Group Project Instructions
ENV710; Spring 2025

For this assignment, you will use your statistical knowledge and skills to analyze a real-world dataset and present your findings both in written and oral form. You will be assigned a group of 2-3 people by your TA. Group members may choose to pursue any research question related to the environmental sciences, and data may come from online databases, faculty- or student-collected data, or your own data collection. The project must include multivariable analyses (i.e., multiple independent variables) within a regression framework.

Abstract: An initial abstract of the group project is due on February 28. The objective of the abstract is for you to work with your group to define a research question(s) and identify an appropriate dataset (min. 80-100 observations and multiple independent variables that may explain the dependent variable(s)). For the final project, you will use more advanced statistics and models to analyze the dataset than those that we have studied to date; therefore, you are not expected to know which models to use to analyze the data at this point. The abstract should be approximately 1-page in length and consist of the following:

1. State your research question(s) and provide 1 paragraph of background information needed to understand the question or issue.
1. Explain why this question is relevant and interesting to environmental science, management or policy in 1 paragraph.
2. Describe your data, including the dependent variable(s) and independent variables in the dataset. What is your unit of analysis (e.g., cities, nations, corporations, individuals, plots of land, rivers, etc.)? What type of data are your dependent variables (categorical, nominal, continuous)? List the data source (website, faculty project, etc.).
3. (Optional) With the understanding that we haven't yet discussed the multivariate statistics and models that you will use to analyze your data, if you do have an idea of the statistical tools/tests you might use (e.g., ANOVA, multiple linear regression, and generalized linear models) briefly discuss them.
4. State any concerns you have about completing the project or specific information or skills you might need that are not listed on the syllabus.

Please turn in a single abstract for the entire group, uploading a Word document to Canvas that is entitled with the last names and first initial of each member of the group (e.g., "*LowmanHPerryKChenXAbstract.docx*").

Figures: Two initial figures are due on March 14. The Duke Library's Center for Data and Visualization Sciences will be visiting our class on March 18th for a consultation and review of group project figures. Ahead of this class, you must turn in a minimum of two figures that may be reviewed and discussed by DVS staff and the class. You may turn in any kind of figure for review – exploratory visualizations, a site map, multi-paneled plot, etc. – whatever you might find most helpful to receive feedback on. Please turn in two jpg files for the entire group, uploading the files to Canvas with the last names and first initial of each member of the group (e.g., "*LowmanHPerryKChenXFig1.jpg*").

Final Paper: The final written assignment is due April 10. Your group's paper should not exceed 4000 words (approx. 5 pages single-spaced) in length, not including any tables, figures, or the bibliography. Papers should be concise and well-written, use active verb tense, and use the first person. There must be a minimum of three figures created in R, and all figures (including tables and graphs) must be accurate, clear, informative, and attractive with relevant captions. Statistical analyses should be pertinent, correctly done, correctly reported, and entirely performed in R while collaborating via GitHub. The paper must include at least as many multivariate models as there are group members; so, each group member must conduct at least one of the analyses. The final written report should be organized as follows:

1. *Introduction* – Clearly define your research question along with a literature review to place your project in a broader scientific context. Be sure to articulate your research question and the hypotheses to be tested.
2. *Methods* – Describe how the data were collected and the data itself (dependent/independent variables, data sources, data missingness, etc.). Describe your data analysis approach including the types of models uses, any data transformations, and how you verified the assumptions of the model.
3. *Results* – The first section of your results should present some descriptive statistics of your data to summarize the dataset (e.g., number of observations, measures of central tendency and variation). The second section should present the findings of your statistical models. Be sure to interpret the strength of the effects of your variables, not just whether they were significantly important or not.
4. *Discussion* – Present your overall findings and their statistical vs. practical significance. What is the scope of inference? To what population can you infer these findings? Compare your results to previous studies as well as your initial hypotheses, and be sure to include some language addressing the limitations/future directions for this study.
5. *Attribution* – Provide a brief description of each participant's role in the project. For an example with appropriate role terminology, see the CRediT author statement published by Elsevier.
6. *Bibliography* – Use a consistent, recognized format for in-text citations and the bibliography. For an example, see the citation format used by the Ecological Society of America's primary journal, *Ecology*.

Please turn in a single file for the entire group, uploading an RMarkdown document as a pdf to Canvas entitled with the last names and first initial of each member of the group (e.g., "*LowmanHPerryKChenXFinalPaper.pdf*"). You must also share the GitHub repository of your code with Heili by April 10.

Final Presentation: The final presentation is due April 15. Your group will be responsible for giving a 5-minute, in-class presentation of your project to the class using Powerpoint or another similar presentation format. Each group member must present as part of the presentation, and it should cover the broader research questions, hypotheses, analysis approach, primary findings, and broader implications of your project. Following your in-class presentation, please turn in a single file for the entire group by the end of the day on April 15, uploading the presentation file to Canvas entitled with the last names and first initial of each member of the group (e.g., "*LowmanHPerryKChenXFinalPresentation.pptx*").