

Reconceptualizing the Measure of International Progress:

Utilizing National Happiness Scores to Assess Quality of Life

DS 4002 | Carson Crenshaw | December 11, 2023



What determines a good life? In a society that is increasingly driven by money and power, happiness and quality of life have the potential to be overlooked and under prioritized. Happiness, however, is a powerful tool that can be used to optimize the well-being of whole populations. Acknowledging this, the General Assembly of the [United Nations began publishing](#) the annual World Happiness Report. This report reviews the current state of happiness in the world and illustrates national variations in happiness for a given year. Ultimately, the research reflects an international demand for a greater understanding of happiness.

As a bright new U.N. intern tasked with studying international development, you have the potential to help determine which issues should be prioritized in future U.N. programs. You believe that people's overall wellbeing is the ultimate good and that the international community should center itself around this goal. World Happiness scores can and should be employed to strengthen societies and achieve better standards of living. As happiness is increasingly considered an important measurement of national progress, you are tasked with examining the relationship between World Happiness scores and life evaluation factors. By figuring out the relative importance of different domains in shaping happiness, you will be able to provide evidence-based guidance for your policymaking supervisors on how to achieve a stronger social fabric and better public welfare.

Deliverable

The objective of this project is to determine which life evaluation factors are associated with high national happiness scores. You will be tasked with presenting your findings to your intern supervisor in order to demonstrate which socioeconomic factor contributes the most to a high national happiness score. This issue will then be the focal point of the next international U.N. development project. In this regard, you want to advocate for the factor which is most likely to lead to a positive growth in public welfare. In order to support your conclusions, you will need to create a (1) GitHub repository to store your data and (2) a brief presentation.

Potential Modeling Approach

This project involves exploring geographical and temporal trends of six different factors in relation to happiness scores and global happiness rankings. Although any number of illustrative methods can be used, choropleth maps, scatterplots, and/or correlation heatmaps can all be utilized to visually convey which life evaluation factors are associated with high happiness scores. Your goal is to demonstrate your findings in a clear and concise manner. Choosing the proper graphical representation is paramount to your success and your persuasive ability.

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	<ul style="list-style-type: none"> ● GitHub Repository <ul style="list-style-type: none"> ○ 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 <ul style="list-style-type: none"> ○ Linked in the FIGURES folder of the GitHub repository. ○ Should contain no more than 10 slides.
GitHub Repository	<ul style="list-style-type: none"> ● 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. <ul style="list-style-type: none"> ○ Source Code (SRC) Folder <ul style="list-style-type: none"> ■ Detail necessary aspects of installing and building code. ■ Should include an annotated markdown file which details the packages used. ○ Data Folder <ul style="list-style-type: none"> ■ 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 <ul style="list-style-type: none"> ■ 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	<ul style="list-style-type: none"> ● 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: <ul style="list-style-type: none"> ○ Title and Outline ○ Motivation and Modeling Approach <ul style="list-style-type: none"> ■ Can also include any research questions generated after reading the foundational materials provided. ○ Data Acquisition ○ Analysis Plan/Tricky Analysis Decisions/Bias and Uncertainty Validation <ul style="list-style-type: none"> ■ Explain to your audience how you reached your conclusions. Limit your use of technical terms. ○ Results (at least 2 slides) <ul style="list-style-type: none"> ■ 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 \(2020\)](#).

Reference Documents for Case Study Project

Data

J.F. Helliwell, R. Layard, J.D. Sachs, L.B. Aknin, J.E. De Neve, & S. Wang, “World Happiness Report 2023 (11th ed.),” *Sustainable Development Solutions Network*, 2023. Available: <https://worldhappiness.report/ed/2023/#appendices-and-data>.

C. Crenshaw, “CaseStudy_DS4002,” *GitHub*, Dec. 11, 2023. [Online]. Available: https://github.com/C-Crenshaw/CaseStudy_DS4002.git.

This GitHub repository contains all information necessary to complete the project specifications, including a downloaded copy of the dataset above.

Content Resources

“About | The World Happiness Report.” Available: <https://worldhappiness.report/about/>.

The previous resource describes the history and development of the World Happiness Report. This report can be utilized to contextualize the analysis conducted and to explain why happiness should be utilized by governmental organizations.

J.F. Helliwell, R. Layard, J.D. Sachs, L.B. Aknin, J.E. De Neve, & S. Wang, “Forward,” *Sustainable Development Solutions Network*, Mar. 18, 2022. Available: <https://worldhappiness.report/ed/2022/foreword/>.

This website is linked in the introductory/hook document. Similar to the “About” article above, this source provides further another history of the World Happiness Report for context.

J.F. Helliwell, R. Layard, J.D. Sachs, L.B. Aknin, J.E. De Neve, & S. Wang, “World Happiness Report 2023 (11th ed.),” *Sustainable Development Solutions Network*, 2023. Available:

<https://worldhappiness.report/ed/2023/#appendices-and-data>.

In addition to providing the foundational data utilized within this project, this source is an extensive documentation of the implications of the World Happiness Report in 2023. More specifically, this report details how income, health, social support, freedom, generosity, and corruption all play strong roles in supporting life evaluations.

Coding Resources

C. Sievert, “Interactive Web-Based Data Visualization with R, Plotly, and Shiny,” *plotly-r.com*, Dec. 19, 2019. Available: <https://plotly-r.com/>.

This website highlights how to use the plotly and shiny libraries within the R coding software. These packages can be used to create complex graphic representations of the World Happiness Data. In particular, interactive data visualizations can be constructed to clearly demonstrate a change over time.

Y. Holtz, “Choropleth Map with R and GGPlot2,” *The R Graph Gallery*, 2018. Available: <https://r-graph-gallery.com/327-chloropleth-map-from-geojson-with-ggplot2.html>.

This article walks the reader through how to build out a choropleth data visualization map. These visualizations are often the best at demonstrating the differences between multiple entities on one variable of interest.

Note: A physical copy is provided for the red colored documents above.