**JOUR 554: Reporting with Data**

**2 units**

**Fall 2021 – Tuesday – 6:30-8:30pm**

**Section:** 21690D

**Location:** ASC230

**Instructor: Matt Stiles**

**Office Hours:** 1 hour before and after class

**Contact Info:** [mattstiles@gmail.com](mailto:mattstiles@gmail.com) | 310.529.8749

**I. Course Description**

The ability to acquire, process, analyze and visualize data is a critical skill for today’s journalists. This hands-on course will provide an in-depth understanding of modern quantitative analysis tools and techniques essential to producing data-driven projects and adding context to stories.

**II. Overall Learning Objectives and Assessment**

The course will be taught into these four subject areas:

* Data analysis using SQL, Python, Pandas and Jupyter
* Mapping analysis using QGIS and GeoPandas
* Data visualization using Python, Pandas and Altair
* Statistical analysis using Python and Pandas

The course will give students a deep introduction to these tools, which can help them probe, understand and derive findings from most data sets. It is designed to elevate the students’ reporting capability and help them craft stronger stories and discover those they might otherwise have failed to see.

To start, it’s critical for students to understand how the basics of how to query a database with a simple and ubiquitous programming language like SQL, or Structured Query Language. SQL is lingua franca of programming languages and is straightforward and logical after some practice.

Using SQL, students will learn how to “interview” their data by asking all the relevant questions with code. This will include common tasks, such as aggregating, filtering and joining data. These skills are key to exploring a data set to find stories, but also to reveal problems and limitations in the data.

The students will them learn to perform similar tasks using Python, an object-oriented programming language widely used in data science and data journalism. They will perform data acquisition, cleaning and analysis with the help of the Pandas library and Jupyter notebook, a powerful tool for maintaining reusable and easy-to-debug records of your work.

Students will also learn how to think about data with geography in mind. Using mapping software, students can import geospatial data and merge it with tabular data to display information on a map. Functions like spatial joins, buffering and density analysis will allow students to increase the power of their reporting.

Finally, students will learn how to visualize data and conduct basic statistical analysis again using Python, Pandas and Jupyter, with an assist from the Altair plotting library. The course will also cover a range of introductory statistical concepts, including standard deviation, regression and correlation.

**III. Course Notes**

The course will use Blackboard for all assignment and project submissions. Copies of lecture slides, data sets and other class information will be also posted on Blackboard. E-mail should only be used by students for questions or clarifications.

**IV. Description and Assessment of Assignments**

There will be a series of practical assignments working with data that align with the different subject areas – data analysis, data visualization, statistical analysis and mapping analysis. Students will be working with data sets they find to produce two data-driven stories that showcase the skills acquired throughout the semester. Story memos must be submitted each week for your two data stories – this will keep students and the instructor aligned throughout the semester. For all assignments, in-class instruction should prepare students to complete the assignment.

Students may also use the data, reporting and graphics from the either of the data-driven stories for their Capstone project.

**V. Grading**

**a. Breakdown of Grades**

| Assignment | % of Grade |
| --- | --- |
| Weekly Assignments | 40% |
| Two data stories, methodologies and graphics | 50% |
| Participation and story memos | 10% |
| **TOTAL** | **100%** |

**b. Grading Scale**

|  |  |  |
| --- | --- | --- |
| 95% to 100%: A | 80% to 83%: B- | 67% to 69%: D+ |
| 90% to 94%: A- | 77% to 79%: C+ | 64% to 66%: D |
| 87% to 89%: B+ | 74% to 76%: C | 60% to 63%: D- |
| 84% to 86%: B | 70% to 73%: C- | 0% to 59%: F |

**c. Grading Standards**

All assignments will be edited on a professional basis and you will be judged first on the accuracy, fairness and objectivity of your stories. You will then be evaluated for broadcast style, editing, production value, originality and the ability to meet deadlines.

**“A” stories** are accurate, clear, comprehensive stories that are well written and require only minor copyediting (i.e., they would be aired or published). Video work must also be shot and edited creatively, be well paced and include good sound bites and natural sound that add flavor, color or emotion to the story.

**“B” stories** require more than minor editing and have a few style or spelling errors or one significant error of omission. For video, there may be minor flaws in the composition of some shots or in the editing. Good use of available sound bites is required.

**“C” stories** need considerable editing or rewriting and/or have many spelling, style or omission errors. Camera work and editing techniques in video stories are mediocre or unimaginative, but passable. Sound bites add little or no color - only information that could be better told in the reporter’s narration.

**“D” stories** require excessive rewriting, have numerous errors and should not have been submitted. Camera work is unsatisfactory or fails to show important elements.

**“F” stories** have failed to meet the major criteria of the assignment, are late, have numerous errors or both. Your copy should not contain any errors in spelling, style, grammar and facts. Any misspelled or mispronounced proper noun will result in an automatic “F” on that assignment. Any factual error will also result in an automatic “F” on the assignment. Accuracy is the first law of journalism. The following are some other circumstances that would warrant a grade of “F” and potential USC/Annenberg disciplinary action:

• Fabricating a story or making up quotes or information.

• Plagiarizing a script/article, part of a script/article or information from any source.

• Staging video or telling interview subjects what to say.

• Using video shot by someone else and presenting it as original work.

• Shooting video in one location and presenting it as another location.

• Using the camcorder to intentionally intimidate, provoke or incite a person or a group of people to elicit more “dramatic” video.

• Promising, paying or giving someone something in exchange for doing an interview either on or off camera.

• Missing a deadline.

**VI. Assignment Rubrics**

These will differ by assignment and will be provided in class.

**VII. Assignment Submission Policy**

A. All assignments are due on the dates specified. Lacking prior discussion and agreement with the instructor, late assignments will automatically be given a grade of F.

B. Assignments must be submitted via Blackboard.

**VIII. Required Readings and Supplementary Materials**

*Computer-Assisted Reporting: A Practical Guide (4th Edition),* by Brant Houston, Routledge, 2014, 264 pages.

*The Curious Journalist's Guide to Data,* by Jonathan Stray, GitBook, 2016, 122 pages. *Available online for free at* [*https://legacy.gitbook.com/book/towcenter/curious-journalist-s-guide-to-data/details*](https://legacy.gitbook.com/book/towcenter/curious-journalist-s-guide-to-data/details)

*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, July 2012, 220 pages. Available online for free at <http://datajournalismhandbook.org/1.0/en/>.

Other online readings will be added throughout the course.

**Required software:** *(all free online and we will install in class)*

* Google Sheets for spreadsheet work and simple graphics
* SQLite and DB Browser for SQLite for databases
* QGIS for mapping analysis
* Python and Jupyter for data analysis, visualization and statistics (installed via Anaconda)
* GitHub for version control and showing your work

**IX. Laptop Policy**

All undergraduate and graduate Annenberg majors and minors are required to have a PC or Apple laptop that can be used in Annenberg classes. Please refer to the [**Annenberg Digital Lounge**](http://www.annenbergdl.org/) for more information. To connect to USC’s Secure Wireless network, please visit USC’s [Information Technology Services](http://itservices.usc.edu/wireless/support/) website.

**X. Add/Drop Dates for Session 001 (15 weeks: TBD)**

* **Friday, TBD:** Last day to register and add classes for Session 001
* **Friday, TBD:** Last day to drop a class without a mark of “W,” except for Monday-only classes, and receive a refund for Session 001
* **Tuesday, TBD**: Last day to drop a Monday-only class without a mark of “W” and receive a refund for Session 001
* **Friday, TBD:** Last day to drop a course without a mark of “W” on the transcript for Session 001. [Please drop any course by the end of week three (or the 20 percent mark of the session) to avoid tuition charges.]
* **Friday, TBD:** Last day to change pass/no pass to letter grade for Session 001. [All major and minor courses must be taken for a letter grade.]
* **Friday, TBD:** Last day to drop a class with a mark of “W” for Session 001

**XI. Course Schedule: A Weekly Breakdown**

***Important note to students:*** *Be advised that this syllabus is subject to change - and probably will change - based on the progress of the class, news events, and/or guest speaker availability.*

|  |  |  |  |
| --- | --- | --- | --- |
| Date | Topics/Daily Activities | Readings and Homework | Final Project Deliverable/Due Dates |
| Week 1 | **Course Overview:**  Class goals. What’s possible with “hacker” journalism?  Brainstorm data-driven story ideas in class.  Software installation in class. | **Reading:** **Readings*:*** *Why Teach SQL?* <http://thescoop.org/archives/2011/07/27/why-teach-sql/>  Sign up for a free a Github account. |  |
| Week 2 | **Data analysis Part 1:**  Introduction to database managers, schema, record layouts, primary keys, etc.  Overview of SQLite, including how to import data and query a database using basic SQL statements. | **Homework**: Browse *A Gentle Introduction to SQL*  https://a-gentle-introduction-to-sql.readthedocs.io/en/latest/ | **Story 1 Milestone:** Submit story pitch for first data story, including ideas for story and graphics. |
| Week 3 | **Data analysis Part 2:** Review Part 1. Download and interview a new data set in class. | **Homework**: SQL assignment 1 with selected data set.  **Weekly story memo due** | **Story 1 Milestone:**  You should be exploring your data set, looking for trends and limitations in the data. |
| Week 4 | **Data analysis Part 3:** Review of Part 2. How to join tables and more complex queries. | **Homework**: SQL assignment 2 with data set downloaded from Web.  **Weekly story memo due** | **Story 1 Milestone:**  Your data should be set, and you should be reporting and writing as you go. |
| Week 5 | **Data analysis Part 4:** Review of Part 3 and in class data assignment | **Homework**: SQL assignment 3  **Weekly story memo due** | **Story 1 Milestone:**  You should be conducting interviews and refining your analysis based on reporting and also writing the story. |
| Week 6 | **Introduction to Python/Pandas:** An overview of a more robust and reusable data analysis system. Intro to Github (basic commands, create a local directory, standards for storing homework there, etc.). | **Homework:** Ensure your Python/Pandas/Jupyter install is working  **Weekly story memo due** | **Story 1 Milestone:**  Rough draft of Data Story 1 due, including methodology. |
| Week 7 | **Basic Python/Pandas:** Importing data (text, csv, Excel, APIs); data types; basic methods for exploring the shape of data.  Learn about best practices for basic data tidying before analysis.  Overview of committing to Github from your terminal. | **Homework:** Create a notebook, import a dataset, ask some basic questions. Then commit it to your Github account and submit a link to your work. (Create directories with the week number to store your work).  **Weekly story memo due** | **Story 1 Milestone:**  You will get edits and feedback from instructor on your Data Story.  **Story 2 Milestone:**  Pitch Story 2 |
| Week 8 | **More Python/Pandas:** Basic analysis and aggregation and creating basic charts.  Data exploration and charting using Altair package. | **Homework:** Use an aggregation method (Groupby, for example) in your notebook. Then commit the notebook to your Github account.  **Weekly story memo due** | **Story 1 Milestone:** Final Draft of Story 1 due  **Story 2 Milestone:**  Obtain data and begin analysis and reporting |
| Week 9 | **Intro to basic scraping:** Use Python to acquire data from the front end of a website. Apply what we’ve learned to the newly imported data.Use Github to retrieve timeseries data from an API. | **Homework:** Scrape a table from a website or grab data from an API, perform some basic cleanup, and then commit your work to Github.  **Weekly story memo due** | **Story 2 Milestone:**  You should be exploring your new data set, looking for trends and limitations in the data. |
| Week 10 | **Review of SQL/Python concepts**  **Intro to Mapping:**  Overview of QGIS. Attribute data, importing, selecting and styling map layers. | **Homework**: Mapping assignment 1 with data set downloaded from Web.  Weekly Story Memo due | **Story 2 Milestone:**  Your data should be set and you should feel confident about using it for your story. |
| Week 11 | **Mapping Part 2:**  Joining tabular data to mapping files. Dissolve, union, buffer, spatial query, spatial joins. | **Homework**: Mapping assignment 2 with data set downloaded from Web.  Weekly Story Memo due | **Story 2 Milestone:**  You should be doing data analysis and reporting + conducting interviews and refining your analysis based on reporting and writing the story. |
| Week 12 | **Mapping Part 3:**  Projections and the finer points of mapping. | **Homework:**  Mapping assignment 3 with data set downloaded from Web.  Weekly Story Memo due | **Story 2 Milestone:**  Finishing the story and any follow-up interviews. |
| Week 13 | **Data Visualization Part 1:**  Overview of using the Altair charting library with Python and Pandas. | **Homework**: Start a simple new notebook and create a chart using Altair. Commit your work to Github.  Weekly Story Memo due | **Story 2 Milestone:**  Rough draft of Story 2, methodology and graphics due |
| Week 14 | **Data Visualization Part 2:**  Advanced Altair: faceted charts, styling, basic mapping. | **Homework**: Add a more complicated chart to your notebook and commit your work to Github.  Weekly Story Memo due |  |
| Week 15 | **Statistics in Python:** Standard Deviation, Z-Scores, regression, correlation, and risk ratios | **Homework**: Perform a simple correlation analysis on two variables in a data set and commit your work to GitHub. | **Story 2 Milestone:**  You will get edits and feedback from instructor on Story 2. |
| FINAL EXAM PERIOD | We will meet for a discussion of projects and class summary. |  | **Story 2 Milestone:**  Final draft of Story 2 is due |

**XII. 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 an amount equal to 1 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 or public relations class.

**Statement on Academic Conduct and Support Systems**

**a. Academic Conduct**

*Plagiarism*

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 Part B, Section 11, “Behavior Violating University Standards” [policy.usc.edu/scampus-part-b](https://policy.usc.edu/scampus-part-b/). Other forms of academic dishonesty are equally unacceptable. See additional information in *SCampus* and university policies on scientific misconduct, [policy.usc.edu/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.

**b. Support Systems**

*Counseling and Mental Health - (213) 740-9355 – 24/7 on call*

[studenthealth.usc.edu/counseling](https://studenthealth.usc.edu/counseling/)

Free and confidential mental health treatment for students, including short-term psychotherapy, group counseling, stress fitness workshops, and crisis intervention.

*National Suicide Prevention Lifeline - 1 (800) 273-8255 – 24/7 on call*

[suicidepreventionlifeline.org](http://www.suicidepreventionlifeline.org/)

Free and confidential emotional support to people in suicidal crisis or emotional distress 24 hours a day, 7 days a week.

*Relationship and Sexual Violence Prevention and Services (RSVP) - (213) 740-9355(WELL), press “0” after hours – 24/7 on call*

[studenthealth.usc.edu/sexual-assault](https://studenthealth.usc.edu/sexual-assault/)

Free and confidential therapy services, workshops, and training for situations related to gender-based harm.

*Office of Equity and Diversity (OED)- (213) 740-5086 | Title IX – (213) 821-8298*

[equity.usc.edu](https://equity.usc.edu/), [titleix.usc.edu](http://titleix.usc.edu)

Information about how to get help or help someone affected by harassment or discrimination, rights of protected classes, reporting options, and additional resources for students, faculty, staff, visitors, and applicants. The university prohibits discrimination or harassment based on the following *protected characteristics*: race, color, national origin, ancestry, religion, sex, gender, gender identity, gender expression, sexual orientation, age, physical disability, medical condition, mental disability, marital status, pregnancy, veteran status, genetic information, and any other characteristic which may be specified in applicable laws and governmental regulations. The university also prohibits sexual assault, non-consensual sexual contact, sexual misconduct, intimate partner violence, stalking, malicious dissuasion, retaliation, and violation of interim measures.

*Reporting Incidents of Bias or Harassment - (213) 740-5086 or (213) 821-8298*

[usc-advocate.symplicity.com/care\_report](https://usc-advocate.symplicity.com/care_report/)

Avenue to report incidents of bias, hate crimes, and microaggressions to the Office of Equity and Diversity |Title IX for appropriate investigation, supportive measures, and response.

*The Office of Disability Services and Programs - (213) 740-0776*

[dsp.usc.edu](http://dsp.usc.edu/)

Support and accommodations for students with disabilities. Services include assistance in providing readers/notetakers/interpreters, special accommodations for test taking needs, assistance with architectural barriers, assistive technology, and support for individual needs.

*USC Support and Advocacy - (213) 821-4710*

[uscsa.usc.edu](https://uscsa.usc.edu/)

Assists students and families in resolving complex personal, financial, and academic issues adversely affecting their success as a student.

*Diversity at USC - (213) 740-2101*

[diversity.usc.edu](https://diversity.usc.edu/)

Information on events, programs and training, the Provost’s Diversity and Inclusion Council, Diversity Liaisons for each academic school, chronology, participation, and various resources for students.

*USC Emergency - UPC: (213) 740-4321, HSC: (323) 442-1000 – 24/7 on call*

[dps.usc.edu](http://dps.usc.edu/), [emergency.usc.edu](http://emergency.usc.edu/)

Emergency assistance and avenue to report a crime. Latest updates regarding safety, including ways in which instruction will be continued if an officially declared emergency makes travel to campus infeasible.

*USC Department of Public Safety - UPC: (213) 740-6000, HSC: (323) 442-120 – 24/7 on call*

[dps.usc.edu](http://dps.usc.edu/)

Non-emergency assistance or information.

*Annenberg Student Success Fund*

<https://annenberg.usc.edu/current-students/resources/additional-funding-resources>

The Annenberg Student Success Fund is a donor-funded financial aid account available to USC Annenberg undergraduate and graduate students for non-tuition expenses related to extra- and co-curricular programs and opportunities.

**XIII. About Your Instructor**

**Matt Stiles** is a senior reporter for the *Los Angeles Times* who combines traditional reporting with computer programming, data analysis and data visualization to find and tell compelling stories about the state. He previously covered Los Angeles County, the nation’s largest local municipal government, after reporting on North and South Korea in Seoul for the newspaper’s foreign desk. Prior to joining the Times, Stiles wrote about national economics at *The Wall Street Journal*, served as data editor at NPR headquarters and was the founding news applications editor at The Texas Tribune. He lives in Los Angeles and tweets as @stiles.

**Bio:** https://www.latimes.com/people/matt-stiles