that you think government's should take over the next three years to address those risks and opportunities.
- *Assessing the state of data driven journalism in N.C.* – You will describe the use of data-driven reporting and presentation at five N.C. newspapers or television stations. Based on interviews with reporters and editors at those news organizations, you will determine and describe their obstacles to data-driven reporting and opportunities. Finally, you will describe a strategy on how each newsroom should deal with those obstacles and opportunities over the next three years.
- *Scrape five public data sites* – You will create scrapers that periodically pull data from N.C. government websites. You will make public the code you use to create the scrapers and you will make the data (feed) publicly available.

### Homework Assignments – 25%

You will have 10 homework assignments throughout the semester are designed to extend the practical classroom instruction. They will challenge you to use data to find precise answers as well as develop your own relevant questions.

You should schedule on your personal calendar right now at least six hours each week to assignments you will do outside of class. Some weeks it will be more. Some less. But if you cannot spend at least 90 hours on this class during this semester you

may want to consider another course. For all homework assignments you must show your work.

I will accept one late assignment from each student for up to 24 hours after the deadline. Further late assignments will be given an F.

You have up to one week to make up a one missed in-class quiz or in-class presentations. Additional missed quizzes or presentations will result in an F grade for that assignment.

Final Project: Story and data viz/app – 25%

Due: May 6, 3 p.m. (Topics assigned March 17)

Individually or with a team of other students you will produce a data-driven journalism project. The topic will likely be N.C. campaigns & elections, but I will accept other proposed topics. The final project should have a defined audience and will have the following elements:

- an interactive map
- an interactive chart
- search and sort interfaces
- a written “So What?” story that makes clear the significance and context of the data. The story will have a lead, news value(s) and quotes.

You may use any tools you wish to create the elements of the final project.

## Class Topics Calendar

Date	Class Topic	HW Assignment Due
Jan. 8	Introduction; tech set-up	
Jan. 13	Examples of Data Driven Journalism	
Jan. 15	Introduction to Excel	
Jan. 20	Data Acquisition Strategies	Describing campaign finance trends with Excel
Jan. 22	Cleaning Data with Excel	
Jan. 27	Cleaing Data with Refine	Public data request & Open-NC.org entries
Jan. 29	<b>Presentations: Numbers in the Newsroom</b>	Presentations: Numbers in the Newsroom
Feb. 3	Web scraping with Google Docs	"cleaned" campaign finance reports
Feb. 5	Web scraping with OutWit	
Feb. 10	Extracting data from PDFs	OutWit or GDoc scrape of public data
Feb. 12	"How They Got the Story"	
Feb. 17	Introduction to databases with MS Access	
Feb. 19	Querying data with MS Access; Intro to SQL	
Feb. 24	Enterprise Joins to Find Stories	Turn a paper police report into an Access database
Feb. 26	Data analysis with Excel Pivot Tables	
March 3	Intro to spatial data with QGIS	
March 5	Census data and geocoding	<b>Midterm Papers</b>
<b>Spring Break</b>		
March 17	Fundamentals of data visualization	
March 19	Intro to Tableau	
March 24	Analyzing data with Tableau	Campaign finance map
March 26	Presenting data with Tableau	
March 31	Google Maps	Tableau of campaign finance
April 2	Google Maps	
April 7	Creating a searchable data website with FreeDive	Google Map example
April 9	FreeDive	
April 14	Work on Final Project	FreeDive search of campaign finance reports
April 16	Work on Final Project	
April 21	Work on Final Project	
April 23	Work on Final Project	
<b>Final Project Due: 3 p.m., May 6.</b>		

## What grades mean

**A** - Mastery of course content at the highest level of attainment that can reasonably be expected of students at a given stage of development. The A grade states clearly that the student has shown such outstanding promise in the aspect of the discipline under study that he or she may be strongly encouraged to continue.

**B** - Strong performance demonstrating a high level of attainment for a student at a given stage of development. The B grade states that the student has shown solid promise in the aspect of the discipline under study.

**C** - A totally acceptable performance demonstrating an adequate level of attainment for a student at a given stage of development. The C grade states that, while not yet showing unusual promise, the student may continue to study in the discipline with reasonable hope of intellectual development.

**D** - A marginal performance in the required exercises demonstrating a minimal passing level of attainment for a student at a given stage of development. The D grade states that the student has given no evidence of prospective growth in the discipline.

**F** - For whatever reason, an unacceptable performance. The F grade indicates that the student's performance in the required exercises has revealed almost no understanding of the course content. A grade of F should warrant an advisor's questioning whether the student may suitably register for further study in the discipline before remedial work is undertaken.

## Student Honor

The University of North Carolina at Chapel Hill has had a student-led honor system for over 100 years. Academic integrity is at the heart of Carolina and we all are responsible for upholding the ideals of honor and integrity. The student-led Honor System is responsible for adjudicating any suspected violations of the Honor Code and all suspected instances of academic dishonesty will be reported to the honor system.

All academic work in this course, including homework, quizzes, and exams, is to be your own work, unless otherwise specifically provided. It is your responsibility if you have any doubt to confirm whether or not collaboration is permitted. If the work is truly your own, you will be able to explain and demonstrate to my satisfaction how you did it.

Do not represent someone else's words, thoughts, or ideas as your own without attribution in connection with submission of academic work, whether graded or otherwise.

Further information about the student Honor Code is available at <http://studentconduct.unc.edu/honor-system>