Stat 240: Teaching and Grading the Project

Fall 2023 – Compiled by Cameron Jones

This document contains all the information that Stat 240 TAs should know to effectively teach and grade the group project administered over the second half of the semester. Any guidance you receive from the professors or in official course communications/announcements takes higher precedence over this document.

Personally, I find that guiding students through their projects is a genuinely fun and rewarding experience, and students seem to relish the opportunity to use their new knowledge to answer a real question about something they care about.

I find it is best to be totally transparent and honest with students on what you expect from them and how you will be grading, so feel free to tell them anything in this document, even outside of the “Present to Students” section!

# Present to Students

The information in this section will be helpful to present to your students at the beginning of discussion this week, as they start to work on their project proposals.

Students will complete the project **in the discussion groups they are currently in**; this should be their second week in the new groups. If you have not switched them to their new post-midterm groups yet, you should do so this week.

## This Week

In discussion sections this week only (week of Monday 10/30), students will NOT be doing a normal discussion assignment on lecture content, they will have **time to begin working on their project proposal with their groups**. Most, if not all, future discussions will return to the normal discussion assignments on lecture content.

To receive credit for the day, they should submit some link, document, or general evidence that they did really think about/work on the proposal. This will be graded for completion; any submission is worth full credit for all attending members.

## The Project Proposal

* The full requirements for the project proposal are enumerated in the Project Information page on Canvas.
  + Students should reference that list and **make sure they have everything before submitting, since that is what we’ll use for grading.**
  + That Project Information page is the definitive source for instructions and rubrics.
* Selecting a dataset and question of interest does **NOT lock them into it** for the rest of the project.
  + Especially if you have concerns about their question or dataset (see: Grading the Project/Proposal), many groups will end up switching their question, dataset, or even their topic entirely.
* **Groups should pick a topic they care about! Encourage creativity!** Not only will it be more fun for them, but they will get more out of it and produce better content if they are enthusiastic about the topic.
* Finally, **projects should be fundamentally solid rather than complex.**
  + Many groups “go big”, aiming to answer a lot of questions or a super broad question (see Grading the Project/Proposal). I highly discourage this.
  + They should **pick one, maybe two questions of interest, which aren’t overly complicated**, so they can comfortably demonstrate what they know.
    - This also ensures all members can contribute to and benefit from the project.
    - Sometimes one particularly experienced group member will try to drag the project towards something out of the scope of this course – like logistic regression or time series analysis. This benefits nobody. We require groups to stay within the scope of the course for this reason.

# The Rest of the Project

No need to overwhelm students yet with information about the later stages of the project – I leave this here for your reference as the project goes on. Feel free to present this information to your students as we enter each new stage of the project if you see fit.

## Overall Project Structure

The final project is a “three and a half” step process for students.

* The **project proposal**, as described above, is due Wednesday, 11/8.
* The **project draft** is a first draft at a comprehensive report, using and writing properly about statistical inference to answer their question of interest – due Wednesday, 11/29.
* After those drafts are submitted, *each student individually, not as a group,* will **peer review** a randomly chosen draft from another group, due Wednesday 12/6 (this is the half step).
* Finally, the **final report** asks groups to incorporate the feedback from their TA and their peers into their draft, and is due Wednesday, 12/13.

For every stage of the project, definitive instructions can be found on the Project Information page.

## The Draft

There are three important ideas behind the project draft, which is their first attempt at a full report.

* The files submitted and the description in the report should enable someone unfamiliar with the general topic and the dataset (for example, you, the TA!) to fully **understand and replicate**their analysis.
* The report should have a **clear and consistent story**- everything should be written with purposeful reference to the question of interest. Anything else is confusing clutter that should be removed.
* There should be a very sharp distinction between statistical jargon and real-life interpretations.
  + The **Analysis section should ONLY contain statistical jargon** (“p-value”, “binomial distribution”, anything a non-statistician wouldn’t know) and no real-life interpretation.
  + The **other sections should ONLY contain real-life interpretations** (to be accessible to a previously unfamiliar reader.)

## The Peer Review

Each student *individually* will be assigned to review another group’s project draft. They do not assign meaningful points, just provide comments. We encourage students to give specific feedback that they would be happy to receive on their own project. More specific instructions found on the Project Information page, there are five specific sections we ask students to review. Every group will receive three or four students’ feedback.

## The Final Report

The instructions for the final draft are exactly the same as the project draft; we just ask that groups make changes according to your feedback and peer review feedback on the first draft. If a group has a great first draft, they won’t have to update much for the final draft. Other groups will have a lot of work to do.

Groups should incorporate all the TA’s feedback; they do not necessarily have to use all the peer review feedback.

# Grading the Project

Your feedback as a TA, especially on the proposal and draft, is critical to students learning from and succeeding on the project. As such, be prepared to spend more time grading these assignments than you do on the weekly homeworks. For example, in the past as a 50% TA with three sections of roughly six groups each, I usually take 10-15 minutes on each proposal/draft – for a total of three to four hours.

Above all, you should be consistent in how you grade. Consistency is key for both fairness and for being able to tell students why they specifically got points off for a specific section.

Students like to squabble for points on the project stuff since they’re more subjective, but you can just tell them you won’t be regrading out of fairness to other groups. (Unless of course you factually missed something or did something incorrect, in which case you should totally give them the points back!)

## Grading: Proposal

The proposals are due Wednesday, 11/8; be aware that you are expected to talk to each group about your feedback in the discussions the next week, the week of Monday 11/13.  
  
 Following the guidelines below I usually end up giving out an average in the low-mid 80s on the proposal, not giving out any perfect scores – maybe one or two. Deviations to be expected, but if your average is in the low 70s or high 90s you might’ve been too harsh/lenient.

### Step 1: Project Info Rubric

The first step in grading a proposal is to **reference the list of expectations on the Project Information page**. Take a small amount of points off if things are missing or not fully described.

Be **lenient with their description of “a possible model** and associated methods of inference or prediction” since they haven’t really learned it yet. Any points you take off here should be due to loose definition of their variables of interest. They should identify if their variable(s) of interest are continuous or categorical, and which (if two variables) they are treating as the cause and which they are treating as an effect. (See more below

Here's a helpful reference for your students for what method they should use if they have two variables; these are fair things to tell your students to guide them in choosing a method they have not yet been taught.

|  |  |  |  |
| --- | --- | --- | --- |
| Bivariate Inference Methods | | Outcome | |
| Binary | Continuous |
| Cause | Binary | Two-Sample Proportions | Two-Sample Means |
| Continuous | ~~Logistic Regression~~ (OUT OF SCOPE) | Linear Regression |

### Step 2: Question of Interest

The next and most important part is setting the group up for success by giving specific feedback about their question of interest. Since the question of interest guides their entire analysis, giving good feedback on the structure of the question itself is critical.

* **Don’t be afraid to take off multiple points for this**. This draws their attention to it and emphasizes how important the question of interest is. I often give out 1/4 or 2/4 in the question of interest section.
* There are a few ways a question of interest can go wrong:
  + It can be too simple. I call these “What’s the biggest number” questions that you can answer yourself just by looking at a table, there is no inference or prediction involved.
    - For example: “Which states had the most deaths per capita from COVID?”
  + It can be too complex. This can happen by involving more than two variables, or requiring methods outside the scope of this course.
    - For example: “What sociodemographic factors cause COVID?”
    - Or: “How does age (continuous) affect the likelihood of dying from COVID?” (Logistic regression, out of scope)
  + Groups can define it too loosely. It should be extremely clear from the wording of the question what the one or two variables of interest are, and if they are binary or continuous.
    - For example: “What is the relationship between position on the field and height among soccer players?”
      * Exactly how is “position on the field” defined? Binary? Continuous? What does that mean?

### Step 3: Dataset

Finally, they should have **specific and detailed knowledge/documentation of their dataset**. They should be able to tell you exactly what variables they’re interested in, what units those variables are in (e.g. “height” is not enough, needs units) – and everything else listed in the Project Information. Most importantly, they absolutely must answer **What is a row of the dataset in real terms?** This question is critical to understanding how they will write code for their analysis, and I usually take off at least a point if not two for failing to answer this question.

* For example, if their only dataset description is “We have a dataset of monthly temperature measurements”, I would take multiple points off and leave a comment like: “Are they Celsius or Fahrenheit, are they averages/maximums/minimums, over what time period were they taken, where are they taken from, how many measurements do you have?”

## Grading: Draft

Historically I have given out an average in the mid 80s on the project draft. Once again, deviations are okay, but it shouldn’t be too high. In particular, the threshold for a perfect score is very high, and even if you give a perfect score you should still prompt with some feedback so the group has something to chew on in the final stage.

Your feedback at this stage needs to be thorough. This is because the final report is graded based on how they respond to your draft feedback – so if you are loose with your feedback, they won’t have much to go on when submitting the final report. If you didn’t penalize a specific thing in the draft and they leave it as it was, it’s unreasonable to then take off for it in the final report.

In each section, full points are only appropriate if there is really nothing wrong with the section. This is a higher threshold than we use for grading homeworks, where minor errors are still tolerated when giving full points. If there are aesthetic changes that should be made to graphs, unclear writing, or incomplete documentation or analysis, don’t be afraid to take a point off. Taking multiple points off in a section should be reserved for fundamental errors or vital components entirely missing.

Just like the proposal and all stages of the project, you should **first reference the Project Information list of expectations** and take small points off for anything missing or not fully described.

After that, keep in mind the big ideas I mentioned in my [Draft description](#_The_Draft) above. **Don’t be afraid to take off points for those big ideas, beyond what’s in the Project Information list**. In particular, students like to fill their reports with fancy-seeming buzzwords and graphs that don’t meaningfully answer the question of interest. While this isn’t on the list, it betrays the idea of a clear and consistent story. I often tell my students **a long report does not mean a good report! In fact, the best reports are relatively short**. They should be fundamentally solid, rather than complex.

The whole thing should be professional in nature. Code should be invisible in the final report, writing should be sharp and concise, graphs should have appropriate titles and axes labels (this can be half a point) and be included very purposefully. If there is a graph, there should always be an accompanying sentence or two explaining what it means and why it’s important.

## Grading: Peer Review

The peer review is not just a completion grade – we want students to show some effort and give thoughtful feedback. They will be asked to give specific feedback in each of five sections. Feedback such as “better graphs” or “more detail” or “this is good” should not receive full credit. It’s alright if students genuinely can’t find much wrong with one or two sections.

**Technical Instructions to grade Peer Review**  
 Open SpeedGrader from the Peer Review assignment page in one tab.

Now, in another tab navigate to the *Project Draft* assignment page. Access the “Peer Reviews” subpage in the top right, shown below.  
  
A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated  
The below screenshot shows what the Peer Reviews page will look like. The highlighted text refers to the fact that Malak was assigned to review Eric Lu’s project submission. Check your Peer Review SpeedGrader tab for who’s peer review you are grading – since there are multiple TAs in each section, you won’t necessarily have all students in this list.

If, for example, Malak is your first student to grade, click on “Lu, Eric” – this will be a link to Eric’s group’s submission.

Once you are there, your student should have a comment there with a peer review. *If you can’t find a comment, check “Show Rubric” in the top right, and then “Show Assessment By:”; sometimes students incorrectly leave it there.*

## A screenshot of a computer Description automatically generatedGrading: Final Report

I grade the final reports split-screen with my draft feedback in the other tab, because as I mentioned in [Grading the Draft](#_Grading:_Draft), the grade you give in the final report should be based on how they incorporate your draft feedback. It’s still important to justify why you are taking points off, but comprehensive written feedback is not as crucial here; most students will just see their grade and accept it and not pay a ton of attention to your comments.

You can still feel free to take off small points for the big ideas, whether it’s sloppy writing/structure or having statistical jargon where they shouldn’t, but points should be mostly based if they successfully adjusted for your draft feedback. Don’t hand out perfect scores like candy, but they’re more common on the final report.

Historically my average assigned grade has been in the low to mid 90s for the final report.