

# Project Guidelines

STAT 5052

March 24, 2021

Data Analysis Term Paper is the final project paper for this course. Each student should identify a project topic, define the research questions, find the data to answer the questions, report the descriptive statistics, analyze the data, and write a paper to summarize the findings. You will be asked to submit final papers on Canvas. The report posted on Canvas is a good example of a final paper in terms of format (your analysis should be much more sophisticated). Here are one website for possible ideas for your project: <https://archive.ics.uci.edu/ml/index.php> You can find any dataset you think is appropriate.

## General Guidelines

- The term project paper should be in MS-WORD or PDF.
- In your report, you must have both descriptive and inferential statistics (it is **required** to use at least one method that we covered for the machine learning part).
- All the statistical tables and charts should be properly labeled. All the models need to be clearly written and algorithms need to be clearly stated.
- You are encouraged to use a reference section to cite reference papers and use the appendix section for showing any other significant information.
- There is no strict limitation on number of pages for this paper. But the average of a report of this kind is about 4 pages.
- The term project will be graded by the following criteria:
  - Appropriateness and accuracy of analysis 60%.
  - Effectiveness of communication:
    - \* Writing: Organization 10%, Professionalism & Style 10% (use the paper writing format of a professional journal in statistics).
    - \* Oral Presentation: 10% (will be credited to the other parts if not assigned to present).

- Thoroughness: 10%. (Good explanation on research design and discuss limitations and possible sampling bias. Literature review with reference. Etc.)
- The statistical software allowed to use including but not restricted to R, SAS, STATA, SPSS. **No raw output to be directly included in the report.** The results should be interpreted and readable to a reader that does not know how to use the software but is interested in the research questions.
- The final project report is due **noon May 10th 2020.**