

For your final exam written project, you will turn in your weekly labs into a draft of a paper. Working on this project in a group is allowed, but not required. If you are collaborating with someone, you should decide as a group how the work will be divided and what you would like to accomplish. I will not require that you finish all (or even most parts), but your collaborators can.

The following is one potential project outline, but you may follow any journal format you prefer. Note that while the ideal paper would include all of the following, **even published papers do not include all of the following**. You and your collaborators (if any) will need to decide what to include.

NOTE: Even published papers do not include all (or even most) of the following. You and your collaborators will need to decide what to include.

- 1. Abstract.
- 2. Introduction: Briefly introduce the problem, techniques, and outline the paper. Try to use as few technical terms as possible (or reference section where defined).
- 3. **Background**
 - o Problem description:

~~Partially~~ Fully describe the problem. **State where data downloaded** Describe how the data is created, **what is its format**, what are issues that one should consider (for example are their different types of noise), etc. Please keep in mind that people from a variety of backgrounds may be interested in your article, so please help them understand your data. Deeper mathematics can be applied if more people have a better understanding of the problem.

- Formatted: Strikethrough
- Formatted: Font: Bold
- Formatted: Strikethrough
- Formatted: Font: Bold

Data Format: How many data points? Is it Euclidean? If so, what is the dimension? How do you calculate distances between data points? How did you clean the data?

- Formatted: Font: Bold
- Formatted: Font: Bold

- o Mathematical background:

One can reference appropriate papers that describe the methods used in your paper or you can provide the background yourself. Including motivation specific to your problem would be particularly helpful.

3-5 pages describing some techniques used in your paper. (summarize some of what you learned, optional to include extra material.

- Formatted: Font: Bold
- Formatted: Font: Bold

- 4. **Results**
 - o Data analysis.

3 - 5 pages per person for group projects

In either the results or discussion section, poorly motivate your choices, both techniques and parameters. Output often depends on parameters. Motivate your parameter choices. Are these choices robust (for example, does one get similar results for different choices of parameter values).

- 5. Discussion
 - o What do your results mean?
 - o Compare your results to results using other types of data analysis tools.
 - o Validation:

Part of validation includes motivating and determining the robustness of the choices you made when analyzing the data and comparing your results to previously published results. But another part is to determine if the results give you useful information for your particular problem. For example, a cluster found in [Topology based data analysis identifies a subgroup of breast cancers with a unique mutational profile and excellent survival by Nicolau, Levine, and Carlsson](#) was validated by further examining the biological properties of this cluster. Of course, this requires laboratory experiments which your collaborators may or may not be able to perform.

- 6. Conclusion

7. Acknowledgement

You should acknowledge anyone who has provided significant feedback. If you publish the results of your project, please acknowledge this course. If I provide you with significant helpful feedback, you are also welcome to acknowledge me.

Formatted: Font: Bold

8. Author contribution

If group project. Summarize who contributed what to the paper (who designed, computed, analyzed, wrote, etc.).

Formatted: Font: Bold

9. Funding sources and conflicts of interest

Can just state none, but if related to another project, you should state so (for example if you have analyzed your data set in another class).

10. References

This is a very important part of your paper. It lets the reader know where to find additional information. One is also required to reference other people's ideas, analysis, conclusions, figures, etc (even if modified, reworded, or redrawn). Using other people's work without acknowledgment constitutes plagiarism. For figures, one may also need to obtain copyright permission if you submit your paper for publication, and redrawing a figure may be discouraged.

Formatted: Font: Bold

You must use a bib file.

Formatted: Font: Bold

Formatted: Font: Bold