Project Guidelines

Overview

The final project will allow you to delve deeper into a topic of interest. Option 1: a research design in which you describe a research question or identify a critique and work through various issues you would need to confront in answering the question or resolving the critique using the techniques that we will cover in class. Option 2: find an existing analysis and replicate it, describing what the author did and exploring alternative approaches or diagnostics. You may use Python or R. The project will comprise 30% of your course grade.

Requirement	**Due**	**Percentage**
Memo 1: Topic and Data	October 7	5%
Memo 2: Analysis Plan	November 9	10%
Insight Report	December 16	15%
Project Total		**30%**

Insight Report

The report must be uploaded to Canvas by 11:59 PM on **Wednesday, December 16, 2020**, and must include (1) all code used to generate the analyses, tables, and plots, and (2) all datasets as imported into Python or R scripts/notebooks. The report must include the following:

- **Executive Summary**
- **Introduction**
  - Provide an overview of your topic.
  - Discuss the current methodological approaches and discuss reasons why these may not be optimal. Potential areas to critique include: research questions, datasets, units of analysis, techniques, and interpretation of findings.
  - Explain how your approach differs from current approaches.
- **Data**
  - Discuss the types of data that would be ideal for analyzing this project.
  - Provide a review of the datasets used for the analysis, including relevant numerical and graphical summaries.
  - Discuss the limitations of the datasets.
- **Methodology**

  - Discuss the intuition behind the technique(s) you are using. You must utilize at least two techniques covered in PPOL 565 and PPOL 566, and one of these must be covered in PPOL 566.
  - Provide examples of how these techniques were applied in a different setting (journal article, news article, blog, video, etc.).
  - Discuss the strengths and weaknesses of each technique.

- **Findings**
  - Discuss the details of your analysis, including estimation and evaluation.
  - Provide an interpretation of the results of your analysis. What insights did you generate from the analysis?

- **Conclusion**
  - Discuss the implications of your insights. How might these insights be applied in your topic domain?
  - Discuss limitations (ethical, computational, data, etc.) and future considerations.

- **Bibliography**
  - The bibliography must be in APA format.

- **Implementation Appendix**
  - Use this section to discuss any relevant, interesting, or innovative aspects of your technical implementation.
  - For example, did you scrape the web or use an API? Develop any new measures? Reshape and/or pre-process the data?

Project Memos

You will submit two memos to Professor Brodnax detailing the progress on the project.

Format

Each memo should be in professional memo format. For information on how to format a professional memo, see: `https://owl.purdue.edu/owl/subject_specific_writing/professional_technical_writing/memos/index.html`. Your memo should be clearly written, with proper spelling and grammar. You must also include appropriate citations and a bibliography formatted in APA style. NOTE: The bibliography will not be included in the 600- to 1000-word guideline. For more information about formatting references, see `http://www.apastyle.org/learn/tutorials/basics-tutorial.aspx`.

Memo 1

The first memo must be submitted to Canvas by 11:59 PM on **October 7, 2020**. This memo should be approximately one page, single-spaced, and discuss your topic and potential dataset(s). You must provide a citation with a web address for any publicly available dataset, and the dataset must have at least 500 observations.

Memo 2

The second memo must be submitted to Canvas by 11:59 PM on **November 9, 2020**. This memo should be up to two pages, single-spaced (not including figures or bibliography), and provide a preliminary review of the data.

You must describe the source and nature of each dataset. Include a table of descriptive statistics for each relevant variable (description, mean, standard deviation, minimum, and maximum for continuous variables, or frequency counts for categorical variables). Discuss how your data are distributed. Provide visual summaries of the data and discuss what these summaries suggest about the appropriateness of the technique(s) you have selected.

Discuss your analytical plan. What techniques will you use? Why? Provide a brief explanation of the intuition behind each model.

Rubric

Area	Expectations
Submission	All components were submitted on time: (1) report (PDF), (2) data, and (3) replication code. All datasets include at least 500 observations.
Coverage	Report content comprises all components of the report outline above, including: an introduction of the problem and contextual background; description of the datasets as well as numerical and visual summaries; explanation of at least two analytical techniques, including a review of examples from outside of class; interpretation of the analysis and evaluation; and discussion of implications and limitations. Report sections are clearly indicated via formatting.
Data	Each dataset includes at least 500 observations. Summaries are provided for all *relevant* variables and/or relationships (minimum of two). Only those summaries needed to answer the question or guide interpretation are included. Numerical values are neatly formatted in tables. The numbers displayed in tables have no more than three digits after the decimal place. Interpretations are provided for all tables.
Graphics	Plots are provided for all *relevant* variables and/or relationships (minimum of two). Only those plots needed to answer the question or guide interpretation are included. The plots include titles, axis labels, and element labels when needed. The labels and table headings are brief and meaningful. The plots are easy to understand without outside information. Interpretations are provided for all plots.
Techniques	Analysis includes at least two techniques covered in PPOL 565 or PPOL 566, one of which is covered in PPOL 566. Content includes an explanation of each technique written for a non-technical audience. The discussion demonstrates understanding of technique and does not regurgitate documentation. Examples of technique applications do not use data or examples discussed in class.
Interpretation	The interpretation discusses key insights derived from the analysis, as well as an evaluation of techniques.
Writing	The text provides a coherent narrative with data and uses proper grammar and spelling. The report includes proper in-text citations and a bibliography in APA format.
Replication	Important data acquisition and pre-processing details are discussed in the Implementation Appendix. The scripts run with no errors and no warnings. The code has meaningful variable names, has liberal comments that provide context, and is easy to follow. The code utilizes conditional statements, loops, and functions. The code also contains no instances of repetition (identical or nearly identical commands appearing more than twice). Script has a logical, intuitive flow.