World Happiness Report Case Study Rubric

DS 4002 – Fall 2023 - Carson Crenshaw
Submission format: Submit a link to your GitHub repository and a PDF copy of your final presentation.

Individual Assignment
General Description: The objective of this project is to determine which life evaluation factors are associated with high national happiness scores. Once the analysis is completed, submit a link to your case study repository and PDF copy of the final presentation.

Preparatory Assignments: A basic programming course, the case study description, and attached resource materials.

Why am I doing this? As a second-year student, building foundational coding and explanatory skills is crucial to your development as a data scientist. Being able to address complex questions and communicate analytical conclusions to a wide audience is the crux of your role. In this regard, practicing executing a project of this size and complexity will familiarize you with your future professional responsibilities. Establishing a strong capacity for developing graphics will enhance your ability to convey information quickly and effectively. Additionally, these simple skills will carry over into future work as you learn more complex analytical methods (regression, predictive modeling) in later courses.

As a second-year student, you are just beginning to explore what data science is as a field of study. While you have not yet covered more complex analytical tools such as regression, conducting a sufficient exploratory analysis of data is an important skill to first develop. Most of the introductory courses within the Data Science school at UVA focus on the presentation of data, so this project should be an extension of the knowledge you currently possess.

- Learning Objective: Organize, condense, and clean publicly available data
- Learning Objective: Use critical thinking to devise the best illustration of information
- Learning Objective: Develop intelligible and efficient code (organized GitHub repo)
- Learning Objective: Prepare findings for presentation to your superiors

What am I going to do? Within this project you will be required to develop a report for the United Nations as if you are a student intern supporting international development goals. After reading through the supplementary GitHub materials, you will individually determine a modeling approach and begin to interpret the data provided. After performing your analysis, you will prepare a presentation to communicate your conclusions and guide an audience through your process. Deliverables include:

- GitHub repository: To provide all annotated code, figures, and resources constructed.
- Presentation: A short report that will quickly demonstrate your findings to your supervisor.

These deliverables will be submitted electronically via a link to a github repository and a PDF copy of the presentation in line with the United Nations' sustainability guidelines.

Tips for success:

- Carefully read and understand the prompt and provided materials. Be sure that your final deliverable specifically identifies which life evaluation factors are associated with high national happiness scores.
- Don't overcomplicate your findings. Making good data visualizations may seem overwhelming because of the extent of information you are given. Work carefully to determine what should be conveyed to your audience.
- Be creative. There is more than one way to illustrate trends in a compelling manner. Think outside the box.
- Remember that you are working with temporal data. The best representations of data will do a good job of incorporating time into their graphics. This could mean working with interactive modeling packages.

How will I know I have Succeeded? You will meet specifications when you follow the criteria in the rubric on the next page.

Spec Category	Spec Details

Formatting	• GitHub Repository
	  ○ The top level page should contain a README.md file, a LICENSE.md file, a SRC folder, a DATA folder, and a FIGURES folder.
	• PDF Presentation
	  ○ Linked in the FIGURES folder of the GitHub repository.
	  ○ Should contain no more than 10 slides.

GitHub Repository	• Goal: This repository should serve as an extensive, complete documentation of your project. Audiences and future researchers should be able to download this repository and produce the same conclusions. Ideally, this website will be used as a foundation for future projects.
	• Use markdown headers to divide content within the README.md file, and link each header to the individual folders referenced.
	  ○ Source Code (SRC) Folder
	  ■ Detail necessary aspects of installing and building code.
	  ■ Should include an annotated markdown file which details the packages used.
	  ○ Data Folder
	  ■ Must contain pre- and post-cleaning datasets
	  ■ Relink dataset provided from original GitHub
	  ■ Include an annotated data dictionary that walks through relevant details of the data
	  ○ Figures Folder
	  ■ Summarizes all figures produced and key findings from them
	  ■ Contains downloadable JPEG versions of all figures
	  ■ Also includes a link to the final PDF presentation
	  ○ License declaration should be included (MIT license).
	  ○ All references should be listed at the very end of the README.md file using the IEEE documentation style. Also include any acknowledgements.

Presentation	• Goal: Demonstrate your project plan and analytic conclusions at the end of the project cycle. This presentation will be your primary method of conveying your findings to your intern supervisors.
	• No more than 10 slides. PDF format submission. No longer than 10 minutes.
	• The structure of the presentation should contain:
	  ○ Title and Outline
	  ○ Motivation and Modeling Approach
	  ■ Can also include any research questions generated after reading the foundational materials provided.
	  ○ Data Acquisition
	  ○ Analysis Plan/Tricky Analysis Decisions/Bias and Uncertainty Validation
	  ■ Explain to your audience how you reached your conclusions. Limit your use of technical terms.
	  ○ Results (at least 2 slides)
	  ■ Focus on the figures. Instead of supplemental text, explain what the graphs represent.
	  ○ Conclusions
	  ○ References

Acknowledgements: Special thanks to Jess Taggart from UVA CTE for coaching on making this rubric. This structure is pulled from Streifer & Palmer (2020).