

# The S-043/Stat-151 Final Project Handbook

This has all the information you need regarding your final projects.

Last updated: Wednesday, October 4, 2023

<b>1. OVERVIEW OF FINAL PROJECT.....</b>	<b>3</b>
LOGISTICS .....	3
FINAL PROJECT COMPONENTS.....	3
FRAMING AND GOAL OF PROJECT .....	4
GRADING .....	4
TWO AND ONE HALF TYPES OF FINAL PROJECTS .....	4
SOME SAMPLE FINAL PROJECT TITLES .....	4
OVERVIEW OF CROSS PROJECT INTERACTION .....	5
<b>2. FINAL PROJECT <i>PRELIMINARY</i> PROPOSAL INSTRUCTIONS.....</b>	<b>6</b>
<b>3. FINAL PROJECT PROPOSAL INSTRUCTIONS .....</b>	<b>7</b>
PART 1 (THE PROPOSAL).....	7
PART 2 (THE MATHEMATICAL MODEL) .....	8
PROPOSAL REVIEW .....	8
RUBRIC (10 POINTS) .....	8
<b>4. FINAL PROJECT EDA (PRELIMINARY RESULTS) ASSIGNMENT .....</b>	<b>10</b>
INTRODUCTION .....	10
THE WORK .....	10
THE REPORT .....	10
FAQs.....	11
<b>5. FINAL PROJECT PRESENTATION .....</b>	<b>12</b>
PEER FEEDBACK .....	12
SLIDE SUGGESTIONS.....	12
EIGHT MANDATES (SOME DOS AND DON'TS).....	13
GRADING CRITERIA.....	14
<b>6. FINAL TECHNICAL REPORT.....</b>	<b>15</b>
PRIMARY DOCUMENT.....	15
SUPPLEMENTARY APPENDICES (OPTIONAL) .....	15
R CODE SUPPLEMENT.....	16
EIGHT MANDATES (SOME DOS AND DON'TS).....	16
GRADING CRITERIA.....	18
<b>7. FINAL PROJECT GRADING RUBRIC .....</b>	<b>19</b>
NOTE ON CALIBRATING THE FINAL GRADES DUE TO PROJECT COMPLEXITY .....	20
<b>8. FORMS AND SHEETS .....</b>	<b>21</b>
PROJECT TEAM STATEMENT OF CONTRIBUTION .....	22

**ALERT!**

**This is a working document. We will add and change things as needed, but the essential elements are currently in place. We will announce any changes, of course.**

## 1. Overview of final project

For the final project, you'll have the opportunity to conduct an original analysis using the methods and tools you've learned in the class. Projects can be actual data analyses using multi-level models, or can be methodological, exploring the properties of the models and estimators that we've covered. We encourage you to work with your own data. If you cannot find data, we suggest you join up with someone who has data to form a group; the likelihood of an interesting project is greatly increased if it is built on an existing research agenda. If you have an idea and want to know if it's feasible, talk to the teaching team as soon as possible.

The two goals of the final project are to

- 1) use the skills you learned in this class to answer an interesting question and
- 2) demonstrate to the teaching team that you've actually learned those skills.

As a result, you should probably choose a project that uses multi-level models in an interesting way. If the final analysis is a model which uses random intercepts to account for the fact that students are nested in schools, we're not going to be impressed, no matter how interesting the substantive question. On the other hand, if your final project allows for more complex models, and then you subsequently discover (and motivate) that this simpler model is the most appropriate, you will be in fine shape.

### Logistics

The final project can be completed in groups of anywhere from one to three people. **We strongly recommend two or three people as an ideal.** All group members are expected to contribute equally and will be given a single grade. We expect that, just in an actual publication, everyone in a group will understand and contribute to generating the entire project. It's fine if different people take the lead on different parts of the project, but we do not want individuals to be solely responsible for any of the parts. Each group will submit a single copy of each project component.

### Final Project Components

Leading up to the final project are a few components

- A preliminary proposal
- A final project proposal (and plan)
- A midterm report with some Exploratory Data Analysis

The final project itself will have the following components:

- A final presentation (and slides for said)
- A technical memo describing your analysis that includes the mathematical model and description of relevant parameters
- Supplementary appendices of NO MORE THAN 10 pages (Optional)
- The R code, neatly formatted.

## Framing and Goal of Project

The final project will take you to a point where you should have work ready to submit to a conference.

## Grading

Final project grades will be based on all of the above components. We will post an overall rubric.

Further details on these components can be found below. Groups of two or more will be asked to sign a statement of contribution that outlines what each member did. You will submit a single form with all signatures of group members.

## Two and One Half Types of Final Projects

There are different types of final projects you might choose to do. Here is a list of the broad categories:

**Analysis of Real Data:** We expect this to be the most common form of final project. In such a project you analyze an actual data set because you are trying to answer a specific science-related research question. These final projects often will be part of a student's dissertation research or Masters thesis. It is perfectly fine in this case to have other people working on the project as far as we are concerned; you should check with your advisor to make sure that it's okay on the other end.

**Simulation study:** Here you would study the performance of some sort of methodology under a variety of different hypothetical, yet hopefully plausible, scenarios. You might compare the performance of multilevel modeling to robust standard errors in small samples, for example.

**Something else:** Here you might consider a hybrid of a simulation study and an analysis of real data to, for example, demonstrate that your approach works in your context. This might add legitimacy to your actual analysis. You might also decide to work on some mathematical properties of these methods. If you propose it, and it's relevant to the course, it's quite likely we will accept it as a valid final project.

## Some sample final project titles

The following were final projects from a previous year:

- Hospital Costs and Insurer Concentration
- Happy Accidents: Ethnic-Racial Identity Intervention Increases Adolescents' Affirmation
- The persistent effects of controversial congressional votes on vote share
- Association of district-level pesticide use and BMI in India
- Examining a universal intervention program's effect on increasing ethnic-racial identity affirmation
- Beverage Perceptions among a Longitudinal Cohort
- Predictors of Peak and Longevity in Professional Tennis
- Nesting' in child growth : Is it social or geographical?
- Investigating Matthew effects for a second grade literacy intervention focusing on vocabulary
- Civil Conflict and Children's Development in Colombia
- External and Internal Financing During Credit Shocks
- Explaining Associations Between Cellular, Broadband, Internet and the Human Development Index through Multilevel Modeling
- Estimation of subglottal pressure from neck-surface vibration in patients with voice disorders
- Analyzing Violations of Multilevel Model Assumptions

Discrimination and Psychological Distress: A Longitudinal Study  
Effect of real-time gaze visualizations on remote learning outcomes  
Is 250:1 the Magic Number? James Bryant Conant's Counseling Ratio and its Relevance for Policy Today  
Drivers of App Usage in Early Childhood  
Women, Business, and the Law: Uncovering Relationships Between Government Support of Women's Rights and Women's Economic Opportunities  
Multilevel Predictors of Anemia Across Ecologically Distinct Regions in Madagascar  
Examining Refugee-Host Relations in Uganda

### Overview of Cross Project Interaction

As part of the final projects you will get to see and respond to other students' work. Here are the currently planned moments

- 1) **Proposal Review:**  
Everyone will get a random handful of the preliminary project proposals to read once submitted, and everyone will give feedback on those proposals.
- 2) **EDA Workshop**  
People will bring preliminary analyses and some exploratory data analysis to a workshop near the end of term. People will be grouped to workshop said preliminary results.
- 3) **Peer Review.**  
You will write feedback on some of the presentations you watch at the end of term.

## 2. Final Project *Preliminary* Proposal Instructions

### **NOTE: THIS IS A VERY LOW STAKES ASSIGNMENT**

The final project preliminary proposal is a description of a project idea that has to do with multilevel modeling. It is a way to give the teaching team an understanding of what you might do for your project. This will allow us to give you feedback on the feasibility and appropriateness of your project before you go too far down the road. You don't have to know exactly what you'll be doing, but this preliminary proposal should give us a clear understanding of how your project will use concepts from the course.

#### **What to do:**

- 1) Read this document
- 2) Talk with friends and form collaborations, if desired
- 3) Write a single page telling us what you are heading towards (see below for details)

**Group members:** Your proposal can be submitted as an individual or as a group (up to three members). Your group members can be tentative, and there will be further opportunities to join other groups or add members to your group after the proposals are in.

**Proposal Review:** The teaching team will read and comment on your proposals.

**Fake Final Project Proposals:** It is perfectly acceptable to come up with an idea and submit it as your final project proposal and then bail on it if you want to work on something else instead, for example if you discover someone doing something that excites you more.

#### **Contents of the Proposal:**

**Form of proposal:** Your final project proposal should be one or two paragraphs – certainly no more than a single printed page. It should be a friendly and accessible text that covers at least the following:

1. A (possibly vague) research question or area of exploration.
2. A brief description of the data being used (if any). This should include sample size and number of covariates (even if approximate).
3. Why the project is likely to be interesting and worthwhile.
4. Why it is connected to multilevel or longitudinal modeling (if this is not obvious).
5. Whether you are open to having people join you on your project (and any human subjects issues with that, given data).

**Note on Data:** If you intend to do an analysis of actual data, you should specify the data source, even if you don't yet have access to the data. If you don't have the data, you should explain how you intend to get access. You should check with any relevant institutions or supervisors about using data.

You should also include any other details you feel are relevant.

### 3. Final Project Proposal Instructions

The final project proposal is a description of a project idea that has to do with multilevel modeling. It is a way to share your ideas with other people in the class as well as to give the teaching team an understanding of your current thinking with regards to your project. This will allow us to give you feedback on the feasibility and appropriateness of your project before you go too far down the road. Your proposal should also give us a clear understanding of how your project will use concepts from the course.

In your proposal you will also present an initial mathematical model, which will help ensure that your research questions are concretely embedded in the statistical tools of the course.

**Group members:** Your proposal should be submitted as a final project group (from 1 to 3 members).

**Form of proposal:** Your proposal should have two parts, the overall proposal part and the mathematical model part. It should be no more than 3 single-spaced pages (2 pages for part 1, 1 page for part 2) with a 12 point font size.

#### Part 1 (the Proposal)

The proposal should be a friendly and accessible text that covers the following. You can mark each section with a bold header (e.g., “**Objective:** etc etc etc”), or make it more narrative if that feels clearer to you.

1. **Objective:** What your overarching goals are.
2. **Data:** If you intend to do an analysis of actual data, you should specify the data source, even if you don’t yet have access to the data. If you don’t have the data, you should explain how you intend to get access in the very near future. You should check with any relevant institutions or supervisors about using data.
3. **Relevance:** Why the project is interesting and worthwhile.
4. **Course Relevance:** Why it is connected to multilevel modeling (if this is not obvious).
5. **Outcomes:** At least one outcome you plan on investigating. (E.g., “One of my primary outcomes is income at 30 years old.”)
6. **Initial Goal:** Describe at least one relationship or statistical feature of the data that you plan on investigating. (E.g., “As a first step, I will describe how income varies by number of job training courses.”).
7. **Initial Research Question:** Propose a specific initial research question that will be part of your analysis.
8. **Simple data exploration:** Report sample size and number of clusters, timepoints, etc., if at all possible. (Guess if not yet sure.)
9. **Initial mathematical model:** See below for details on this part.

You should also include any other details you feel are relevant.

## Part 2 (the mathematical model)

In this part you present the (possibly tentative) primary model you plan to use in your analysis in neatly formatted mathematics. Give a single line of interpretation for all of the parameters that you care about (if you have groups of parameters that are all the same, e.g., 20 coefficients for 20 different level 1 control variables, you can describe them all as a group).

### Necessary Components of the model one-pager

The mathematical model (multilevel form)

Mathematical model (reduced form or lmer command)

Defined indices (e.g., student  $i$  in school  $j$ )

Short descriptions of the primary parameters of interest in your model

Short description of how you may fit different versions of your model (e.g., adding interaction terms) to answer extensions to your primary research questions

Justification of using MLM or cluster-robust approaches (depending on what you are doing). Is MLM needed for your project? Or do you just like to use it?

At this point, you can also provide other salient details you want feedback on!

## Proposal Review

The TFs will read and comment on your proposals. These proposals (without focusing on the math part) are also going to be read by your peers, so please write clearly so they can understand you! Know that people outside of your field will be trying to understand what you are up to.

Things the TFs, in particular, will be attending to:

- Is the model clearly stated?
- Is the research question in alignment with the model?
- A high level assessment of whether your proposal is likely to work out, or not, given the realities of the data and your stated purpose.

## Rubric (10 points)

Proposal (5 points)

Completeness: All parts present      2

Clarity: Easy to understand?      2



Coherence: Choices appropriate 1

(E.g., random effects model with too few observations per group.)

Math model (5 points)

Math is correct 1

Indices defined 1

Primary parameters defined 1

Both reduced and multilevel form 1

(Reduced and multilevel forms also in agreement.)

Justification of modeling approach 1

## 4. Final Project EDA (preliminary results) assignment

### Introduction

This assignment is to primarily serve as a stick to get you going on your final project. It's also going to lead up to a workshop day, where you are going to share your progress with other teams, and exchange feedback and brainstorm next steps. This feedback process mimics the real research process of getting feedback as you conduct an evaluation.

### The work

You should start exploring and analyzing your data. This means identifying some research questions, writing some general descriptions of your data, and ideally generating some preliminary models. It is better to focus on a smaller range of outcomes and research questions at this stage. Do not spread yourself too thin.

### The report

The result of this exploratory analysis is a report that consists of the following:

- 1) One paragraph overview of background and some small number of primary research questions of interest.
- 2) One paragraph describing your data sources: What is the structure of your data files (e.g., student-level, school-level, teacher-level, county-level)? Do the variables come from surveys, administrative records, classroom observations, etc.? Please feel free to discuss any data decisions you've made or are considering (e.g., you're dropping certain observations, creating new measures, or are unsure of which measures to use). **Make sure you include the sample sizes at each level** (e.g., number of time points per unit, average number of students per school, etc.).
- 3) Two **publication ready** plots of your data. *Publication ready* means having things like nicely labeled axes and captions.
- 4) An analysis of the variability in your data (primarily your outcome). You can do this by fitting an unconditional model (one without any extraneous covariates) to obtain a variance decomposition (allowing for ICC calculations and the like). These unconditional models are good starting point for getting a handle on where variation is, and what your data structure is.
- 5) One paragraph describing trends in your data, referring to your two plots and your variability analysis.
- 6) An initial mathematical model describing the primary model you are planning on fitting (or have fit). Try to focus attention on your primary covariates and outcome; keep these initial models simple and straightforward.

As you write your model, be sure to define your subscripts! I.e., at the beginning of your model write something like "For time  $t$  for student  $i$  in school  $k$  we have..."

Also, to be correct you would then subscript as  $tik$ , keeping your levels in order.

- 7) One paragraph of initial findings. If you have not yet fit a model, you can still describe preliminary findings in terms of trends, etc., in your plots and initial exploratory analysis.
- 8) One paragraph describing next steps, blocks, barriers, concerns, or other things you would like to discuss and get feedback on.

**There is a 2.5 page limit, not including the plots with captions, but including your full mathematical model.** This is single space, 12 point font with margins of at least 1 inch on all sides.

## FAQs

Q: What if I have missing data issues?

A: For a preliminary analysis, simply subsetting your data to complete cases (once you have identified your primary variables of interest) is fine. You can also start thinking about how to handle missing data in other ways.<sup>1</sup>

Q: What if my outcome is not a continuous, but is something like binary or a count?

A: For a preliminary analysis, simply treating your outcome as a continuous outcome should be fine.<sup>2</sup> We will learn methods for handling other kinds of outcomes soon.

Q: What if I don't yet have my data?

A: You should seriously consider joining another group or person and working with them. At this point, not having data is a very risky situation since it often takes longer to get data than you think.

---

<sup>1</sup> See the missing data handout in the handout textbook to get some further help on missing data.

<sup>2</sup> In fact, econometric approaches typically do this all the way through. In fact, this approach can often give more interpretable coefficients on your regression models.

## 5. Final Project Presentation

You will be assigned to one of three presentation blocks. During this block you and your teammates will give your presentation and watch other presentations.

After you give your presentation, you will talk to a TF, which will give you a chance to clarify a few points and to get extra feedback. The TF will then look through your submitted technical document to make sure everything is in order.

On the day of the final presentations, we will have a few minutes at the beginning where everyone can set up and load their slides.

Each presentation will have the following time allotment:

1 person	12 minutes + 3 for questions
2 people	15 + 3
3+ people	17 + 3

Each presentation will be timed and given a 2 minute warning and then a hard stop. After the hard stop we will have 3 minutes of questions. There will also be 2 minutes for transitions between presentations. Every group member should actively participate in the presentation. **To respect other presenters, these times are relatively hard limits. If you do not make it to the end of your presentation by the time given, that is not good for you.** A bell will ring at the end of your time. You can then, if you must, use up to 1 minute of your question time to wrap up.

Also be warned that **if something is unclear the instructor or TFs may ask brief clarifying questions during the talk.**

The room where the presentations are being held will include a projector that you can use to present slides. If you want to hand out documents, you will need to bring them yourself. Expect about 15 other people.

The schedule of which presentations are in which block will be announced in advance.

### Peer Feedback

In addition to giving a presentation, you will be actively watching other presentations. This means all students are expected to attend all the presentations in their time slot. Students will each review the other presentations, giving constructive feedback on a provided peer evaluation sheet. These will be given directly to the reviewed group. The TFs and I will also scan these sheets for further reference and review.

### Slide Suggestions

You should never have more than 1 slide per minute of presentation. We strongly recommend only a single slide for background if possible.

You should have a slide which shows your primary model, mathematically, if possible.

You might start off with an arc of: title slide, background slide, research question slide, data description slides 1 & 2 (with plots!), (general) model specification slide, results slides (table), results slide (plot), results slide (plot), conclusion slide = 10 slides!

## Eight Mandates (Some DOS and DON'TS)

### 1. *Keep your narrative clean*

DO specify research questions on the slides.

DO NOT try to squeeze too much into the presentation. Identify the crucial background and results, and focus on those.

### 2. *Clearly present your work*

DO cite your data source(s) early on, and be sure to briefly describe or show the structure of your data.

DO specifically show the models you fit, and connect the analysis to the research questions.

### 3. *Justify your choices*

DO justify your approach so that the connection between your research question and your analysis is clear.

DO explain why your model is correct (or at least a good) choice. *Justify* your choice.

### 4. *Guide your audience's attention*

DO clean any regression tables you show on your slides. Add reference lines and circle relevant numbers as helpful.

### 5. *Make your plots impactful and useful*

DO use graphical displays of your results. It is generally easier to compare things graphically than using tables.

DO ensure your displays are effective in communicating your point.

DO use thick lines on plots for slides so they are visible.

DO plot the estimated growth curves for different subgroups of interest and the overall mean, if analyzing longitudinal data.

DO NOT show a plot if it is not important. If you can't explain why you're showing it, don't.

### 6. *Make your slides easy to read*

DO check that someone from 30' away can read the figures and the bulk of text in your slides.

DO avoid pale colors and pastels; they tend to wash out.

### 7. *Avoid distracting clutter and give a polished feel*

DO NOT discuss data cleaning unless it's important (you can have one sentence of complaint)

DO NOT include any raw output on slides.

DO NOT put R code in your presentation unless you have an extremely good reason.

DO NOT rely on huge blocks of text to communicate your results. You want our attention to be on you.

DO NOT try to present an exhaustive account of every analysis you conducted, or even every analysis which made it into your final paper. The presentation is for summarizing and making sense of your results. 10 minutes is not a lot of time, and you'll need to be strategic about what you include.

8. *Polish and edit*

DO prepare an interesting and well-organized presentation.

DO practice your presentation at least once to work out any bugs and make sure it doesn't take up too much time.

### Grading criteria

See the grading rubric for how the presentation will be assessed.

## 6. Final technical report

In addition to the presentation you will turn in a PDF of your slides (NOT powerpoint, whatever) and a technical supplement, described next. This report has three parts:

The **primary document**. This document should be 7 to 10 pages, double-spaced, not including appendices, should explain your research question, how you tried to answer that question, and what your results were. This should be 1500 words or fewer, not including any supplementary material or the R code, all described below.

The **appendices**. These are optional. These can have extra stuff that you want to show off or have extra analyses that somehow demonstrate that you did the right thing in the main body of your text.

The **R code supplement**. This might be a separate document or documents. They can be raw markdown files or R scripts that, if run on your data, would generate all the output used in your report.

### Primary Document

The **primary document** will consist of three parts.

- 1) A *brief* introduction that motivates and provides your research question (this should not be the bulk of your manuscript).
- 2) A data section (if applicable) that quickly describes your data. With this include a Table 1 (or similar) describing your data.
- 3) A methods section that explains the approach you used to answer the research question.
- 4) A results section that reports the results of the analysis. This section should include
  - a. 2 figures
  - b. 1 table of results we might care about (this is *not* your Table 1 of data stats).
- 5) A conclusion that includes some identified next steps and some limitations or concerns with your work so far.
- 6) A declared word count (of the abstract, but not the code or appendices) posted prominently at the top of your manuscript, near your title and name.

The paper should not be an exhaustive account of everything you did, but should give us enough information that we could replicate your analyses. The main text should include any tables and charts of primary importance. It *supplements* your presentation, and should have some of the technical details that the presentation might not have. It should include your primary figures and tables as well (figure and table captions should not be double spaced). These figures and tables should be in the text, near where they are being discussed, and *not* at the end.

### Supplementary Appendices (Optional)

These are optional. These can have extra stuff that you want to show off or contain extra analyses that somehow demonstrate that you did the right thing in the main body of your text. **No more than 10 pages**. We will penalize appendices that do not seem well organized or well thought out, or seem to primarily serve to intimidate or overwhelm with output.

These appendices are for supplementary analyses, tables, and charts that further support your primary claims. For example, you might have a primary model displayed in the text, and then mention that four related models with different covariates (in appendix) show similar findings. You might also put residual plots here, referring to them in the text. Finally, you might put an alternate analysis here that you can't bear to cut. In this case you would mention the overall finding in the primary report, such as "We determined our final quadratic growth model with an iterative model-fitting process, as described in Appendix A. We found the quadratic growth model had substantially improved fit as compared to a linear growth model."

## R Code Supplement

This might be a separate document or documents. They can be raw markdown files or R scripts that, if run on your data, would generate all the output used in your report.

Please have clearly formatted and ideally somewhat commented code. This does *not* mean comment every line of code. Do think about putting your code in sections, with commentary as to what each section does. E.g., "The following code makes all of my factor variables that I use" or "Missing data imputation block."

For a reference for the more extreme end of commenting, consider the typical documents the instructor has supplied as examples for lectures. Alternatively, if your report is a Markdown file, simply submit the raw markdown file.

## Eight Mandates (Some DOS and DON'TS)

### 1. *Keep your analysis direct and to the point*

DO keep an eye on your research questions: only do analyses that further those goals.

DO NOT fit a million models and try to discuss them all. Pick a few core models, and use the rest as sensitivity checks in your appendix.

### 2. *Keep your narrative clean*

DO give some context so we understand why this is an interesting research question.

DO clearly specify your research questions.

DO NOT try to squeeze too much into your abstract. Identify the crucial background and results, and focus on those.

DO NOT write your report as an adventure story describing the series of steps you took in your analysis. You want to provide a coherent description of your final analysis.

That being said, DO NOT hide any "fishing expeditions" where you try a series of models until you get one with an answer you like. Circumvent this by acknowledging in a sentence or two other models or paths taken in your report.

DO spend most of the paper on the methods and results.

DO be explicit about the models you're fitting and how they relate to the question.



DO NOT throw everything into your primary report, regulating all your tables and charts to the appendix to make as much space as possible, so as to get more space.

3. *Clearly present your work*

DO cite your data source(s) early on, and be sure to briefly describe or show the structure of your data.

DO specifically show the models you fit, and connect the analysis to the research questions.

DO be precise in the language you use to describe your results.

DO NOT give the impression that you've estimated a causal effect when in fact you haven't.

4. *Justify your choices*

DO justify your approach so that the connection between your research question and your analysis is clear.

DO explain why your model is correct (or at least a good) choice. *Justify* your choice.

5. *Guide your audience's attention*

DO clean any regression tables you show and remove unnecessary clutter from displayed output.

6. *Make your plots impactful and useful*

DO use graphical displays of your results. It is generally easier to compare things graphically than using tables.

DO ensure your displays are effective in communicating your point.

DO plot the estimated growth curves for different subgroups of interest and the overall mean, if analyzing longitudinal data.

DO NOT show a plot if it is not important. If you can't explain why you're showing it, don't.

DO put your plot in the main text, so it is easy to connect to your narrative.

7. *Avoid distracting clutter and give a polished feel*

DO NOT discuss data cleaning unless it's important (you can have one sentence of complaint)

DO NOT include any raw output in your report.

DO NOT put R code in your primary report unless you have an extremely good reason.

8. *Polish and edit*

DO prepare an interesting and well-organized document.

DO have someone read it, if possible, for clarity, before you submit.

### Grading criteria

Your technical report is a supplement to your presentation. See the grading rubric for how it will be assessed.

## 7. Final Project Grading Rubric

We will grade the final report, final presentation, and final project together. The grades for the final presentation and report are focused on communication and will be used to assess the project itself. Including the initial proposal and EDA, the final project grade represents 25% of your final grade.

<i>Category</i>	<i>Points</i>	
<b>Presentation of Work</b>	<b>___ out of 12</b>	
Introduction	1	Is your introduction brief and clear? Are your research questions present and easy to understand?
Data	1	Is your data clearly described?
Methods	2	
- Clarity		Is your approach clear?
- Model selection		Do you motivate & explain why you selected the models you did?
Results	3	
- Clarity		Are your materials easy to digest, are you tables clear, etc.?
- Use of plots		Are your plots clearly labeled, attached to a purpose, easy to read?
- Interpretation		Connected results to substance without implying false causality?
Conclusion	1	Reasonable sensemaking of results. Making a case without overselling.
Writing	1	Is your report easy to read, not full of grammatical errors, etc.?
Math 1-Pager	2	Is your math 1 pager clear and correct?
R Code	1	Is your submitted R code easy to examine?
<b>The Work Itself</b>	<b>___ out of 8</b>	
Transparency	1	Can we easily understand what you did?
Consistency	1	Is discussion of your methods, results, and conclusions generally consistent with what you actually did (your R code)?
Correctness	1	Is what you did correct?
Appropriateness	1	Did you use approaches appropriate for your research question?
Modeling	1	Did you design and select a model with reasonable care and thought?
Challenge	1	Is the problem you're working on challenging in some way? <sup>3</sup> Did you tackle challenges as they came up, or ignore them?
Statistical care	1	Did you conduct your analysis in a careful manner, attending to the assumptions underlying the models you fit or the techniques you used? E.g., model sensitivity checks, bootstrap check, etc.
Interest	1	Do you make an effort to show that your research is interesting and worth caring about and not just a way to show off modeling skills?

<sup>3</sup> You don't need to create an artificially difficult analysis, but you should go beyond just using the simplest models we've discussed.

### Note on calibrating the final grades due to project complexity

Some projects call for more complex modeling that push beyond the scope of course material. If this results in more errors in your work, the teaching team may calibrate your grade to take this additional complexity into account. *This does not mean that just making your project more complex will be rewarded.*

## 8. Forms and Sheets

The following pages have forms that we use for grading and structuring feedback.

You should print out a copy of the contribution form and sign it, if you are a group of 2 or more people.

We will provide feedback forms at the final project presentations.

### Project Team Statement of Contribution

**The undersigned members of this project team affirm that the source of all data is as represented, and that each member of the team has contributed to this project.**

In the space below, each member of the team should indicate their contribution to the project and sign below that description of work.

Team member 1 Contribution to the project:

---

---

Team member 2 Contribution to the project:

---

---

Team member 3 Contribution to the project (if applicable):

---

---

Team member 4 Contribution to the project (if applicable):

---

---

Team Name/Canvas Group Number

---

Team member 1 signature

---

Team member 2 signature

---

Team member 3 signature

---

Team member 4 signature

---