THE ASSIGNMENT: This is your primary goal with your work; review it often (and try to think from the perspective of your client). You need to use this to keep your work on target.

> Given the data set, characterize the populations and determine a good predictive model.

RESEARCH JOURNAL

Best practices in research include a systematic method of *journaling* – keeping a record of research-related activity including observations, conjectures, results (both positive and negative), insights, and reflections.

Journals should be kept by each member of the team, and your contribution to work will be evaluated using the journal. Your journal should be digital, be able to be submitted as a Word document or a PDF, and should demonstrate the following characteristics.

1. **Reflection:** The student will succinctly reflect on research activities including work attempted, insights and results gained, and further paths of inquiry to be pursued.

2. **Organization:** The student will create a single document (useful to self and others) that will serve as a quick reference of their work .

3. **Documentation:** The student will document when (dates) and for how long (approximate elapsed time) course-engagement occurred.

GUIDELINES FOR JOURNALING

1. **Style:** As long as all other elements are addressed, the student may use personal style in developing their research journal, although your professor reserves the right to direct changes to improve effectiveness. Journal entries are not meant to be formal pieces of research. Entries are to be useful rather than polished. Informal writing, abbreviation, and bulleted phrases are all acceptable as long as they can be easily deciphered by another mathematician. Daily entries should take less than five minutes.

2. **Frequency:** The student should journal on each day that they engage in research activities.

3. **Date/Time:** Each entry should start with the date, an approximation of total elapsed time for that day, and a running total of time spent on research.

4. **Entry:** The entry for a day should be a brief synopsis of and reflection on the research activity. Some prompts that could be addressed are . . .

   (a) What was the question you focused on?
   (b) What attempts were made? with what outcomes?
   (c) What impact might these outcomes have on other avenues of inquiry?
   (d) What questions arose? Do they need immediate attention, future attention, or setting aside?
   (e) What challenges appear to be hampering progress? How might these difficulties be resolved?
   (f) What should you work on next time?

5. **Submission:** The Research Journal is to be submitted to Canvas (or elsewhere if directed) with each project submission. Submissions should include your name in the file name as well as on the first page of the file. Failure to submit the journal will result in a 0 on that portion of the submission (your professor will not go searching for it).

Planning an Inferential Study

Statistical Tools

It will be easier to proceed if you understand what statistical tools are available and the different types of questions that they address. Complete the handout provided with the goal of understanding what specific research questions a given technique addresses.

Big Picture, Hypothesis Testing (10 pts)   Due online: 3/10/26, 11:59pm

Analysis Plan

Review the data definitions and the general research question. Usually there will be more specific questions that help address the general research problem. Clearly define the specific research questions you would like to address, and create a plan to address them. See the Analysis Plan template to help you organize the information needed.

Analysis Plan (10 pts)   Due online: 3/13/26 11:59pm

Research Meetings

Schedule a meeting with your professor during the weeks of March 16 and April 6 to discuss your progress and your plan for next steps. See the rubric for the expectations.

Meetings Completed (10 pts)   by: 3/20/26, 4/10/26, 3:00pm

Population Characterization and Predictive Modeling

R Initialization

Create an R project to contain all your work and save your environment/history and an R Markdown file to contain your code and the outputs. You will use the R Markdown file to run and record the results of your analysis.

** Files in R do not autosave, so save both your R Markdown and R Project files often.

Using section labels (#, ##, etc.) in R Markdown, create an R code section and a short text description/formulas for each question and analysis you identified in your revised Analysis Plan (group items by how they relate to the research questions, not based on the type of analysis). Impose structure/hierarchy to make it easier to follow and search through your code (use the Outline feature). Work on learning to format the text portions so that it knits attractively (Latex is helpful here). Import the provided data.

Mathematical Evidence of Analysis

Your RMarkdown/knitted files will provide the supporting work to justify your conclusions and models, desciptive statistics, details on your HT/CI/regression models, and any other supporting computational or graphical evidence.

It should be a well-documented summary of all the analysis you did, the decisions and conclusions you made and why they were justified. Summary statements before each code chunks should be included. One should be able to follow your file and understand your analysis and the conclusion you drew from each piece (even if it doesn't get included in the final report).

Submission: You will submit a draft with as much of the analysis as you can get done by the first deadline for feedback as well as a final draft. The same grading rubric will be used for both submissions. Your submission should include your journal, your RMarkdown file, and a knitted HTML file of your RMarkdown file.

Mathematical Evidence Draft (25 pts)   Due online: April 7, 2026, 11:59pm

Mathematical Evidence Final (150 pts)   Due online: April 23, 2026, 11:59pm

Report to Decision Maker

As you draw conclusions, draft a typed, attractive (but professional) report that summarizes your characterization of the relevant populations with all of the discussion that a non-technical audience needs to understand the analysis and how it can be used.

Assume the audience is well versed in the data and context (but you still need to be clear about variable definitions and terms you use) but does not spend their time doing statistical analysis. The average person should be able to read your report and understand what the results were, and the statistician should be able to read your report and be convinced based on the statistics that your conclusions are accurate. Appendices are permitted as appropriate/necessary.

This document should be drafted in parallel with your analysis run in R. Multiple revisions of the final document are expected (include a note at the end of the report indicating your total number of revisions). The primary goal is clarity to the audience while clearly justifying your conclusions and presenting all the details they need to make decisions without adding anything unnecessary.

Expected mathematical components for final report:

- Relevant/appropriate descriptives (using APA formatting, graphs as appropriate) included prior to and in support of your inferentials - descriptives should be used to lead the reader to pose the question answered by the inferential

- Results from inferential analysis (HT, CI, Regression - using APA formatting) - conclusions for each question should include a discussion of what the statistic tell us is most likely true about the population (use English, not jargon - no "reject the null" language)

- The implications of your conclusions relevant to the general research question

SUBMISSION: You will submit the best rough draft you can write of what you have done so far (after you have initially written and then revised it). Feedback will be given and a final draft will be submitted. Your submission should include your journal and your report.

DRAFT (25 PTS)   DUE ONLINE: APRIL 10, 2026, 11:59PM

FINAL (150 PTS)   DUE ONLINE: MAY 1, 2026, 11:59PM

MEETINGS ∼ 10 POINTS, RUBRICS WILL BE SCORED IN CANVAS

_____/ 35% Project Process: Journal maintained, appropriate time invested and progress made

Good progress: A/B Journal received by the meeting, time invested was adequate, progress made is sufficient to not be behind

Acceptable: C/D Journal received by the meeting, time invested was less than expected, progress was made but the work is behind schedule

_____/ 35% Journal Quality: Journal objectives, style , frequency (including date/times), entries quality

Good Document: A/B Journal objectives are met, style is effective, frequency is complete (including date/times), entries are substantive (addressing relevant prompts; all prompts should be addressed, but not every day)

Adequate: C/D Some journal objectives are met, style is adequate but can be improved, frequency is incomplete and/or does not include date/times, entries do not address all relevant prompts and/or some prompts are never addressed

_____/ 15% Meeting Preparedness: Prepared for the meeting with updates of progress and questions

Well Prepared: A/B Gave a clear update, had specific questions that were relevant and important to the work, had a plan for future work

Adequate: C/D Gave an update, had at least a question but was somewhat superficial/could have been answered by using the text, had some idea for what was next

_____/ 15% Meeting Management: Managed the meeting well

Well Handled: A/B Took initiative in the discussion, was comfortable discussing the work accomplished, asked questions clearly, did not need prompting, interaction was comfortable and conversational in nature while still being on task

Developing: C/D Participated in the discussion, was somewhat uncertain in places about the discussion, needed some prompting, interaction did not seem comfortable and conversational in nature or did not stay on task

_____/ 10 pts

MA 442 MATHEMATICAL EVIDENCE SUBMISSION $\sim$ 25 OR 150 POINTS

_____/ 20%   Presentation of mathematical work to an external audience (aka me):

Quality of organization/clarity of presentation - use of formatting/chunks/etc, logical presentation of content

F —— D^- —— D —— D^+ —— C^- —— C —— C^+ —— B^- —— B —— B^+ —— A^- —— A —— A^+

Quality of commentary - purpose and conclusions for each chunk/subchunk of code is clearly expressed in a logical and concise manner

F —— D^- —— D —— D^+ —— C^- —— C —— C^+ —— B^- —— B —— B^+ —— A^- —— A —— A^+

_____/ 80%   Met expectations for the mathematical components - mathematical work is comprehensive, compelling, and complete (made sufficient progress for the draft)

- Relevant/appropriate descriptives in support of your inferentials - descriptives should be used to lead the reader to pose the question answered by the inferential

- Results from inferential analysis (HT, CI, Regression) - conclusions for each question should include a discussion of what the statistic tell us is most likely true about the population

- The implications of your conclusions relevant to the general research question

F —— D^- —— D —— D^+ —— C^- —— C —— C^+ —— B^- —— B —— B^+ —— A^- —— A —— A^+

_____/ 0%   Penalty: poor presentation, failure to follow directions, etc.

_____ % = _____ / 25 or 150 points

Ma 442 Draft ∼ 25 points

_____/ 20%   Effective use of visuals and clear presentation of data for an external audience (aka management/decision maker)

$$F — D^- — D — D^+ — C^- — C — C^+ — B^- — B — B^+ — A^- — A — A^+$$

_____/ 40%   Sufficient information (both technical and non-technical) was presented (population characteristics were clearly stated/supported, audience of various technical skill has all they need to make relevant decisions)

$$F — D^- — D — D^+ — C^- — C — C^+ — B^- — B — B^+ — A^- — A — A^+$$

_____/ 40%   Clear/compelling narrative that explains and justifies the conclusions, easy to read and understand, all necessary information was presented and in the order needed to understand the argument; actionable conclusions were presented

$$F — D^- — D — D^+ — C^- — C — C^+ — B^- — B — B^+ — A^- — A — A^+$$

_____/ 0%   Penalty: poor presentation, failure to follow directions, etc.

_____ % =_____ / 25 points

The final draft rubric will contain all of the above, as well as include two additional pieces:

Y/N were appropriate revisions from rough draft feedback done, and

Y/N were sufficient revisions of the written document completed (you will be asked to estimate the number of times you revised the paper (you may include both rough draft and final draft revisions) - there is not a hard number I am looking for here, but I want to see that you put effort in critiquing your own work and improved your communication based on your own critique, as opposed to just responding to my own feedback).

Ma 442 Final Report ∼ 150 points

_____/ 8%   Revisions:

Sufficient revision based on feedback?   NO ———- YES

Final draft should be notably improved based on the feedback you received on the first draft.

Sufficient number of personal revisions?   _____   NO ———- YES

Evaluation is based on quality of the paper and effort invested in the revision process.

_____/ 16%   Effective use of visuals and clear presentation of data for an external audience (aka management/decision maker)

$$F — D^- — D — D^+ — C^- — C — C^+ — B^- — B — B^+ — A^- — A — A^+$$

_____/ 38%   Sufficient information (both technical and non-technical) was presented (population characteristics were clearly stated/supported, audience of various technical skill has all they need to make relevant decisions)

$$F — D^- — D — D^+ — C^- — C — C^+ — B^- — B — B^+ — A^- — A — A^+$$

_____/ 38%   Clear/compelling narrative that explains and justifies the conclusions, easy to read and understand, all necessary information was presented and in the order needed to understand the argument; actionable conclusions were presented

$$F — D^- — D — D^+ — C^- — C — C^+ — B^- — B — B^+ — A^- — A — A^+$$

_____/ 0%   Penalty: poor presentation, failure to follow directions, etc.

_____ % =_____ / 150 points