CSS Capstone Project Guide

UCSD CSS Program

This document is intended to serve as a single point of reference for how a CSS Capstone is supposed to work, for students, affiliated faculty, and Capstone Hosts. Details about registration, requirements, and responsibilities of students are found in the UCSD CSS Student Handbook. Although it provides guidance as to the expectations of the program for both students and faculty, faculty are free and encouraged to exercise their own judgement as to how to best guide their students through this process.

Contents

1 General Capstone Project Guidelines

The CSS Capstone project is meant to serve three important purposes:

1) The project should give students **'real world' experience** working with social science data and answering concrete and useful questions
2) The project should provide a **means of demonstrating student ability** to do computational social science in the real world
3) The project should **serve as a piece of their portfolio**, which they can use while applying to jobs, academic and otherwise.

Thus, all 'requirements' for a given project should be viewed as stemming directly from these three desires, and so long as you're able to explain how a variation or deviance from the below is reasonable and serves these goals, we're open to modification to best suit the individual needs of the project.

Finally, it's worth noting that **your entire project depends on a successful final submission**. Although you'll be graded as Satisfactory/Unsatisfactory in CSS 296 for each of three quarters, these intermediate quarters are simply a way to give you credit for your time spent working towards your capstone. Ultimately, you must complete *an entire capstone project* to earn this degree.

1.1 Projects involve Real World Experience

CSS Capstone projects are meant to give students actual experience with actual data as found in the real world. Similarly, one of the key distinctions between CSS and more generic data science programs is in the nature and complexity of questions asked. Thus, a successful project should minimally involve:

- **Working with actual, collected social science data**
  - Ideally, students will demonstrate experience with the entire life cycle of data, with elements of data cleaning, tidying, transformation and formatting, and storage
  - Some projects will start with clean data, and that's fine, but please don't feel like having 'clean' data is a requirement for the program
  - We're open to projects involving large amounts of simulation, but some work with real-world data is a desirable aspect
- **Addressing questions and hypotheses grounded in the real world**
  - There will likely be many analyses in a given project, and many will be 'summary' in nature (e.g. 'What is the distribution of nationalities of users of this tool?' or 'What's the average age of people without homes in San Diego?'), and that's fine and these should be reported as well
  - Students should also demonstrate some engagement with actual hypothesis testing (e.g. 'Does the nationality of a participant predict their likelihood to recommend the tool?' or 'Are anti-homeless laws enforced more often against younger people without homes?')
  - Students in projects without inherent hypotheses (e.g. those helping an organization to better understand the demographics and nature of the people they serve) are expected to work with their capstone hosts and find hypotheses to test using the data which they find.
  - Although this is framed in terms of classical statistics, machine learning approaches are very welcome as well, provided the models are well explained and (ideally) tested and investigated

1.2 Projects demonstrate Student Ability

CSS Capstone projects will vary greatly in their nature, but from the perspective of the program, a great capstone will show what the student has learned from the program to the Capstone Host, CSS Faculty Advisor, and designated readers. To this end, a successful project should involve:

- **Demonstration of understanding** of the data collection, cleaning, and analysis process
  - A final project consisting of a large number of model outputs and visualizations with no further commentary is not acceptable
  - Students are expected to describe the state of the data, any efforts taken to transform it into analyzable form (e.g. collection, cleaning, transforms, and anything else)
  - Including the code which accomplishes the analysis (e.g. in a Jupyter Notebook) is encouraged but not sufficient, please explain what it does
  - For each analysis, include discussion of the questions, the method for arriving at the chosen modeling approach, and the model
  - Include some discussion of *what each model actually shows*, interpreting the figures and coefficients in a way which would be understandable to a future employer without a strong data science background
- **Demonstration of a variety of different methods**
  - A project which shows that students are able to take the mean of 50 different variables isn't a great demonstration of your knowledge
  - Projects should demonstrate different analytical approaches where possible
  - Students are encouraged to address different kinds of questions (e.g. questions suited to both inferential and predictive/ML analysis)
  - If a project truly only needs one or two types of analysis (e.g. a series of logistic regressions), students should describe the other methods they considered and why this is the best choice
- **Demonstration of skill with data visualization**
  - Analyses should include visually compelling, understandable, and readily interpretable plots and graphs where possible
  - Ensure that axes are labeled and interpretable (where possible/reasonable)
  - Avoid deceptive axis scaling for comparisons, or where necessary (e.g. for variables with different effect sizes), highlight it and explain the choice
- **Clear highlighting of novel approaches or unusual elements**
  - Although many projects will fall under the umbrella of 'data science', some may involve different kinds of computation and approaches outside the conventional 'data analysis' realm, and these aspects should not be ignored
  - Where students have acquired new skills or used unusual techniques, highlight them, and discuss what they are, why they're useful, and how they were learned

1.3 Projects serve as a piece of the student's portfolio

Finally, these capstone projects are meant *specifically* to serve as a demonstration of the student's skill with computational social science methods. As a result, successful projects will include:

- **Public, Open, and Non-Proprietary Data**
  - Although your project may involve the use of proprietary, privileged, or IRB-supervised data, your final project should be composed and formatted such that it can be freely shared, released, and publicized with no negative consequence or legal concern
  - Human participants should be discussed only as coded identifiers (e.g. 'Participant 45'), or psuedonyms, and any released sensitive data should be de-identified.
  - All released data, data labels, and code in the project should be non-sensitive
  - If proprietary or sensitive data which cannot be released is used in one or more steps, students should

give a simulated, publicly shareable example of the formatting and nature of the proprietary data, and explain roughly what it contained and how it was used.
  - The Capstone Host **must** review the final project in its final form before it is released to the CSS Faculty Advisor and CSS Reader, to ensure that no proprietary or sensitive data is contained in the released version of the final project, allowing students to freely distribute the project
- **Interpretability of the data and questions** for people outside of the field of practice
  - In cases where the project is working in a very specific domain, students will want to include some background information such that a future employer could understand what is happening and the questions asked with only a basic understanding of statistics and technology
  - Students will also want to take time to explain the nature and interpretation of specific variables in use, again assuming little background in the domain of study
- **Clear explanations of analyses for a lay person**
  - Students should write this for somebody who has relatively little statistics background, and explain what each test is examining, and what we can reasonably conclude
- **Clean, interpretable visualizations** to help tell the story
  - Students should use visualizations to tell as much of the story as possible, using model outputs to reinforce the message
  - Prefer visualizations which are understandable with a clear caption
  - In cases where esoteric visualizations are needed (e.g. QQ Plots, AUC-ROC curves), explain what they mean, or explain what they show, in captions
- **Clean overall formatting**
  - Although students are welcome to use whatever writing and style conventions they're most comfortable with, please make the formatting meaningful, helpful, and such that it 'gets out of the way' and simply helps to communicate
- **Demonstration of teaching**
  - Students will be explaining data for the rest of their lives, which almost always involves teaching, and as such, good writing is good teaching
  - Students should teach that which a reasonably sophisticated human would need to understand the work

1.4 Students must complete a Full Project

Finally, it's worth noting that **your degree depends on a successful final capstone submission**. Although you'll be graded as Satisfactory/Unsatisfactory in CSS 296 for Fall and Winter quarters, these intermediate quarters are simply a way to give you credit for your time spent working towards your capstone. Ultimately, you must complete *an entire capstone project* to earn this degree, and without a successful final project submission, graduation is not possible.

This means that students who fail to produce a Satisfactory capstone project in Spring even after two 'Satisfactory' quarters may still be asked to repeat *the entire capstone process* in a future year with a new project/host, at the discretion of the education committee.

1.5 A Note on Terminology

For the purposes of this document, there are five different roles, which may overlap depending on the project.

- **Students:** CSS M.S. Students, participating in the project individually or as a team
- **Capstone Host:** This is the person (or team) who the students are working 'for', who is designating the project, providing the students with data, questions, and domain support, and interacting with the students on a weekly basis to keep them on track.
  - For Off-Campus Projects, the Capstone host is the person within the host organization (not UCSD) who is managing the capstone.

  - For On-Campus Projects hosted by faculty or staff who are not CSS Affiliates (e.g. a UCSD Professor of Literature, or UCSD Staff Librarian), the Capstone host is the faculty member who proposed and is 'running' the project. Students will also be assigned a 'CSS Faculty Advisor' who is a CSS Affiliate and will serve in the role detailed below.
  - For CSS Affiliate Projects, the Capstone host is the same person as the CSS Faculty Advisor below, and will take on both roles and complete the tasks assigned to both.
- **CSS Faculty Advisor:** This is a designated CSS Faculty Affiliate who will meet with Students on a regular basis, help guide the project, provide students with input, feedback, and assistance, and read the students' final paper, assigning a 'Satisfactory' or 'Unsatisfactory' grade.
- **Capstone Reader:** This is a CSS Faculty Affiliate, chosen by the students in consultation with the CSS Faculty Advisor, who will serve as a second reader on the project, also evaluating the project as 'Satisfactory' or 'Unsatisfactory'.
- **CSS Director:** This is the Director of the CSS Program, who has no formal role, but may be called upon to help resolve conflicts, intervene when students are struggling, or serve as a tie-breaking vote if the CSS Faculty Advisor and Reader differ in their assessment of the paper.

2 Capstone Timelines

Each year, the Capstone process will follow a similar timeline, with exact dates varying year-to-year.

2.1 Spring and Summer before the Program Starts

- Students and Faculty may submit capstone ideas (e.g. "I'd love to work with Dunder Mifflin, a local paper company with a major data science problem…" or "My colleague Ethan is doing some interesting computational sociology and could use some help…") to the CSS director.
- All Capstone proposals should be submitted in late August

2.2 September during the Bootcamp

- Proposals are finalized in Early September and students are given the opportunity to indicate their project placement preferences.
- During this time, students are also expected to fill out the CSS Capstone Student Participation Agreement
- Students are matched with projects in Mid-to-Late September, and begin regular meetings with their project hosts and CSS faculty advisors

2.3 Fall Quarter

During the Fall Quarter, **Students** will…

- Focus on gaining the necessary background to understand both the problem(s) and the organization in which they're situated
- Formalize the problem(s), identifying the sorts of methods by which it can be addressed or investigated
- Work to understand the type and nature of the data they're working with and collecting
- Think about the kinds of data cleaning which may be required to put this data in a usable state
- Plan the approaches which will be brought to bear on the problem(s)
- Ensure that the CSS faculty advisor (where applicable) is 'on board' and fully understanding the task, the process, and areas where students need particular support.
- Work with their CSS Faculty Advisor to select a Capstone Reader for the project, reach out to that person to confirm, and then notify CSS advising of their choice
- **Submit their Fall 'Questions and Methods' docment (See 'Project Deliverables') to their Faculty Mentor by a mutually agreed-upon date**

  - Please send as completed copy to the CSS Program Director as well, but your Faculty Mentor is the person whose read-through matters.

During the Fall Quarter, **Capstone Hosts** will…

- Assign readings and/or research tasks which will prepare the students to work in the specific field of study where the capstone is occurring
- Work with the students to identify specific questions, hypotheses, ideas, or fruitful regions for research
- Contact the CSS Director to arrange intervention and support in the event that one or more students are doing 'Unsatisfactory' work, are not communicating, or are not engaged with the project
- Review the students' 'Questions and Methods' paper to ensure that it seems plausible to lead to a good capstone
- **On or near December 1st**, Capstone Hosts should inform the CSS Director whether student work has been 'Satisfactory' or 'Unsatisfactory'
  - This will form the basis of their Fall CSS 209 grade, which is entered by the Director for convenience
  - 'Unsatisfactory' is equivalent to 'failing the class' and will block their progress, and should be assigned only in situations where the student is not engaged, doing low-quality work despite repeated corrections, or is otherwise not satisfactorily performing the tasks.

During the Fall Quarter, **CSS Faculty Advisors** will…

- Meet regularly with students and discuss the project, the nature of the analyses needed
- Start working with the student(s) to form a reasonable project idea
- Contact the CSS Director if the student drops off the radar, or fails to reach out or respond
- Assist the students in identifying a Capstone Reader among our CSS Affiliated Faculty (in any department) who has experience relevant to the work, or failing that, would be able to fairly judge the output.

During the Fall Quarter, **Capstone Readers** will…

- Be approached by students, and confirm their willingness to be a reader

2.4 Winter Quarter

During the Winter Quarter, **students** will…

- Implement the core components of their modeling and data analysis, creating formal, robust, and defensible models
- Revise their deliverables list with their project host based on any additional difficulties or opportunities which may arise
- Identify any particularly difficult steps still ahead and address them with their project hosts and faculty advisor, soliciting help and support as needed
- Begin outlining their final paper, considering the organization and format (e.g. paper, website, Jupyter Notebook with integral prose, etc.) which would best suit the needs of the project
- **Submit their Winter 'Full Outline' docment (See 'Project Deliverables') to their Faculty Mentor by a mutually agreed-upon date**

During the Winter Quarter, **Capstone Hosts** will…

- Ensure that the students have settled on a *concrete, deliverable final project*
  - They should also ensure that the scope is ambitious enough to be both useful and something students can be proud of
  - But they should also ensure that the project is tractable in the coming months
- Contact the CSS Director to arrange intervention and support in the event that one or more students are doing 'Unsatisfactory' work
- Discuss the students' outline with them, again to ensure a successful project

During the Winter Quarter, **CSS Faculty Advisors** will work to ensure that…

- Students have regular meetings
- The project as described will meet the guidelines for a CSS Capstone Project as laid out here
- The hypotheses and questions being investigated are being evaluated properly by the proposed work and format
- The analyses under discussion are adequate and well-suited to the task
- The scope is both sufficient to show learning, but also reasonable in the allotted time

During the Winter Quarter, **Capstone Readers**…

- May choose to meet with students to discuss the project and offer input at early stages
  - Readers are not expected to regularly meet with students, and are being generous with their time if they offer to do so.

2.5 Spring Quarter

During the Spring Quarter, **students** will…

- Complete their projects, identifying and then creating a final set of deliverable analyses, models, and outcomes
- Focus on finalizing their analysis with rigorous, reproducible methods, compelling visualizations, and transparent explanations
- Work with project hosts to put this into a format useful and usable to the host
- Create their final project and presentation, constituting their digital portfolio, to be submitted to UCSD

During the Spring Quarter, **Capstone Hosts** will…

- Contact the CSS Director to arrange intervention and support in the event that one or more students are doing 'Unsatisfactory' work or if there's a concern that students won't finish by Week 7 (e.g. Mid-May)

During the Spring Quarter, **CSS Faculty Advisors** will…

- Help ensure that the proposed project is being completed on a reasonable timeline to complete by Week 7 of the quarter (and contact the CSS Administration Team if this is in question)
- Ensure that the project being discussed will meet the requirements outlined here
- Offer support to students as they finalize their work
- Advisors may elect to request or review paper drafts
- Advisors are expected to serve as one of the graders for the final project, as described below.

During the Spring Quarter, **Capstone Readers**…

- May choose to meet with students to discuss the project and offer input prior to submission
  - Again, readers are not expected to meet with students, and are being generous with their time if they offer to do so.

3 Project Deliverables

Each quarter, students will be required to 'Deliver' a portion of their capstone to their capstone advisor.

3.1 Fall Quarter - Questions and Methods

At the conclusion of the Fall Quarter, students will be asked to submit a preliminary draft of their final project's Introduction, Background, and Methods sections which include:

- The research questions or hypotheses they plan to address in their project, enumerated clearly

- A few references to relevant literature, where applicable
- The data they intend to use to address these questions
- The methods which they intend to use to address these questions

Although this document is *not binding*, and students are welcome to revise their questions, data, and methods later on, this should allow students and their mentors to make sure they're on the same page, addressing interesting and answerable questions, and using methods which are reasonable and interesting. This will be graded on completion, but poor performance here can be grounds for an "Unsatisfactory" Fall Quarter grade.

The exact manner by which faculty evaluate this paper is up to them, but the program's goal for this paper is to provide strong indication that the student is 'on the right track' and thinking reasonably about how to move forward towards a completed project.

3.2 Winter Quarter - Full Outline

At the conclusion of Winter Quarter, students will submit a complete outline of their final project, discussing in detail: - The anticipated format of their capstone project (e.g. Paper, online notebook, etc) - The final form of their research questions and/or hypotheses - The literature they're planning to engage with - The data they're using - The methods they'll be using - The specific analyses they'll be conducting to address each question - The kinds of plots they anticipate using to show the data

Students, where possible, should make direct reference to the rubric below to ensure they're on track for a successful project.

As before, this is *not binding*, and students who find problems with a question or line of analysis in Spring are encouraged to work with their capstone host and faculty mentor to revise and come up with a better approach even if it differs from their proposed outline. Instead, this serves to demonstrate to your faculty mentor that you're on the right path, and that your current planned project has the potential to lead to a successful capstone. This will be graded on completion, but poor performance here can be grounds for an "Unsatisfactory" Winter Quarter grade.

3.3 Spring Quarter - Project Submission

Four weeks prior to Spring Final Exams (i.e. Week 6), **students** will submit a draft of their final project to their Capstone Hosts (if separate from their CSS Faculty Advisors). At this point, **Capstone Hosts** are expected to look over the final project write-up to ensure that no private, privileged, confidential, or IRB-protected data or information is included, and that the entirety of the document can be publicly presented and distributed.

Three weeks prior to Spring Final Exams (i.e. Week 7), students will submit their final project paper to the **CSS Faculty Advisor as well as a Capstone reader** designated by the CSS program.

At this point, **CSS Faculty Advisors and Capstone Readers** are expected to read the project, offer feedback, and if any revisions or rewrites are required, describe their nature to the students by **Week 8** of Spring quarter. **Students** must complete revisions or rewrites **prior to the start of Final Exam Week**.

During Spring Quarter Finals Week, students will present their work at a public event where CSS faculty, capstone partners, industry representatives, and others are welcome.

Following the Presentations, **Capstone Hosts** are encouraged to communicate with the CSS Director to give feedback on the capstone process, and discuss whether they're interested in hosting a subsequent year's capstone.

4 Project Format

The exact format of the project is not crucial so long as it can be easily shared with readers and employers, and should be chosen by the students, capstone host, and CSS Faculty Advisor to optimally serve the needs of the project. The rubric assumes a classical academic paper, but alternative formats are very much able to meet these goals, and may better serve the goal of demonstrating knowledge for some types of project. Some reasonable formats could be…

- A PDF document (e.g. generated from LaTeXor Google Docs)
- A Camera-Ready Journal Article, in the format and style expected by the journal with embedded plots, and links to supplementary material
  - In these cases, please let us know what requirements were imposed by the journal (e.g. 'all plots must be black and white' or '5000 word limit'), and the target journal
- A compiled HTML Jupyter or RMarkdown Notebook containing literate code
  - Simply submitting code and output is not sufficient, the project should include explanation and discussion interspersed with code and visualizations, using Markdown blocks
- An interactive hosted website, containing multimedia content or interactive plots
  - This can also be an uploaded Jupyter Notebook output
  - Students are responsible for finding hosting (e.g. through Github pages, or through their Capstone hosts)
- A working Google CoLab instance containing a Jupyter notebook which includes the code, data, and writeup
  - Students must ensure that the instance is publicly available
- A (sizable) presentation or video, including all the details discussed in the rubric
  - This would be delivered separately from the capstone presentation

In some cases, the project host may request a particular type of output be generated to make this project's output useful to them. As such, students may include as a part of their project alternative formats, such as a YouTube video, an infographic, a page hosted on the company's website, or a publication formatted and sized for a particular journal. These are welcome, will be considered as a part of the students' academic output, and can be referred to or included as a part of the formal project writeup, but students should ensure that the requirements for the project are met, regardless of the chosen method, and a series of infographics (e.g.) is unlikely to meet the rubric items below.

Finally, students are responsible for ensuring that all necessary links are working, all downloads are downloadable, and that the final project is readable, accessible, and usable without requiring additional or proprietary software downloads by the reader.

5 Project Rubric

Ultimately, the projects will be reviewed by both the student's faculty advisor and a third-party reader from the CSS program. The goal of this process is to ensure that the student has done satisfactory work, and thus, has completed the Capstone (CSS 296) element of the CSS M.S. program. To help students understand some of the elements which go into this decision, we've created the following rubric, where 'Masterful' and 'Acceptable' push a project towards 'Satisfactory', and 'Novice' and 'Way Off' push a project towards 'Unsatisfactory'.

Ultimately, students will be evaluated according to faculty's overall judgement of the sum total of the work, and sufficient insufficiency in any one area could push the entire work over the line. One should treat this rubric as an explanation of how to produce an excellent paper, rather than a binding law which will govern how faculty evaluate the paper, and naturally, faculty will apply their own perspective, which this is not intended to supercede or override.

Finally, note that this rubric is written for a typical 'data analysis' sort of project. If your project involves a different

aspect of computation or engages with data differently, students can work with their CSS Faculty Advisors to suggest alternative rubric interpretations.

5.1 Cover Letter

All projects are required to include a cover letter.

- **Masterful:** Completes and includes a cover page including at least:
  - Project Title
  - Name of all participants
  - List of any supplementary files or urls (e.g. source code links, website link, etc)
  - Any instructions for readers (e.g. 'Start here on the website, then go to this page, then…')
  - Brief description of the relative contributions of each group member, ideally itemized where possible, and where all work was joint, expressing the division of labor.
- **Way Off:** Any element of the cover page is missing.

5.2 Real World Experience

- **Masterful:** Project includes collection, aggregation, cleaning, or analysis of real-world data or computational endeavors, representing real problems in computational social science, and engages with the real-world utility of the work.
- **Acceptable:** Project uses real-world data, which has been previously cleaned, scrubbed, or partially simulated, but still shows engagement with the real-world utility of the data, and involves additional transformation for analysis.
- **Novice:** Project uses existing data or analytical scripts previously prepared for analysis without any cleaning, transformation, or re-factoring, or consists solely of applying an existing analysis to a new or expanded dataset. The utility of the work in a social science context is not well addressed.
- **Way Off:** Project uses manicured 'off the shelf' datasets with 'off the shelf' analysis or ML to answer questions with little relevance to the social science.

5.3 Hypotheses and Questions

- **Masterful:** Project includes two or more hypotheses (or sub-hypotheses), making specific predictions and using inferential statistics, machine learning, or implementation of computational modeling with post-hoc analysis to establish whether the hypotheses are supported or not by the results. All discussed hypotheses are addressed cleanly and clearly.
- **Acceptable:** Project engages with one or more hypothesis, but without making predictions, or with relatively little discussion of the hypothesis following analysis. Also possible when several hypotheses are discussed, but are weakly specified.
- **Novice:** Project engages only with weakly specified hypotheses, fails to adequately evaluate the hypotheses tested in light of the data, or leaves explicit hypotheses unaddressed.
- **Way Off:** Project evaluates no specific questions, presenting only summary statistics, engages with multiple unfalsifiable hypotheses without acknowledging this fact, or engages with a hypothesis without performing analyses which actually address it

5.4 Scope

- **Masterful:** The project includes sufficient and concise information about the background, questions, data cleaning, code, analysis, results to allow the reader clearly understand what's going on, alongside a suitable introduction, discussion, and conclusion, without feeling 'bloated' or unwieldy
- **Acceptable:** The project provides good information about the above elements of the analysis, but is weaker in some areas or lacking suitable discussion, introduction or conclusion. Also relevant where the project is

either too limited in scope as to feel insufficient for a year-long project, or is so broad that no one section has sufficient detail or richness.

- **Novice:** The project is missing or greatly deficient in discussion of one or more elements listed above, or the scope is so small as to poorly reflect the student's learning and work over the course of the year.
- **Way Off:** The project is substantially lacking in description of more than one of the above areas, or the scope of the project is itself deficient, reflecting work which could feasibly be performed in a week or two.

5.5 Demonstration of Knowledge

- **Masterful:** The project shows sufficient complexity and variety in approach, methods, and visualization to demonstrate skill and knowledge and reflect an entire academic year's work and training. The project is completely free of major misconceptions and analytic errors, shows few if any minor misunderstandings, and reflects excellent command of the CSS M.S. program curriculum.
- **Acceptable:** The project shows reasonable complexity and variety in approach, methods, and visualization to demonstrate skill and knowledge and reflect an entire academic year's work and training. The project may contain some misunderstandings or analytic errors, but they don't substantially invalidate the work, and the student demonstrates considerable attention to the CSS M.S. program curriculum.
- **Novice:** The project doesn't have enough complexity in scope, methods, approach, or visualization to adequately demonstrate the student's learning, or contains substantial misunderstandings, errors, or problems with the analytic approach which render the analysis unreliable in part or whole. The student shows some knowledge of the CSS curriculum, but with major deficiencies and evidence of struggle with many concepts.
- **Way Off:** The project demonstrates little sophistication or understanding of the kinds of methods, approaches, visualizations, and argumentation expected of a graduate level student. The analysis may also contain serious problems which render most or all of the work unreliable, or demonstrate substantial struggle with the CSS curriculum on the whole.

5.6 Clarity and Comprehensibility

- **Masterful:** The project is written in such a way that it contains the necessary background to understand the core questions even for somebody outside of the domain of research and inquiry, and contains sufficient explanation of the models, code, visualizations, and findings such that an intelligent person with basic familiarity with statistical nuance could understand the core ideas and findings. All steps are explained (or demonstrated with code) such that they could be reproduced by another researcher, and any included code is commented.
- **Acceptable:** The project is written such that it is largely understandable to somebody with only basic background in the domain of research and computational data analysis, albeit with some elements insufficiently explained and some background knowledge questionably assumed. The process is detailed sufficiently to allow reproduction of the work, although some additional follow-up questions may be required, some elements may be unclear, and some complex code may be uncommented.
- **Novice:** The project makes regular and substantial assumptions of strong background knowledge of the reader in domain or analysis, and would be largely incomprehensible to a reader outside of the field. Elements of the process may be unclear, and insufficient detail is present to be able to reproduce even basic elements without additional information from the author.
- **Way Off:** The project is written with the assumption that the reader is the writer and has all the knowledge needed, leaving important elements unexplained, expecting readers to infer the results from unclear plots or model outputs, and/or providing insufficient information for the reader to *understand* what was done, let alone to reproduce it.

5.7 Visualizations

- **Masterful:** Visualizations are easy to interpret, readable, of sufficient size and resolution, easy to follow, clearly labeled, and accessible (e.g. not using color as the sole differentiation between lines, or using colorblindness friendly color palettes). There are no plots which are unclear or deceptive, and formatting elements are consistent across plots (e.g. orange dotted lines referring to the same condition across a cluster of plots). Most importantly, plots 'tell the story', explaining what's going on in parallel with model or ML output and making the findings more easily accessible.
- **Acceptable:** Visualizations are generally clear and interpretable, and well labeled, with few areas of confusing or deceptive formatting, and visualizations have general consistency with few exceptions. Plots are used as an effective device for explaining the data, are described and associated with text, and supplement the analysis well. Some desirable visualizations are missing, but they're not crucial.
- **Novice:** Visualization is present, but often difficult to interpret, poorly labeled, and unclear. Formatting is inconsistent or misleading, plots are undersized or blurry or difficult to read, or deceptive. Plots may also feel 'tacked on' or like they don't integrate well with the discussion, or just 'happen' without further explanation. Important points are entirely lacking figures, and the story is very 'incomplete' without them.
- **Way Off:** Visualizations are absent, present only in limited parts of the project, uninterpretable in context, unreadable, or unrelated to the story.

5.8 Length and Formatting

Note: It is exceptionally hard to give a meaningful length estimate for a project which will include model output, visualizations, chunks of code, and otherwise. I encourage you to focus on how long it 'should' be given the scope, the fact that it is the culmination of a full year of work, and your own desire to show knowledge, rather than aiming for an arbitrary guideline. Note also that larger group projects will be expected to have greater scope and length and depth than individual projects, so plan accordingly.

- **Masterful:** The project is sufficiently long to demonstrate knowledge, show the richness of the work, and show proper scope, but without dragging or feeling like words are being added for the sake of talking. Formatting is reasonable, readable, and enhances the text. Citations, where given, are reasonably formatted according to an accepted standard, and contain the information needed to be followable. Additionally, all analyses and plots and model outputs are explained in sufficient depth.
- **Acceptable:** Project length is sufficient, although perhaps a bit too brief or too wordy. Formatting is acceptable, if occasionally distracting or adding difficulties. Not all analyses or plots are captioned and described, but this doesn't hinder the comprehension of the work.
- **Novice:** Project is 'a bit too short' to accomplish the needed demonstration of knowledge, or buries lack of knowledge in a mountain of text. This could also apply if the formatting is distracting, problematic, or hurtful to the argumentation.
- **Way Off:** The paper is so short as to be unable to address the topic well, appears 'last minute' or 'low effort', or is not given enough space to demonstrate knowledge. This could also apply if the paper is presented without formatting or sectioning, citations are missing crucial information leading to inability to confirm sources, or if the project fails to include sufficient explanatory text and is just a bucket of model outputs and graphs.

5.9 Structure and Organization

- **Masterful:** The project is well organized, with the introduction discussing the structure of the project to guide the reader. Sections are numbered and cross-referenced throughout the paper where needed. Transitions between sections are clear, and the sections make sense.
- **Acceptable:** Structure of the project is acceptable, if somewhat difficult to understand or poorly explained in a few places. The reader has a clear understanding what's going on, even if it's not always fully explained.

- **Novice:** The project's structure is counterintuitive and not understandable, with the structure (or lack thereof) hurting the argumentation considerably.
- **Way Off:** The project has no discernable structure nor organization.

5.10 Language and Argumentation

- **Masterful:** Student language use is consistently clear and understandable, with sufficient formality for the task, but while avoiding obtuse, over-wordy academic writing. Additionally, the project was clearly proofread, with typoes and English language grammar issues relatively rare and not considerably affecting the argumentation.
- **Acceptable:** Student language use is largely understandable, but in places unclear, or with a sufficient number of typoes or English language grammar issues as to start to hinder understanding of the argumentation.
- **Novice:** The argumentation is regularly unclear, and our ability to understand the content of the paper is considerably hindered by typoes, difficult to understand argumentation, or English language grammar issues.
- **Way Off:** The project is not written in English, is largely incomprehensible, or shows evidence of automatic machine translation or unedited Large Language Model use (e.g. 'hallucinated' data or information)

5.11 Academic Integrity and Citation

Note that plagiarism or other academic integrity issues will result in an automatic 'Non-Satisfactory' on the project, as well as potential suspension or dismissal from the program.

- **Masterful:** All direct quotes or external references are cited, with the source given clearly in a 'Works Cited' section, in the format of your choice (APA is a fine default) including the necessary information to track down the resource. All other words are your own. External data sources are clearly discussed, with links (where possible) and date of access.
- **Acceptable:** All direct quotes or external references are cited, but the citations do not include enough information for the reader to locate one or two original resources.
- **Way Off:** Sources are not cited, or there is evidence of academic dishonesty. Be careful to cite all your sources and quotes.

6 Project Grading and Rewrites

Your capstone project will be graded, as discussed, by both your CSS Faculty Advisor and a Designated Reader, during Week 8 of the quarter.

6.1 Revisions and Rewrites

Following grading, there are four possible outcomes:

- **Satisfactory**
  - This means that your project has been deemed 'Satisfactory', and if any minor revisions were suggested, you do not need to resubmit after completing them, and they're given only for you to strengthen the work as a writing sample.
  - At this point, your capstone project is complete, pending your presentation.
- **Satisfactory pending revisions**
  - This means that your project is quite nearly ''Satisfactory", but there are some more more substantial revisions to elements of the project which need to be completed
  - These required revisions will be detailed when the project is returned to you
- **Unsatisfactory, rewrite needed**
  - This means that your project requires substantial re-writing, expansion, or reworking to reach the bar of 'Satisfactory'
  - Your Advisor and Reader will provide guidance as what should stay from the existing paper, and which elements need to be re-written, added, or removed, and where the deficiencies exist in the current project
- **Unsatisfactory**
  - 'Unsatisfactory' is awarded in the event that student(s):
    * Do not submit a project by the expected deadline without good cause
    * Have not engaged with the capstone project tasks or process
    * Do not produce a 'Satisfactory' project at the end of the revision/rewrite process

In the event of a **Satisfactory pending revisions** or **Unsatisfactory, rewrite needed**, you will be expected resubmit your project to your Advisor and Reader with these revisions completed **before Finals Week**, at which point they'll evaluate whether further revisions are needed, or whether the project is now 'Satisfactory'.

In the event of an intractable disagreement between the CSS Faculty Advisor and Capstone Reader as to the nature of a paper, a member of the CSS Educational Committee or the CSS Director can serve as a third, tie-breaking vote.

6.2 Group Grading

In group projects, **all group members will be assigned the same grade**, and as such, all group members are responsible for ensuring that any revisions or rewrites are concluded and submitted, and that the names on the final project reflect the work of the students who participated.

In the event that a group member is not participating adequately in the capstone process, remaining group members should work with their CSS Faculty Advisor to address the issue, and if necessary, work with both the CSS Director and their advisor to ensure that assigned grades reflect the work performed.

7 Capstone Project Presentations

As a formal end to the CSS M.S. Program, you'll attend our Capstone Project Presentation event. During this event, each Project Group (or individual) will give a brief, 10 minute overview of their project. During this, you'll

want to describe…

- What questions motivated your work
- What data you analyzed
- The basic method of data analysis you used
- Your most interesting or revealing findings, discussed in the context of your questions
  - This is a great opportunity to show off your favorite visualizations!
- What you learned from your project, and how your capstone host benefited from the work

For this overview, you're welcome to use slides, project a poster, or show interactive plots you've generated in a notebook. *We'd like to hear from each group member, so please plan transition points where one person will hand the talk over to the next, who will continue.*

Although there will be a moment for questions, for more detailed discussion, you can use the mixer period following the presentations to talk with your colleagues, faculty, and capstone hosts.

All students are expected to attend and to present their capstone at the event, barring good-cause difficulties. In the event that you cannot attend, please contact the CSS Advising Team ASAP to find alternate ways to complete this requirement.

8 Specific Guidance for Capstone Hosts

Thank you for being a Capstone Host for UCSD's CSS Program!

Your role as a capstone host is, in large part, to help the students gain experience working in the field, as well as in *your* field. As such, your main roles are…

- To regularly meet with students, ideally weekly, to ensure that they're making progress, addressing good questions, offering you the support you need, and getting the support they need.
- To show the students some of the problems or questions that you're facing (as discussed in your proposal)
- To propose any specific questions which you'd like answered, or point students towards rich and interesting areas of questioning
- To give students basic guidance on how to attack those problems (e.g. using what data)
- To ensure that the students are trained as to any specific confidentiality, data privacy, intellectual property, HIPAA, FERPA, or IRB complications
- To review the final project to ensure that the contents of the writeup, code, and any released data are able to be made public with no concerns
- To give students the background they need to succeed, whether in terms of understanding the organization, understanding the area of study, understanding the concepts and methods underlying the research, and otherwise

In addition, you'll also be a point of contact about the project for the CSS Administration, and help us keep an eye on the process. Please contact the CSS Director if…

- You've not heard from one or more student(s) in more than two weeks without prior agreement or understanding
- There are other, fundamental issues which are hampering the students' ability to succeed
- You're experiencing difficulties or problems in your interactions with your students
- At any point, you have a student doing 'unsatisfactory' work and not responding to your guidance to get them back on track

9 Specific Guidance for CSS Faculty Advisors

This section is being written primarily for CSS Faculty Advisors for Off-Campus and Non-CSS-Affiliate-Hosted Projects. For CSS Affiliate Directed Capstones, CSS Faculty Advisors will play the role both of Capstone Host and Advisor, and can combine these duties as they find optimal.

Thank you for being a CSS Faculty Advisor for UCSD's CSS Program!

As a CSS Faculty Advisor, you'll effectively be serving as a research mentor or 'thesis advisor' for these students, but with the key distinction that you'll be less focused on guiding the questions, data, and work (as the majority of this will be handled by the Capstone Host), and you'll meet less regularly with the students than a thesis advisor might. At risk of defining a position you know well, some of the expected roles would be…

- To regularly meet with students, at least once per month, and as needed
- To interact with the Capstone Host where it's useful (e.g. to help find the best project approach balancing the needs of the host and the needs of the students)
- To keep track of the academic, project side of the capstone and ensure that the work the students are doing could lead towards a workable final project
- To ensure that students are testing specific social science hypotheses or addressing social science tasks, whether as part of, alongside, or in addition to their Capstone Host assigned work
- To help guide students as they face methodological or conceptual difficulties in the analytical process
- To ensure that the project is rigorous, well-conceived, demonstrates their learning, and is able to pass muster with the CSS program and to lead to a 'Satisfactory' project grade
- To help students identify another CSS Affiliated Faculty Member to serve as the Capstone Reader who has relevant expertise, or who can fairly judge the paper, and then ask students to meet with the proposed reader to confirm.

Additionally, you'll handle student evaluation by…

- **In Fall Quarter**, where the CSS Director is instructor of record for CSS 296, you'll give your recommendation (S/U) to them rather than directly entering it in eGrades
- **In Winter Quarter**, you'll assign a CSS 296 grade in eGrades for each student (in consultation with the project hosts), again using Satisfactory/Unsatisfactory
- **In Spring Quarter**, you'll handle the grading of the CSS Final Project, as described above in 'Project Grading and Rewrites', and enter the resulting Satisfactory/Unsatisfactory grades for CSS 296 grade into eGrades.

And finally, you're asked to reach out to the CSS Director with any concerns, fundamental issues, absentee students, or problems with the ongoing project.

Although you have been given many guidelines here around timeline and procedure, ultimately, you're trusted to work with the students to ensure the best outcome, and we encourage compassion. Please work with the Capstone Reader in the event of any timeline changes for project grading and submission, and if you feel major departures from the plan are what's best for the student, please work with the CSS Director as well.

10 Specific Guidance for Capstone Readers

Thank you for being a Capstone Reader for the CSS Program!

Although you are welcome to participate in the project to whatever extent sparks joy, meeting with students, reading drafts, or discussing the work, your only required roles are to consent to being a reader in Fall, and to evaluate the paper in Spring.

As such, your required work should be relatively minimal, and confined to Weeks 8-10 of Spring Quarter. During this time, you'll work with the student(s) CSS Faculty Advisor to evaluate the students' final project according to

the guidelines above in ‘Project Grading and Rewrites’.

Your role here is to serve primarily as a second reviewer, to ensure that the project ‘passes the bar’ for the CSS program. You’re encouraged to work directly with the students’ CSS faculty advisor and collaborate to come to a conclusion about the papers, as well as to ensure that you both can accommodate any deviations from the expected schedule.