

Tulsi Patel
Prof. Alonzi
DS 4002: Prototyping
April 28, 2025

FINAL PROJECT RUBRIC

Submission Format: GitHub Repo with finished outputs and scripts, as well as a clear ReadMe file that details where the student had trouble with the instructions or the code and documents their process.

Why are you doing this?: The goal of this project is to get the student aware of issues that exist on a global scale and how data science can be applied to these larger issues.

Spec Category	Spec Details
Data Cleaning	<ul style="list-style-type: none">• Use Stata to clean the data• Load the dataset correctly from an Stata file (.dta) into the R workspace and explore its structure• Handle or remove missing values• Rename variables if necessary for clarity and consistency.• Document each cleaning step clearly in comments or markdown.• Ensure that the cleaned dataset is ready for analysis without errors or inconsistencies.
Data Analysis	<ul style="list-style-type: none">• Generate time series plots of HDI and/or GDI over time to identify trends, seasonality, or structural breaks.• Examine distributions to understand the range and shape of the data.

	<ul style="list-style-type: none"> • Calculate and interpret summary statistics (mean, median, standard deviation) by year, country, or region. • Comment on any missing or inconsistent data and how it might affect interpretation. • Use EDA to justify the modeling approach and whether a time series model is appropriate. • Support visual findings with text commentary that connects to the research question.
Data Visualization	<ul style="list-style-type: none"> • Use appropriate plots to explore both HDI and GDI across countries or regions. • Include clear axis labels, titles, and legends to aid interpretation. • Use color, shape, or facetting to compare groups (e.g., continents, development categories, gender differences).