**BIOST 536/ EPI 536 2024**

**Data Analysis Project**

**Extended Abstract – Guidance 11/4/2024**

**Format**:

* **Maximum of two pages of text** **total** for Sections 1-4
* Minimum 1-inch margins on all sides, minimum font size 12, minimum line spacing 1.5 for pages 1 and 2.
* File format can be MS Word (preferred) or PDF.
* At the top of your document, include a **title** and your **group number** but **no names**.
* Filename convention: Label your submitted file according to this convention: Group03.docx for group 3; Group11.docx for group 11; etc.
* Only one person in the group should submit the final report via Canvas.

Your extended abstract should contain the following sections, which should be labeled:

**1. Background**: a *brief* background for the problem and the question you will address. (Reminder: a literature review is beyond the scope of this project).

**2a**. **Study Design**: Briefly summarize the study design and the available data (e.g. sample size(s)).

**2b. Statistical Methods**: describe the statistical methods that you used. Your description should have enough detail that a knowledgeable reader who had access to the data could, in principle, re-create your analysis from your description. Examples of such details include: variables adjusted for and the method of adjustment; how standard errors are estimated for regression parameter estimates (model-based or robust standard error estimates); the type of hypothesis test performed, if any. If any covariate is especially critical (strong confounder) or possibly controversial in terms of its role in the analysis, explain your reasoning for the role for the covariate in your analysis (e.g. why you did or did not adjust for the variable). (Note: there may not be any such covariates)

**3. Results**: Present results in a way that would be appropriate *for a general scientific audience*. You can assume your audience is knowledgeable about confidence intervals and P-values. Provide careful interpretation for key parameter estimates. Present and interpret key inferential results.

**4. Discussion**: Summarize the conclusions from your analysis. Discuss key strengths and weaknesses of the study.

**5. Use of AI tools** (does not count against the 2 page maximum). A reminder that using these tools for this project is allowed as long as you cite their use, as discussed on the first day of class. This citation should include a description of which AI tool was used and what it was used for. Reminder: it is never permissible enter datasets into AI tools and it is never permissible to copy text output verbatim into your report. If you never use AI tools in your report, you should state this in this section.

**6. Tables and Figures** (these do not count against the 2 page maximum). You can include a maximum of 4 total tables plus figures. Each table or figure should have an informative caption. Each table and figure should have a label (e.g., Table 1, Figure 1) and your text should refer to these labels. Tables and figures should be constructed thoughtfully and with careful attention to detail. As always, unedited software output is not acceptable. Note that tables can be single-spaced because tables with lots of white space can be difficult to read.

For example, you might choose to include two tables in your document, where Table 1 provides descriptive statistics summarizing the subjects in the study and Table 2 provides the results of the main analysis. You might include a figure if you believe it adds importantly to your Extended Abstract. Alternatively, you might decide that figures are not especially helpful for summarizing the study or your analysis, and thus decide not to include any figures. General advice: focus on quality over quantity for tables and figures.

6. **Code** (does not count against the 2 page maximum; a separate file is preferred). Provide a copy of the code for your analysis. Your code should have adequate documentation so that a TA or instructor could replicate the reported analysis.

**Additional Guidance**:

Each group should work together to produce a single Extended Abstract that addresses the study question. A grading rubric will be provided, but here is some additional guidance.

* Although groups will collaborate on the project, only one person per group should submit the final paper on behalf of the group.
* Your job is to develop and appropriate analysis plan, analyze the data accordingly, and report the methods and results clearly and succinctly.
* Methods to handle missing data have not been covered in this course. There are no expectations for handling missing data other than transparency regarding the extent of any missing data and how missing data were handled (e.g. complete case analysis).
* Direct questions to Katie or Pam. Use the course discussion board to ask questions if appropriate. If you are unsure, email your question to [pshaw@uw.edu](mailto:pshaw@uw.edu) and/or [katiek@uw.edu](mailto:katiek@uw.edu). You should not consult with students in the class outside your group, and you should not consult individuals outside the class.
* A grading rubric will be provided, and you should review it to understand how your Extended Abstract will be evaluated. Here are some additional guidelines:
  + Do not plagiarize, including from this document or from AI tools.
  + When data come from an observational study, be careful to avoid language that sounds “causal” or that could lead readers to think causal conclusions can be drawn.
  + Take care with the word “significant”. Do you mean clinically significant? Statistically significant? It is often best to reserve the word “significant” for its technical meaning in statistics. Etymologically, “significant” means “signifying something”. If you really mean “large” or “substantial” or “important” then use one of those words instead of “significant.”
  + If you say that unmeasured confounding is a study limitation, you should be more specific: what potential confounders do you have in mind?
  + Proofread your Extended Abstract for spelling and grammar.