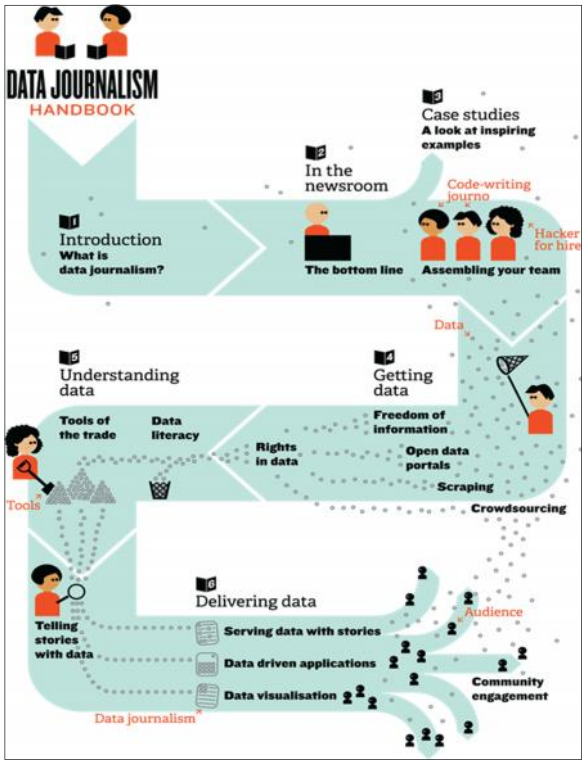
USC Annenberg

JOUR 322: Data Journalism
2 Units

Spring 2017:	Wednesdays, 2-3:40 p.m.
Section:	21168D
Location:	ANN L116
Instructor:	**Dana Chinn, Lecturer**
Office:	ASC 227
Office Hours:	By appointment on Tuesdays, Wednesdays and other days
Email:	chinn@usc.edu

Course Description and Learning Outcomes

Proficiency with gathering, analyzing and visualizing data is essential in journalism today as commodity content becomes increasingly ineffective in both serving the public interest and engaging audiences.



Infographic by Lulu Pinney
from "The Data Journalism Handbook"

This two-unit introductory course is an overview of how to go beyond anecdotal reporting with:

- **Basic quantitative analysis and data visualizations** to make comparisons, put data into context, and examine trends, and

- **Data science concepts** such as preparing databases, and analyzing or "interviewing" them as you would any other source.

You'll use some of the many data journalism tools, especially Excel. However, this foundational course is about data, not software and coding, because "…despite advances in technology, the fundamentals of basic reporting still apply. These include a devotion to accuracy, clarity, fairness, solid news value and, ultimately, good storytelling." *(Jennifer LaFleur, Center for Investigative Reporting)*

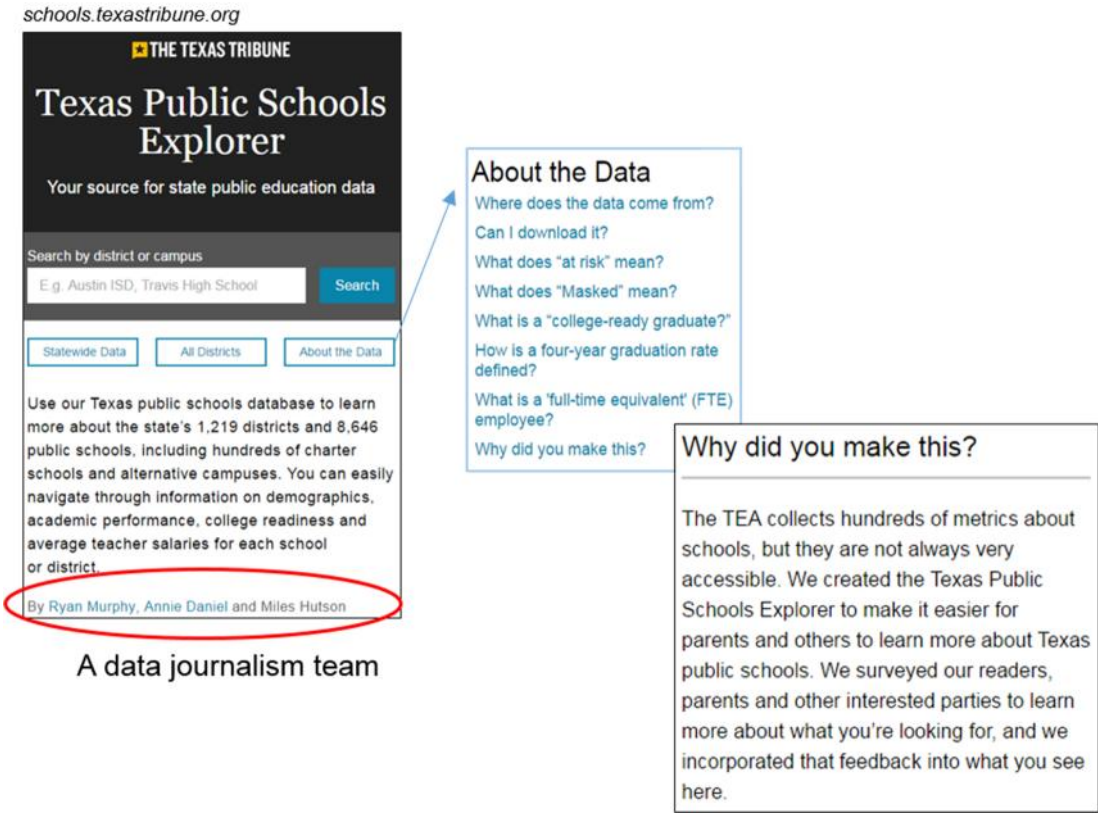
Course Outcomes
There's an old saying that "there's no data without a story, and no story without data." After successfully completing this course, you will have:

1. **Demonstrated your understanding of basic quantitative analysis and data visualization concepts** such as:
   - Percentages, rates, averages vs. medians, indexes and rankings
   - Line, bar and pie charts; tables; maps

2. **Gathered and structured at least one raw dataset using data integrity and transparency standards followed by professional journalists (midterm project).** You will have practiced extensively with Excel formulas and functions, and will understand what's required to use other tools such as SQL and QGIS.

3. **Produced a data journalism story package (final project)**
   Your final project will build from the dataset you prepared for your midterm. It will include a story or infographic with charts, maps and/or illustrations that incorporates interviews and other reporting; a searchable database that helps audiences engage with your story; and an "about the data" page that explains your methodology.

4. **Learned whether you want to be a data journalist,** and if so, what topics and skills you need to pursue with coursework and internships.



A data journalism team

Learning Objectives and Assessment

- **Define components in news stories that need to be quantified**, and label good, bad and "so what" uses of data in stories. Describe what basic quantitative analysis and data visualization concepts should be used.

- **Demonstrate knowledge of basic quantitative analysis and data visualization concepts** in Excel with small datasets.

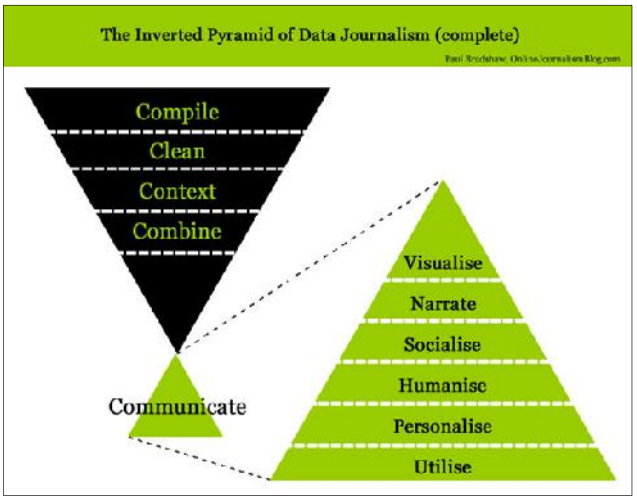
- **Illustrate your knowledge of the components of a data journalism story by writing story pitches** that include:
  - Datasets as sources
  - Plans for an interactive database

- **Examine different types of datasets and how they're structured.**

  - **Government data** through open data portals, interviews, and/or Freedom of Information Act (federal) and California Public Records Act (state and local) requests

  - **Sources** such as the Investigative Reporters and Editors, Inc./National Institute for Computer-Assisted Reporting Data Library

  - **Datasets you compile** from hard copy documents, web scraping, observations and secondary sources.



- **Construct a simple relational database schema or table structure, and prepare the datasets for analysis and visualization tools.**
  - Document why the data in a dataset was collected and how it's currently used; identify what's missing or needed in a codebook or record layout.
  - Extract, clean, and code and otherwise prepare data to be analyzed in Excel, MySQL, Google Fusion Tools, QGIS and other tools.

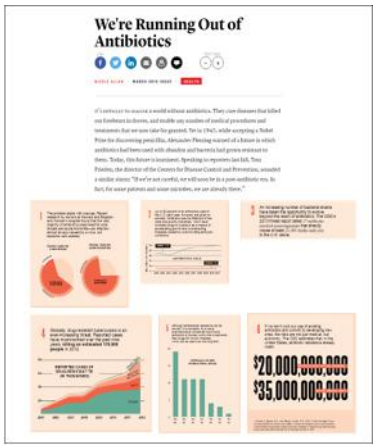
- **Analyze datasets** with basic descriptive statistics. Demonstrate knowledge of what type of chart, table or map explains the findings or patterns effectively. Identify any additional coding or datasets that are needed. Interview sources based on your analyses.

- **Create a data journalism story package** that synthesizes your data analysis, interviews and other reporting.

Course Grade Components and Breakdown

	Percent Of Grade
1. Assignments (10)	**40%**
2. Midterm project: Dataset analysis	**20%**

- **Database schema** or record layout
- **Raw dataset**
- **Restructured and coded dataset**
- **Descriptive statistics calculations**
- **Infographic with a summary and two charts**, each with an explanatory paragraph



"We're Running Out of Antibiotics," by Nicole Allan
The Atlantic Chartist, March 2014

3. Final project: Data journalism story package **30%**

The final project builds on the dataset you prepared for your midterm project.

- **Story or infographic with at least three data visualizations** (chart, map, illustration and/or searchable database wireframe); interviews with at least two approved sources.
- **An "about this story" write-up** that includes definitions of terms, the assumptions you made when you cleaned and coded the data, and information about your interviews and other research that's essential to make your methodology transparent to readers.
- **Revised dataset analysis (midterm project)**: database schema and record layout; both the raw and the cleaned-up, coded datasets; Excel workbook with descriptive statistics.

4. Attendance and class participation	**10%**
Total:	**100%**

Attendance and class participation
Part of your course grade will be based on your coming to class on time, being prepared, and actively participating during the lectures and in-class workshops.

Grading Scale
Each assignment, quiz and story will be worth 100 points.

A	95-100	B+	87-89	C+	77-79	D+	67-69
A-	90-94	B	83-86	C	73-76	D	63-66
		B-	80-82	C-	70-72	D-	60-62
						F	59 and below

Grading Standards
There will be grading rubrics for each assignment and project based on the following:

A
The story uses the correct data, is concise and complete, and suitable for publication. Calculations are correct and appropriately formatted (e.g., rounded). Sources (data, people, secondary sources) are correctly presented. Any data integrity issues and assumptions are identified.

The story is clearly written, is free of spelling and grammatical errors, and adheres to AP Style. It includes relevant data visualizations with explanatory text. There are no 3D and other types of graphs in formats that obscure the trends or data points. Charts, graphs and/or maps use the correct colors, data labels and font sizes, and have unobtrusive backgrounds.

B
The story has most of the attributes of an "A" story but is missing a significant element.

C
The story uses the correct data and calculations, but is missing one or more significant elements and is poorly written. It may have some unsubstantiated statements such as "Many people think…" or "Most of the group…"

D
The story uses only some of the correct data or calculations and is missing one or more significant elements. It would not be published.

F
The story has a material factual error or doesn't use the correct data or calculations. It may have other major flaws such as stating hypotheses as facts.

Late assignment policy
I will accept late assignments but will deduct 10 points or a full letter grade.

Required Tools and Readings

1. **Laptop (PC or Mac) with as much RAM and hard disk space as possible.** Please bring your laptop to every class. Being on non-class websites during class will affect the class participation component of your course grade.

2. **Software**
   a. **Microsoft Office 2016** (Word, Excel, PowerPoint) is available for free to USC students at itservices.usc.edu/officestudents.

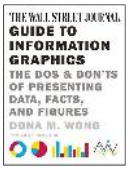
   You will be using Excel to learn the building blocks of quantitative analysis and database design. This is a two-unit introductory course, so the emphasis is on learning fundamental concepts rather than struggling with software.

   Excel can be difficult for journalists to learn, but it's an essential tool for all aspects of data journalism. For data analysts, Excel is a sketchbook where you can quickly upload a dataset and experiment with different calculations. It's also often the easiest and most efficient way to prepare datasets before loading them into mapping, database and statistical analysis programs.

   We'll be reviewing programs such as MySQL and QGIS, and you'll also see examples from our guest speakers.

   b. **Charts, tables, infographics and maps**
   Your assignments and projects will be graded based on the principles in *The Wall Street Journal Guide to Information Graphics,* listed below. I'll be showing you how to do charts and tables in Excel and PowerPoint, and maps in Google Fusion Tables. However, for your assignments and projects, you are welcome to use any program or combination of programs that you'd like.

3. ***Numbers in the Newsroom: Using Math and Statistics in News,*** *Second Edition, E-version,* by Sarah Cohen for Investigative Reporters and Editors, Inc., 2014, 134 pages. Available via the IRE website store at store.ire.org.



4. ***The Wall Street Journal Guide to Information Graphics: The Dos & Don'ts of Presenting Data, Facts, and Figures****,* by Dona M. Wong, W.W. Norton, 2010, 158 pages.

5. ***The Data Journalism Handbook: How Journalists Can Use Data to Improve the News***, edited by Jonathan Gray, Liliana Bounegru and Lucy Chambers, O'Reilly Media, 2012, 220 pages. A digital version is free at datajournalismhandbook.org/1.0/en/.

Suggested Resources

- **AP Stylebook:** Hard copy, online or any other source, e.g., tipsheet

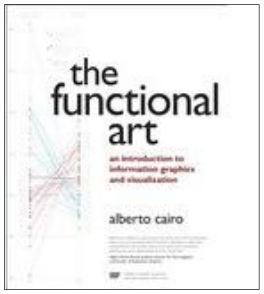
- **Become a member of Investigative Reporters & Editors, Inc.** The IRE website (ire.org) gives members access to the National Institute for Computer-Assisted Reporting library of databases and many other data journalism resources. Student membership is $25 a year. It's $10 a semester per student if our entire class signs up, but I recommend you sign up for a year so you can access the materials after the course ends.

    - **IRE Resource Center** - ire.org/resource-center - Tipsheets, story suggestions, stories and other resources.

    - **NICAR-Learn** - learn.ire.org - Short videos on specific techniques and tips. Some are free; others are available with an annual paid subscription of $25/year for IRE members; $40/year, nonmembers.

- ***The Functional Art: An Introduction to Information Graphics and Visualization***, by Alberto Cairo, New Riders/Peachpit/Pearson Education, 2013, 364 pages with DVD introductory information graphics video course.

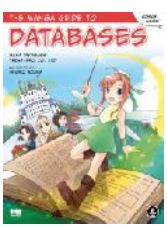
  Website: thefunctionalart.com
  Twitter: @albertocairo



- The work of **Edward Tufte**, "arguably the most influential theoretician in information design and visualization," according to Alberto Cairo.

  Website: edwardtufte.com
  Twitter: @EdwardTufte

- ***The Manga Guide to Databases, Third Edition,*** by Mana Takahashi, Shoko Azuma and Trend-Pro Co., Ltd, No Starch Press, 2012, 224 pages



A fun way to learn about database design and data integrity!

Course Schedule

Guest speakers will be added to the schedule throughout the semester.

Codes to required readings
NN: Numbers in the Newsroom
WSJG: WSJ Guide to Information Graphics
DJH: The Data Journalism Handbook

Week	Topics	Readings	Homework due the following week
1	**Course overview** **What is data journalism?**	DJH Chapters 1, 2 & 3	Skills and interests survey on Blackboard If needed: • Update to Office 2016 • Excel tutorials on Lynda.com
2	**What is a data journalism story?** **Quantitative analysis with Excel 1** • What's a number? 7 tips • The Pers: Fractions, rates, percents	NN Chapters 1 & 2 NN Fractions and Percentages Conversion Chart, p. 132	Assgn. 1
3	**Quantitative analysis with Excel 2** • Averages, medians, ranges • Rankings • Per capita • Excel filters for sorting	NN Chapter 2 WSJG Chapter 3	Assgn. 2
4	**From raw dataset to story and visualizations** • Line vs. bar vs. pie charts	WSJG Chapters 1 & 2 DSJ pgs. 165-176, 186-190	Assgn. 3
5	**Quantitative analysis with Excel 3** • Pivot tables **Intro to dataset sources**	DJH Chapters 3, 5 & 6 NN Chapter 6	Assgn. 4

Week	Topics	Readings	Homework due the following week
6	**Dataset sources** • Government open data portals • Public records requests • Datasets you compile: web scraping; converting PDFs; logging observations. Primary sources on DocumentCloud.	DJH Chapter 4	Assgn. 5
7	**Preparing datasets for analysis and accurate reporting** • Database schemas or table structures; labels • Data types, codebooks or record layouts	Handout to be determined	Assgn. 6
8	**Midterm project workshop – dataset preparation** • Revise schema and record layouts • Clean data. • Compile code sheet; label rows		Draft midterm.
9	**Midterm project workshop – infographic with at least two charts** • Review cleaned and coded data and descriptive statistics. • Draft infographic.		Complete midterm project – due at the beginning of class in Week 10
10	**Maps and Google Fusion Tables** • When to use a map • Types of maps; examples of map abuse • Overview of mapping programs • Database schemas and record layouts – Google Fusion Tables vs. QGIS and others	WSJG pgs. 90-91 DJH Chapter 6	Assgn. 7
11	**Final project overview**	To be determined	Assgn. 8 (final project draft component 1)

Week	Topics	Readings	Homework due the following week
12	**Searchable databases**	To be determined	Assgn. 9
13	**Topic to be determined** (based on final story pitches)	To be determined	Assgn. 10
14	**Topic to be determined** (based on final story pitches)		
15	**Final project workshop**		

The final project is due via Blackboard on Tuesday, May 9, 2016, 10 a.m.

Policies and Procedures

Internships
The value of professional internships as part of the overall educational experience of our students has long been recognized by the School of Journalism. Accordingly, while internships are not required for successful completion of this course, any student enrolled in this course that undertakes and completes an approved, non-paid internship during this semester shall earn academic extra credit herein of one percent of the total available semester points for this course. To receive instructor approval, a student must request an internship letter from the Annenberg Career Development Office and bring it to the instructor to sign by the end of the third week of classes. The student must submit the signed letter to the media organization, along with the evaluation form provided by the Career Development Office. The form should be filled out by the intern supervisor and returned to the instructor at the end of the semester. No credit will be given if an evaluation form is not turned into the instructor by the last day of class. Note: The internship must by unpaid and can only be applied to one journalism class.

Statement on Academic Conduct and Support Systems
Academic Conduct
Plagiarism
Presenting someone else's ideas as your own, either verbatim or recast in your own words - is a serious academic offense with serious consequences. Please familiarize yourself with the discussion of plagiarism in *SCampus* in Section 11, *Behavior Violating University Standards* https://scampus.usc.edu/b/11-00-behavior-violating-university-standards-and-appropriate-sanctions/. Other forms of academic dishonesty are equally unacceptable. See additional information in *SCampus* and university policies on scientific misconduct, http://policy.usc.edu/scientific-misconduct/.

USC School of Journalism Policy on Academic Integrity
The following is the USC Annenberg School of Journalism's policy on academic integrity and repeated in the syllabus for every course in the school:

"Since its founding, the USC School of Journalism has maintained a commitment to the highest standards of ethical conduct and academic excellence. Any student found plagiarizing, fabricating, cheating on examinations, and/or purchasing papers or other assignments faces sanctions ranging from an 'F' on the assignment to dismissal from the School of Journalism. All academic integrity violations will be reported to the office of Student Judicial Affairs & Community Standards (SJACS), as per university policy, as well as journalism school administrators."

In addition, it is assumed that the work you submit for this course is work you have produced entirely by yourself, and has not been previously produced by you for submission in another course or Learning Lab, without approval of the instructor.

Support Systems
Equity and Diversity
Discrimination, sexual assault, and harassment are not tolerated by the university. You are encouraged to report any incidents to the *Office of Equity and Diversity* http://equity.usc.edu/ or to the *Department of Public Safety* http://dps.usc.edu/contact/report/. This is important for the safety of the whole USC community. Another member of the university community - such as a friend, classmate, advisor, or faculty member - can help initiate the report, or can initiate the report on behalf of another person. *The Center for Women and Men* http://www.usc.edu/student-affairs/cwm/ provides 24/7 confidential support, and the sexual assault resource center webpage https://sarc.usc.edu/ describes reporting options and other resources.

Support with Scholarly Writing
A number of USC's schools provide support for students who need help with scholarly writing. Check with your advisor or program staff to find out more. Students whose primary language is not English should check with the *American Language Institute* http://ali.usc.edu/ which sponsors courses and workshops specifically for international graduate students.

The Office of Disability Services and Programs
http://sait.usc.edu/academicsupport/centerprograms/dsp/home_index.html provides certification for students with disabilities and helps arrange the relevant accommodations.

Students requesting test-related accommodations will need to share and discuss their DSP recommended accommodation letter/s with their faculty and/or appropriate departmental contact person at least three weeks before the date the accommodations will be needed. Additional time may be needed for final exams. Reasonable exceptions will be considered during the first three weeks of the semester as well as for temporary injuries and for students recently diagnosed. Please note that a reasonable period of time is still required for DSP to review documentation and to make a determination whether a requested accommodation will be appropriate.

Stress Management
Students are under a lot of pressure. If you start to feel overwhelmed, it is important that you reach out for help. A good place to start is the USC Student Counseling Services office at 213-740-7711. The service is confidential, and there is no charge.

Emergency Information
If an officially declared emergency makes travel to campus infeasible, *USC Emergency Information* http://emergency.usc.edu/ will provide safety and other updates, including ways in which instruction will be continued by means of Blackboard, teleconferencing, and other technology.

About Your Instructor
I've been a full-time faculty member at Annenberg since 2002, and have focused on digital analytics and open data since 2007. This semester, in addition to teaching JOUR 322/Data Journalism, I'm teaching JOUR 477/Web Analytics for News and Nonprofit Organizations.

I've run many data journalism initiatives at Annenberg over the years, including mentoring Annenberg's first Google Data Journalism Fellows in 2014 and 2015. I'm thrilled to be teaching Annenberg's first required data journalism classes as I work with news organizations on audience engagement strategies and defining the role of journalism in open data.

I have an undergraduate degree in journalism and an MBA, both from USC. I continue to work with various nonprofit news organizations and e-commerce companies as a consultant, and previously worked at Gannett, the Los Angeles Times and other media organizations.