Final Project Rubric

PPOL670 – Introduction to Data Science

Fall 2020

Student:_____

Project Name: _____

Total Score: _____ / 50

Project Materials

4 points

- Report was posted to Canvas as a `.zip` containing the following items:
  - Report was rendered using RMarkdown as any one of the following file types: `.pdf`, `.html`, `.docx`. File was titled `lastname_firstname_final_report.pdf`. (___/1 point)
  - `.Rmd` file containing all the code used to generate the analytics in the report. File was titled `lastname_firstname_final_report.Rmd`. (___/1 point)
  - Student included the data used in a `Data/` folder. (___/1 point)
  - Student included an `.Rproj`. (___/1 point)

Document Presentation

16 points

- **Student used professional looking visualizations in the report:**
  - Figures were easy to understand? (___/1 point)
  - Figures made sense within the context of the report? (___/1 point)
  - Student described the purpose and the insight drawn from the figure in the text? (___/1 point)
  - Figures referenced in the t ext are labeled, i.e. references to "figure 1" correspond to the figure title (e.g. "Figure 1: Title")? (___/1 point)
  - Figures include titles? (___/1 point)
  - Figures labels/axes/text are readable? (___/1 point)
  - Color scheme made sense; easy to differentiate between colored items (___/1 point)
  - Figures were appropriately proportioned to the document? (___/1 point)
- **Student used `R` Markdown for a professional looking report:**
  - Report was rendered without errors or warnings. (___/1 point)
  - No code was visible in the report. (___/1 point)
  - No raw output was visible in the report. (___/1 point)

- Report includes a title, author byline, and word count. (___/1 point)
- Report is 12 pages in length (double-spaced; 12 pt font; if rendered as `.pdf`/`.docx`) or 3000 words.[1] (___/1 point)
- Report contained no (or few) grammatical/spelling errors. (___/1 point)
- Report reads as a single cohesive document. (___/1 point)
- Student cited academic, data, and package sources. (___/1 point)
  - To cite a package, use `citation("package_name")` to get a the citation information for a package, e.g. `citation("ggplot2")` will yield "*H. Wickham. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016.*"

Content

Points 30

The student's project sufficiently addressed these general areas.

- **Introduction** (___/5 point)
  - Student clearly established the aim of the project.
  - Student offered a clear roadmap of the report (i.e what is covered in the report).
- **Problem Statement and Background** (___/5 point)
  - Student offered a clear and complete statement of the problem and/or aim of their analysis.
  - Student included a brief summary of any related work (i.e. a *light* literature review)
- **Data** (___/5 point)
  - Student outlined where their data came from.
  - Student clearly specified:
    - the unit of observation;
    - variables of interest;
    - potential issues in the data (e.g. missingness, coverage, etc.)
  - Student articulate the steps they took to wrangle the data.
- **Analysis** (___/5 point)
  - Student described the methods/tools they explored in their project.
    - Justified the tools/methods that they used.
    - Adequately described what the tools/methods are doing.
    - Note: Assume the reader is smart but doesn't know `R`/Machine Learning well. That is, be crystal clear about what you're doing and why.
- **Results** (___/5 point)
  - Student gave a detailed summary of their results.
  - Student presented their results clearly and concisely.

[1] Note that your citations do not count against your word/page count.

  - Student used visualizations (and tables) whenever possible/appropriate.

- **Discussion** (___/5 point)
  - Student spoke on the "success" of their project (as defined in their proposal).
    - "Did you achieve what you set out to do? If not why?"
  - Student articulate how they would expand the analysis if given more time.