Data Journalism Degree

Master’s Projects Guidelines and Deadlines



# 

# INTRODUCTION

At Columbia, the data journalism degree master’s project is an ambitious, serious and sustained data-driven work of journalism. Students design, undertake and complete the project independently or in pairs, working with individual advisors. To meet these requirements, projects need to demonstrate mastery of obtaining, analyzing and presenting data to a broad audience.

The exact form of the master’s project is flexible, with various options outlined later in this document. All data degree master’s projects have to include a data component, along with a separate, detailed description of the methodology followed during the research process as well as code documentation (details below).

The Master's Project is a tutorial-style course. The instructor works with the student on a one-to-one basis, shepherding the story through each stage of the process from narrowing down their initial ideas to the completion of their thesis. Students look to instructors for help in a number of areas, including:

* Which sources to seek out
* What technology and methodology to use
* What sorts of questions to ask
* How to navigate roadblocks
* How to test the validity of their reporting or data product.

Helping students structure their story, their data analysis and compose a narrative arc will likely encompass the majority of the time.

# MEETINGS

As students settle on a topic and begin the reporting process, we recommend students and advisers meet either individually or as a group frequently to refine and develop ideas. Once the main body of writing and reporting is underway, meetings can be held as necessary to ensure consistent development.

The ideal frequency varies from one adviser to another, but we recommend weekly meetings in the earliest stages of the project with a potential decrease to every two weeks later in the process.

While deadlines for each stage are outlined below, Master’s Projects should be thought of as iterative documents that undergo continual improvement. From one meeting to the next, **students should provide evidence of having made significant progress in their reporting**, and later, in their writing.

Meetings can be held at the school, in a public place such as a coffee shop or library, or even at your place of work.

# CHOOSING A TOPIC

Students are expected to come up with a variety of proposals for a topic.

Instructors are expected to pass judgment as professional editors. Bases for evaluation might include:

* Is the subject newsworthy?
* Would it be interesting and/or important to an informed audience?
* Is it manageable; i.e., can the student get the necessary access?
* Is the selected data enough to prove the hypothesis?
* Is the data quality adequate for the project?
* Is the proposed methodology adequate?

Advisors should keep in mind that pivots and dead-ends are a common occurrence, and encourage students to maintain backup plans and a sense of flexibility.

# THE PROCESS

Work on the project can be roughly divided into three sections:

1. Picking a topic
2. Research, data collection, analysis, and reporting
3. Writing and rewriting

The first few weeks of the semester should be spent hammering out a workable idea, during which instructors are encouraged to challenge the students to make sure they can deliver the topic they have chosen.

Once a project is approved, the next phase involves research and initial reporting. Students should test their topics to make sure they stand up in terms of access, viability, data quality, and public interest value. Students should stay flexible and ready to pivot if a story does not prove viable, but be mindful of making steady progress each week.

Some instructors wish to see written evidence of this progress in the form of updated outlines or story proposals. Others prefer to gauge their students’ progress in discussion-oriented meetings. Unlike students in the general MS curriculum, Data Journalism MS students have a light course load during this semester and should be able to dedicate ample time each week to making headway on their project.

Memos and drafts can be handed in or emailed to the adviser, and should also uploaded into the Master’s Project course management system. In order to accommodate an efficient schedule, instructors should return edited drafts and hold conferences within one week of submission.

For the final project, one copy is due at noon of the deadline day **in the Dean of Students office**. This is the version of the Master's Project that goes on record in the Journalism Library. A second copy should be emailed or delivered to the adviser. Students will receive detailed formatting instructions a week or so prior to the deadline.

*Advisers may give students permission to hand in their projects prior to the deadline if they are satisfied with the final version, but the Dean of Students Office needs to be alerted.*

# 

# DATA MASTER’S PROJECTS DEADLINES

Each project will require different amounts of investment in the various stages of the project. Some may be front-loaded and involve a large amount of initial research or data work, while others may get to the draft stage quickly and include more rounds of edits.

With this in mind, **individual deadlines may be coordinated between students and their advisors to best serve each individual master’s project**. Chances are you have your own preferred workflow, and it’s fine to completely sidestep our process.

The deadlines below are provided as broad guardrails to help ensure a finished, viable submission at the end of the summer semester. Advisors are welcome to accelerate the schedule if suitable.

Beyond the dates below, advisors may also request official second and third drafts, while others prefer frequent meetings with a continually-updated document.

|  |  |  |
| --- | --- | --- |
| Date | Submission | Description |
| April 14 | First project memo | An initial attempt to flesh out the project idea(s).  This requires (at least) formulating a hypothesis, doing a clip/archival search, and consulting a few sources, even if they are just advocates or experts willing to talk.  Please review the format of the final memo [here](https://docs.google.com/document/d/1tJ0e27-raEAoQjHeQMYvvMv4k1l-u3ZSmlQBLUMNRD0/edit). |
| May 19 | Second project memo and final approval | A fleshed out memo that includes timetable, source list, possible characters. The memo should follow the [format guidelines](https://docs.google.com/document/d/1tJ0e27-raEAoQjHeQMYvvMv4k1l-u3ZSmlQBLUMNRD0/edit).  Sample project memos can be found [here](https://drive.google.com/file/d/0B7UAuEqoequKYnlWUEVRRi1VVTA/view?usp=sharing). |
| June 16 | First draft | Write a draft based on the gathered material.  Expectations for this draft will be set by the advisor. Some might just want an opening with a nut graf and an outline, while others will want a fully fleshed-out document. |
| *Adviser’s discretion* | Interim drafts | Each advisor will have their own methods of providing feedback and assessing progress. Students may be required to submit interim drafts, work through a continually-updated document, or other processes. |
| August 12 | Final draft | The final version of the project, ready for submission. |

# FORMAT

**All Master’s Project formats must include a data component**. At least one of the sources for the stories has to come from any kind of data (structured or unstructured).

Students must clear with their adviser which format they will choose. Advisers will work with them in determining whether it is an appropriate subject for the selected format, and also how to divide their material between the narrative, app or visualization, tool or platform portions of the project. The data product can also include video, audio or animation. The School will **evaluate case by case if it can** provide the resources (equipment or specialized adviser) to produce them.

|  |  |
| --- | --- |
| **Format** | **Description and examples** |
| **Regular data-driven story**  (Story between 2500-5000 words, plus documentation and methodology) | Students will use one or multiple datasets from reliable organizations to supplement traditional reporting in proving a hypothesis. Data may be acquired through downloading, scraping or filing FOIA requests, among other methods.  **ProPublica:** [Temporary Work, Lasting Harm](https://www.propublica.org/article/temporary-work-lasting-harm) - [methodology](https://www.propublica.org/nerds/how-we-calculated-injury-rates-for-temp-and-non-temp-workers)  **The Washington Post:** [Left with nothing](https://www.washingtonpost.com/sf/investigative/2013/09/08/left-with-nothing/), [Suuspicious Bidding](https://www.washingtonpost.com/sf/investigative/2013/09/09/suspicious-bidding/), [Mistakes put homes in peril](https://www.washingtonpost.com/sf/investigative/2013/09/10/mistakes-put-homes-in-peril/)  **The Marshall Project:** [Afraid of Jail?](https://www.themarshallproject.org/2017/03/09/afraid-of-jail-buy-an-upgrade) - [methodology](https://www.themarshallproject.org/2017/03/09/how-we-crunched-california-s-pay-to-stay-data) |
| **Creating a data collection/survey**  (Maximum 2500 words, plus documentation and methodology) | Students design a methodology to create a non-existing dataset or series of datasets to prove a hypothesis. They can write their own scripts to automate the data extraction process from websites or social media. They can also automate requests, queries or posts to social media or emails to get responses from third parties. This format also covers those stories that require OCRing documents to create a dataset.  Students may experiment with machine learning, artificial intelligence and other innovative technology and devices to collect their data such as sensors, drones, GPS receptors, flight trackers, etc.  **BR Data and SPIEGEL:** [No place for Foreigners](https://www.hanna-und-ismail.de/english/index.html) - [methodology](https://www.hanna-und-ismail.de/english/methodology.html)  **BuzzFeed News:** [We Trained A Computer To Search For Spy Planes](https://www.buzzfeed.com/peteraldhous/hidden-spy-planes?utm_term=.hhZ78Z79a#.sjk5Xq5KP)  **ProPublica:** [How to Get Censored on China’s Twitter](https://www.propublica.org/article/how-to-get-censored-on-chinas-twitter) - [methodology](https://www.propublica.org/article/how-we-observed-censorship-on-sina-weibo) |
| **Algorithmic accountability/ platform analysis**  (Maximum 2500 words, plus documentation and methodology) | Students examine an algorithm or platform in detail to decode it, discover its assumptions and the results it produces on people’s lives.  In the process, students analyze algorithmic accountability and find the outcome(s) of its application: discrimination, denial of service, censorship, violation of privacy, false prediction. Students can also program their own algorithm to analyze data, metadata, videos, pictures, etc.  **ProPublica**: [Machine Bias](https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing)  **The Markup:** [Pixel Hunt](https://themarkup.org/series/pixel-hunt) |
| **News data platform/tool**  (Maximum 2500 words, plus documentation and methodology) | Students produce a data platform or tool to collect, clean, aggregate, translate, consolidate, analyze and/or present data from one or multiple sources. The platform or tool must have a journalistic purpose or solve an existing journalistic problem.  When the product is completed, students must use concrete journalistic cases to prove and describe how they solved the problem. They should complete the reporting and writing of at least one of those stories.  **elections.**  **ProPublica:** [Electionland](https://projects.propublica.org/graphics/electionland)  **ProPublica:** [How We Are Monitoring Political Ads on Facebook](https://www.propublica.org/article/how-we-are-monitoring-political-ads-on-facebook)  **The Markup:** [Citizen Browser](https://themarkup.org/series/citizen-browser) |
| **Interactive, visualization, or visual narrative**  (Maximum 2500 words, plus documentation and methodology) | Students produce an interactive or static visualization or animation based on one or several datasets.  The interactive/visualization must be combined with a story and must accomplish a journalistic purpose. Data must be a significant part of the project.  **Reuters:** [Point of No Return](https://www.reuters.com/investigates/special-report/myanmar-rohingya-return/) **Wall Street Journal:** [The Impact of Vaccines](http://graphics.wsj.com/infectious-diseases-and-vaccines/) **FiveThirtyEight:** [What Would It Take To Turn States?](https://projects.fivethirtyeight.com/2016-swing-the-election/) |

# 

# SOURCES AND METHODOLOGY

A list of all sources (including human sources, data sources, public records) and contact information should accompany the final draft, to allow instructors time to conduct a random check of the story's veracity.

In addition to the list, students must deliver a detailed explanation of the methodology and editorial decisions they applied to their data component. Code and data must be provided to the fullest possible extent, with the aim of complete reproducibility by a third party.

# [Here](https://www.propublica.org/article/surgeon-level-risk-short-methodology) is an methodology example from ProPublica.

# FURTHER LOGISTICS

The Master’s Project is a for-credit course in the Summer. All advisers must use the online grading system to submit a pass/fail grade for each advisee at the end of the semester.

In addition to submitting their memos and drafts to their advisors, students will also be submitting interim materials through CourseWorks. We’ll also be conducting quick check-ins for both advisors and students. This will help us take stock of how a lighter workload with an accelerated schedule is working.

# DISCIPLINE AND ACADEMIC WARNINGS

Students are required to arrive on time to meetings with their advisers, make deadlines, and follow the adviser’s instructions. Since day one, the adviser will set up the standards for the master’s projects and the consequences for not meeting those standards.

Advisers will communicate with the students immediately if the advisor considers that the student is underperforming. In that case, students should expect clear and detailed instructions to follow up. Students must let their advisers know immediately if they find the instructions unclear.

Both students and the adviser benefit from clear, written performance parameters set at the beginning of the term and the adviser’s enforcement of the policies he or she has set for the academic year.

Here are the steps that we ask advisers to take to handle students who are not meeting the standards as detailed in the written instructions:

* Verbal warning from adviser.
* Written warning via email from adviser. Includes language informing the student that if there is no improvement by particular date, the matter will be referred to the Dean of Students Office for official warning.
* Academic Warning from the Dean of Students office. Letter includes specific targets/assignments with due dates that must be met to be removed from warning status.
* Academic Probation from the Dean of Students office. Letter includes specific targets/assignments with dates that must be met to be removed from probation. Letter also informs the student that not improving will result in failing the class.
* Student fails class and is notified by the adviser

# ADDITIONAL SUGGESTIONS

**Review other Projects:** the last five years of Master’s Projects are kept in the Journalism Library, which is remarkably difficult to access. I will work on getting digital copies that we can all access.