STAT 231 Project 3: Blog

Spring 2025

Timeline

All deliverables must be submitted by 11:59 PM on the dates provided at the appropriate submission location (except the project showcase, which is during class). Adherence to these deadlines is required to help us keep on pace for the semester.

Activity	Date	Submission
Blog plan	Thursday, April 10	New GitHub Issue (Blog plan)
Status update 1	Thursday, April 17	New comment on GH Issue (Blog plan)
Status update 2	Thursday, April 24	New comment on GH Issue (Blog plan)
Project showcase	Thursday, May 1	In class
Final blog post	Tuesday, May 6 by 5pm	Website (using GitHub Pages)
Reflection II	Tuesday, May 6 by 5pm	Gradescope

Assessment

The project components will carry the following weight:

Component	Points
Blog plan	10
Status update 1	5
Status update 2	5
Project showcase	15
Final blog post	75
Reflection II	10

Criteria for each component are shared below and will be assessed in the following way:

Achievement	Rubric	Max points
Excellent (Always)	Demonstrates criteria in full; could be used as an exemplar	100%
Good (Usually)	Nearly demonstrates criteria in full with a sound grasp	90%
Intermediate (Somewhat consistently)	Has some demonstration of criteria, but not consistently enough across all areas.	80%
Needs improvement (Sometimes)	Has some demonstration of criteria, but needs improvement across multiple areas	70%
Unsatisfactory (Rarely)	Features minimal demonstration of the criteria	60%
Unacceptable (Never)	Criteria largely overlooked	40%

[10] Blog Plan

- ☐ Plan contains all requested content
- ☐ Plan is submitted on time and in requested format.

[5] Status Updates x 2 (5 points each)

- ☐ Status Updates are submitted by their due date as instructed
- ☐ Status Updates contain all requested content (including any deviations from the original proposal, your progress on the project to date, and any major concerns the professor should be aware of)

[15] Project Showcase

Presentation:

- ☐ Group provides succinct summary of blog.
- ☐ Group is prepared with specific areas where they want feedback.
- ☐ Group presents in allotted time.
- ☐ All group members contribute to presentation and Q&A with good balance among speakers.

Peer feedback:

- ☐ Feedback is complete, thoughtful, and submitted on time.

[75] Final Blog Post

[5] File management

- ☐ Files are logically organized within folders in your blog repo.
- ☐ Folder and file names follow file-naming expectations (no spaces or special characters except "-" or "_") and make the content or purpose clear. It would be easy for anyone to navigate the contents of the repo without additional context.
- ☐ All files necessary for reproducing your project are on GitHub (qmd, PDF, Rproj, datasets, other imported files, …).

- ☐ There are no extra copies of files, "messy"/scratchwork files, or other files not used for your final project.
- ☐ .gitignore is used effectively to ignore unnecessary files (e.g., .DS_Store)

[20] Code

Readability: Code is formatted in a consistent style (consistent use of spacing, indentation, and naming conventions).

- ☐ Code follows recommendations for spacing (around operators, between arguments, etc.)
- ☐ Code follows recommendations for placement of line breaks (after pipe operators, between plot layers, logical breaks within functions if line is 80+ characters otherwise).
- ☐ Use of indentation, white space between sequences of code, and separate code chunks promotes readability.
- ☐ Object and function names make the purpose or content clear.

Documentation: Code is well-organized. Style, organization, and documentation make the code easy to follow. Location of code chunks is relevant to document section.

- ☐ Code chunks are split logically and located in relevant places in the document.
- ☐ Code is well-documented with the use of code comments. Documentation explains the *why* of the code.
- ☐ Code is well-organized and easy to follow.
- ☐ Only code that is evaluated is included in the final submission.

Functionality and code choice

- ☐ Code is fully reproducible without modification. Filepaths are relative to the location of the R or qmd file.
- ☐ Code chunk options are used correctly and customized as appropriate.
- ☐ Code runs without error and correctly produces desired output.
- ☐ Code is efficient with little or no redundancies in processes or stored objects.
- ☐ Code demonstrates proficiency with packages learned in class (primarily **tidyverse** and **tidyverse**-friendly functions)
- ☐ No (or only minimally necessary) data wrangling is present within any webpage's *.qmd* file.

[50] Content

For any statistical methods implemented:

- ☐ Choice of statistical method is appropriate and carried out correctly.
- ☐ Communication of statistical methods and results is correct.
- ☐ Communication of statistical methods and results is clear (easy for a reader to follow and understand).
- ☐ Communication of statistical methods and results is appropriate to audience level.

For any data visualizations:

- ☐ Visualizations are relevant and use effective visual cues.
- ☐ Graphs are well-suited to the type and scale of data they represent.
- ☐ Visualizations are easy to read, with clear labels and legends.
- ☐ Width and height of figures are aesthetically pleasing.

☐ Interactivity, if used, enhances the chosen representation of the data and is user-friendly.

For any tables:

☐ Tables are relevant and appropriate.
☐ Tables complement visualizations and are not redundant.
☐ Tables are publication quality, easy to read, with clear labels and nice formatting
☐ Number of digits displayed for numeric columns are consistent and appropriate.
☐ Interactivity, if used, enhances the chosen representation of the data and is user-friendly.

References:

☐ Data sources and any other relevant sources are appropriately cited.
☐ All necessary references appear to be included, and each reference list item is as complete as possible (e.g., title, author, date, link).

Other content:

☐ Blog incorporates at least one of the methods introduced in the second half of the semester: text analysis, network science, unsupervised learning, and/or spatial data.
☐ Purpose of blog is clear and elements are cohesive.
☐ Blog is compelling (interesting, creative, thought-provoking, or meaningful) with a solid introduction.
☐ Data sources and context are clearly stated, including, but not limited to: size of the data, how/when it was acquired by the blog team, how/when the data was originally acquired if from a survey, experiment, or other study. Relevant variables are explained.
☐ Claims or conclusions are adequately supported.
☐ Language of the blog is appropriate to the audience and clear. Context and description the topic is sufficient for a reader without subject-matter expertise to follow along.
☐ Little or no grammatical errors or spelling mistakes.

Formatting:

☐ Site is well-organized and uses headings and subheadings as appropriate to make the document easy to navigate.
☐ Text and analysis is effectively interwoven.
☐ Good choice of when to display code. Any code displayed is relevant to the audience.

[10] Reflection II

☐ Reflection is complete, thoughtful, and submitted on time.