Future Flare

Project presentation

Team Members

Ajinkya Chafekar
Kavya Nagesh
Pratyush Singh
Dolsy Arora

Contents

01/ Introduction

02/ Dataset

03/ Deliverables

04/ AI Models

05/ Visualization

Thank You

01/ Introduction

In our dynamic world, climate change deeply impacts us and the planet. Embracing the circular economy minimizes waste, seen in innovations like car-sharing and modular designs. As urgency grows, finding impactful solutions becomes crucial. Yet, evaluating these solutions overwhelms human evaluators due to their vastness. AI EarthHack steps in: leveraging AI's power to create a decision-support tool. This tool streamlines evaluations, boosting accuracy. Combining human judgment with AI analysis, our goal is to highlight innovations effectively addressing climate change through the circular economy.

02/
Dataset

We were given access to the dataset named
AI EarthHack Dataset csv file.
Which contains 3 columns and 1300 rows.

Columns: [”id”,”Problem”,”Solution”]

03/
Deliverables

To Provide a comprehensive summary of our tool, highlighting its technical intricacies and creative aspects. Additionally, showcase the tool's functionality through a demonstration.

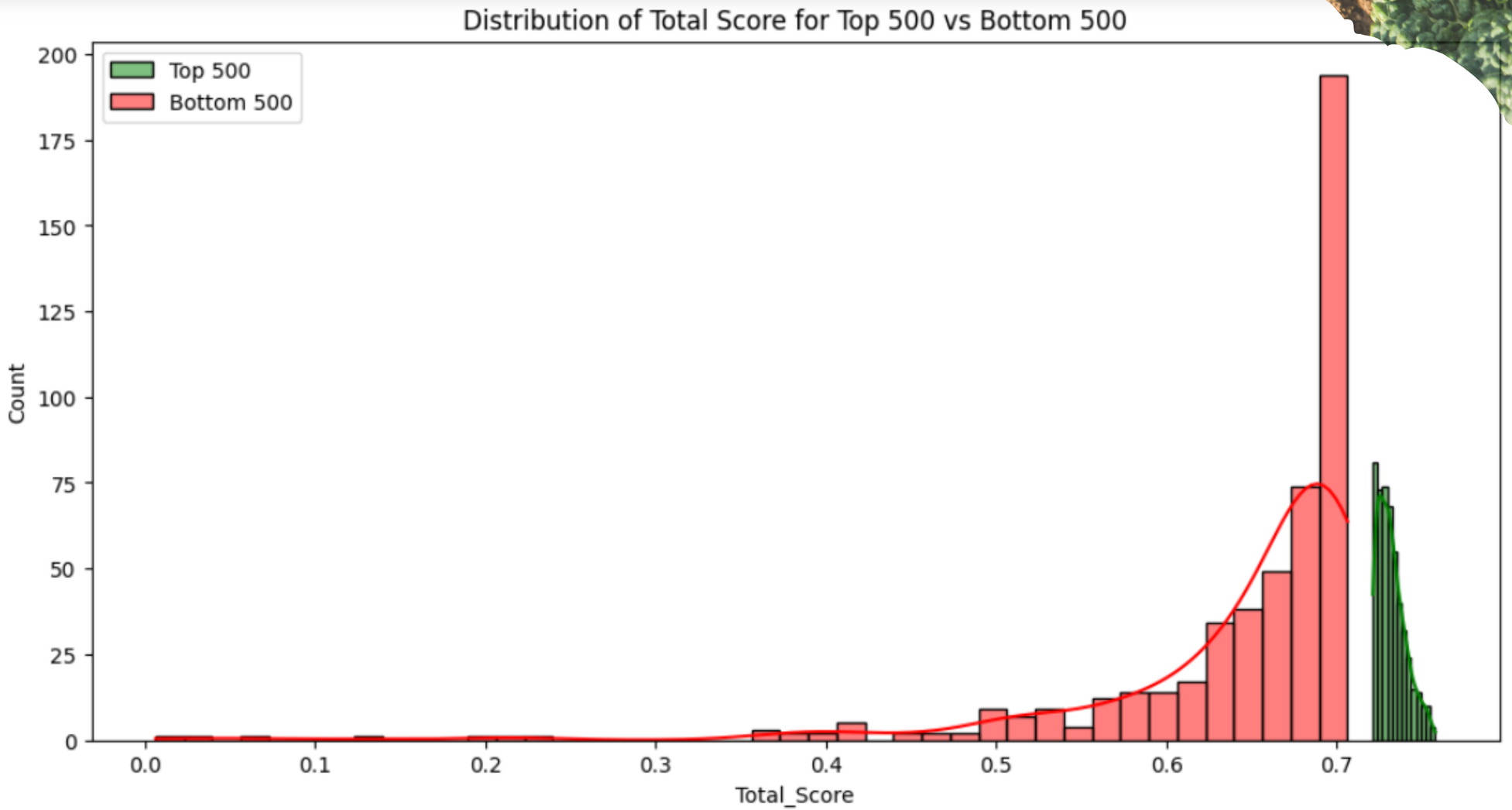
04/
Generative AI model taken into consideration

1. BERT(Bidirectional Encoder Representations from Transformers)- A pre-training technique in natural language processing (NLP), employs transformers to grasp word context in a sentence, enhancing language comprehension through consideration of both preceding and following words for a given term.

2 ROBERTa- ROBERTa, a refined version of BERT, optimizes pre-training methods, addressing limitations and achieving superior performance in natural language understanding tasks.

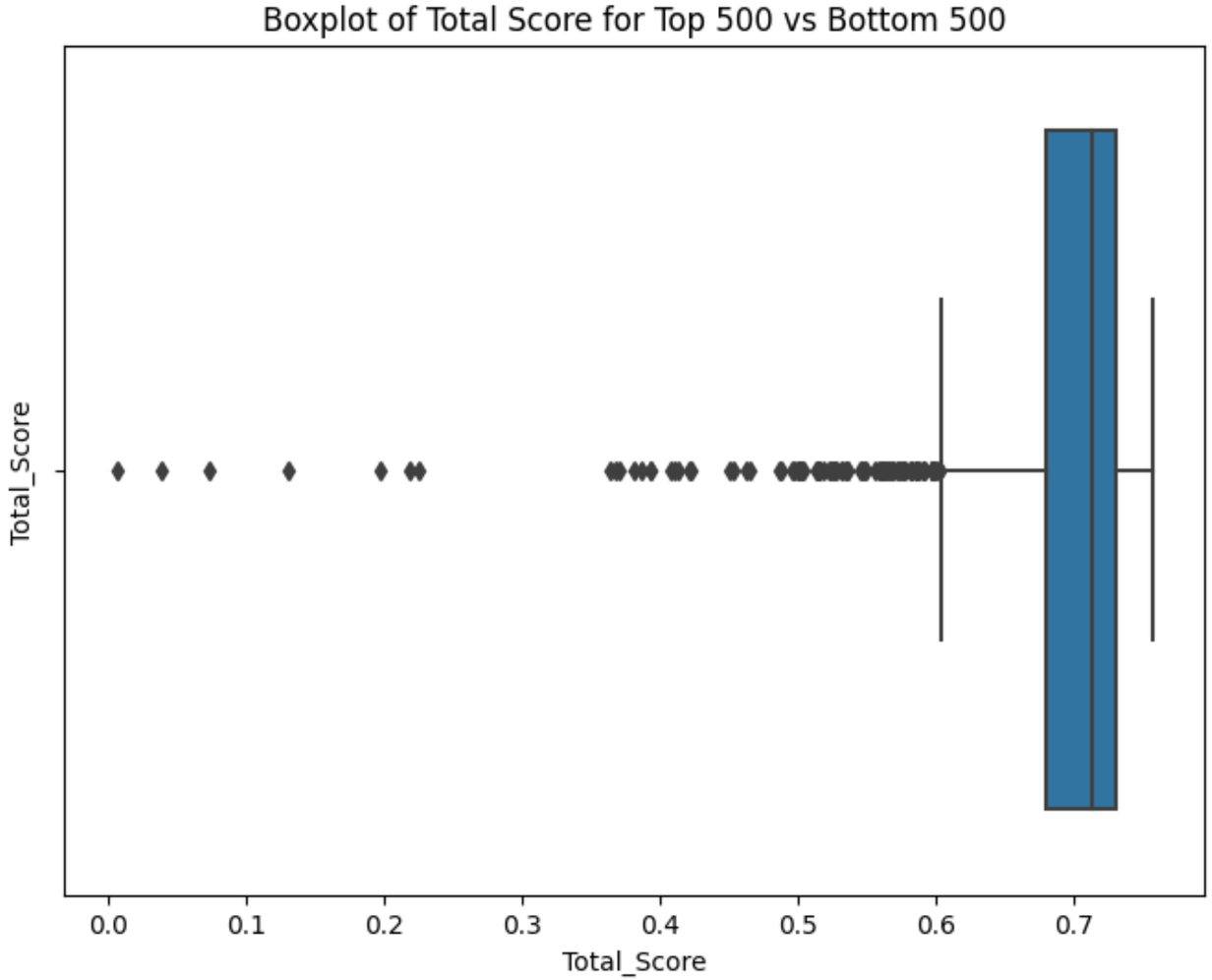
3. DistilBERT- DistilBERT, a streamlined adaptation of BERT, maintains contextual understanding in language with reduced model size and quicker inference speed, catering to diverse natural language processing applications under resource constraints.

Histogram



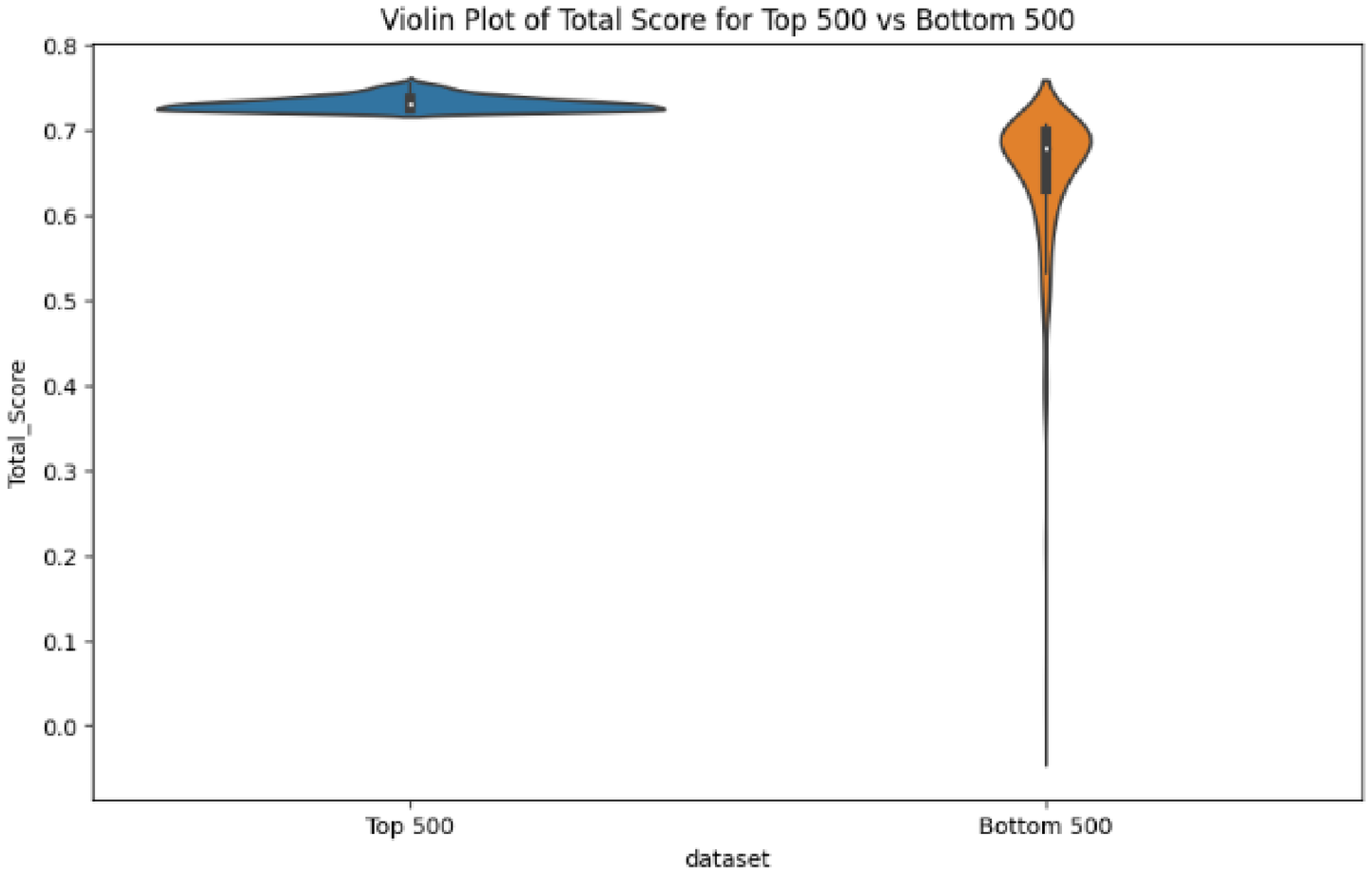
In the above graph, it shows how the top 500 and bottom 500 datasets are distributed based on the Total_Score, that they have acheived.

Boxplot



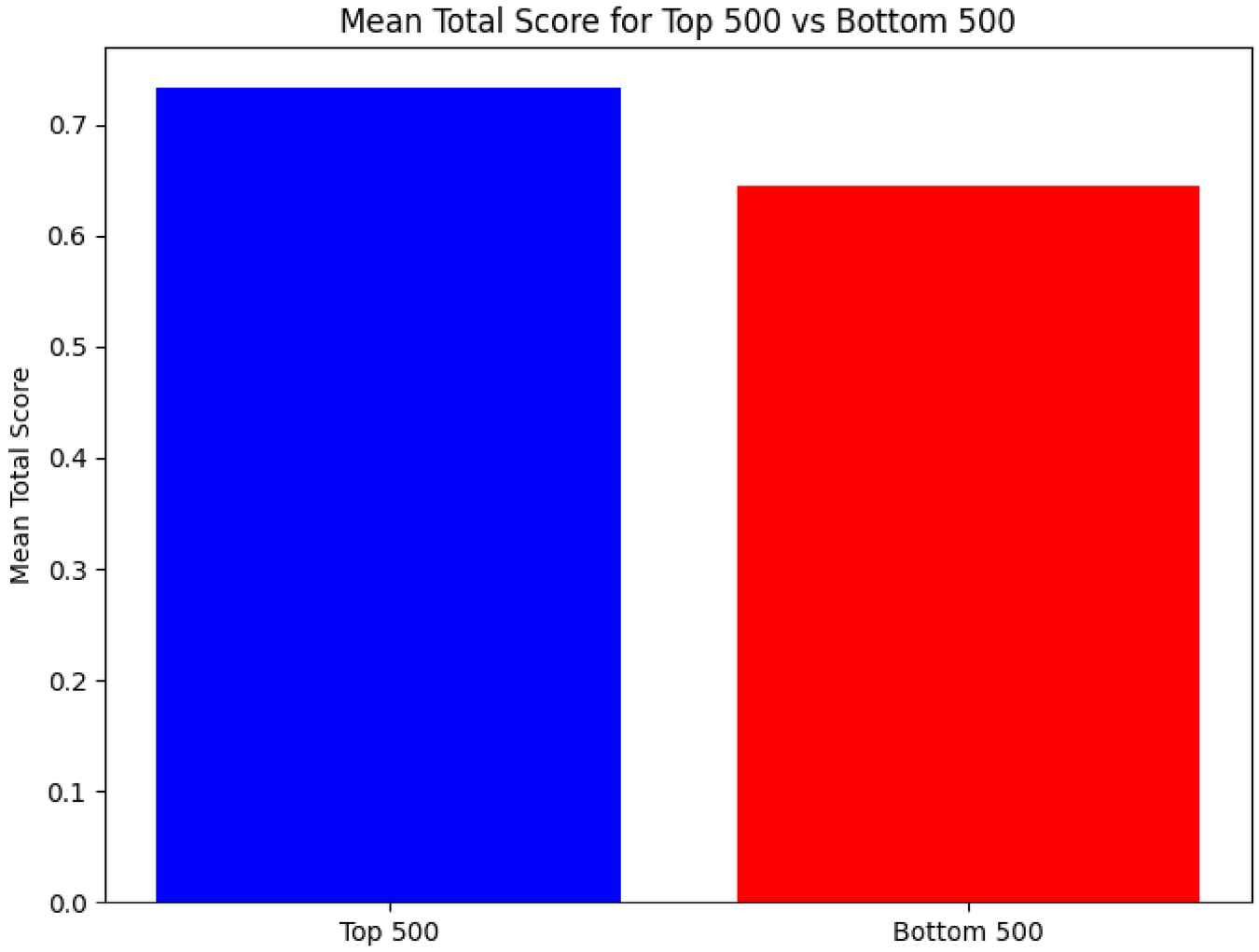
As it is clear from the boxplot how the data are distrbuted and the mean dataset fall within the rage of 0.7 and the black diamonds signifies the outliers from the dataset whose Total_Score value is 0.0. This can be verify by the below output of the same record.

Violin Plot



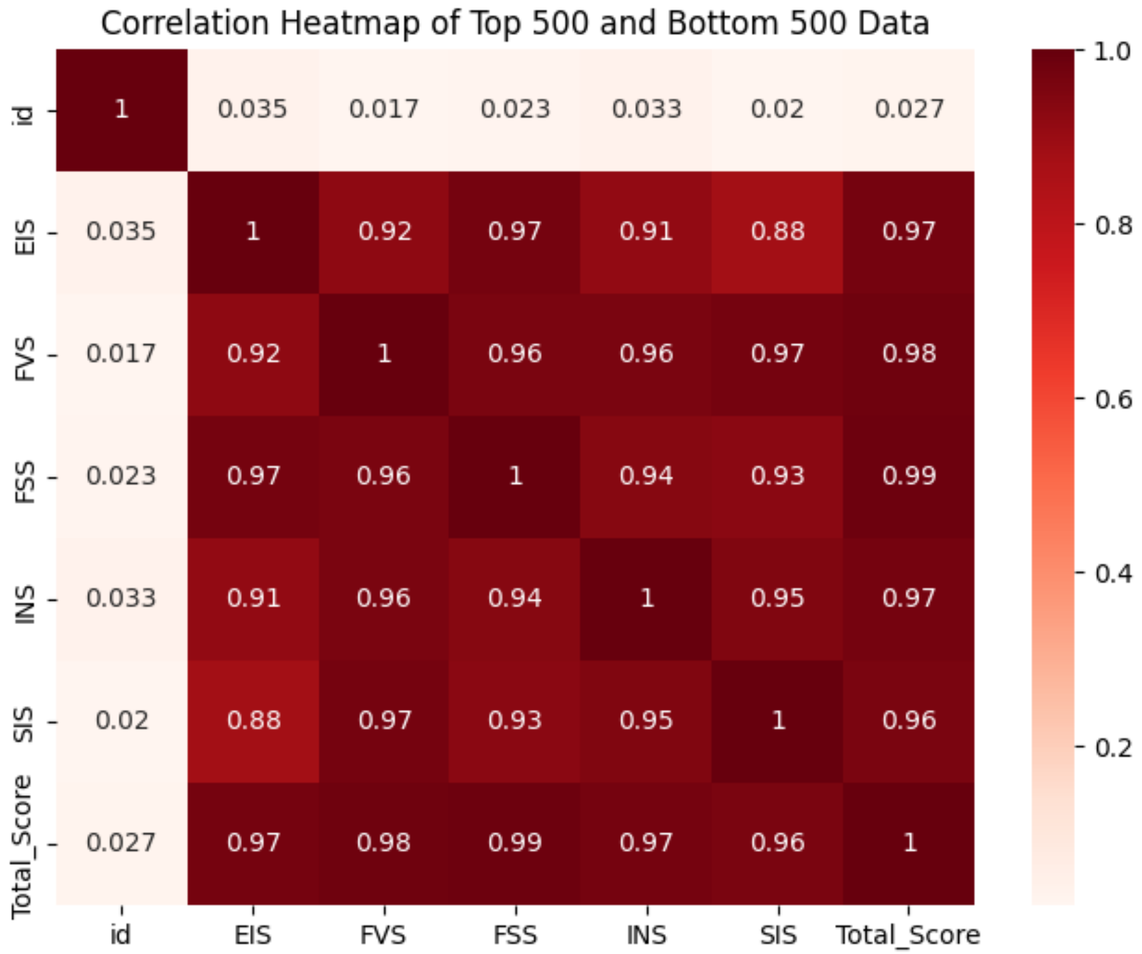
In the above Violin Graph, it can be clearly seen for the Top 500 Data, the Median is close to 0.75, whereas for the Bottom 500, it is close to 0.68 as the data in bottom 500 has different Total_Score which decreases the median value. Also the Black line interprets how the data is distributed, so for the Bottom 500 the minimum value is 0.0 & whereas maximum value is around 0.74.

Bar Graph

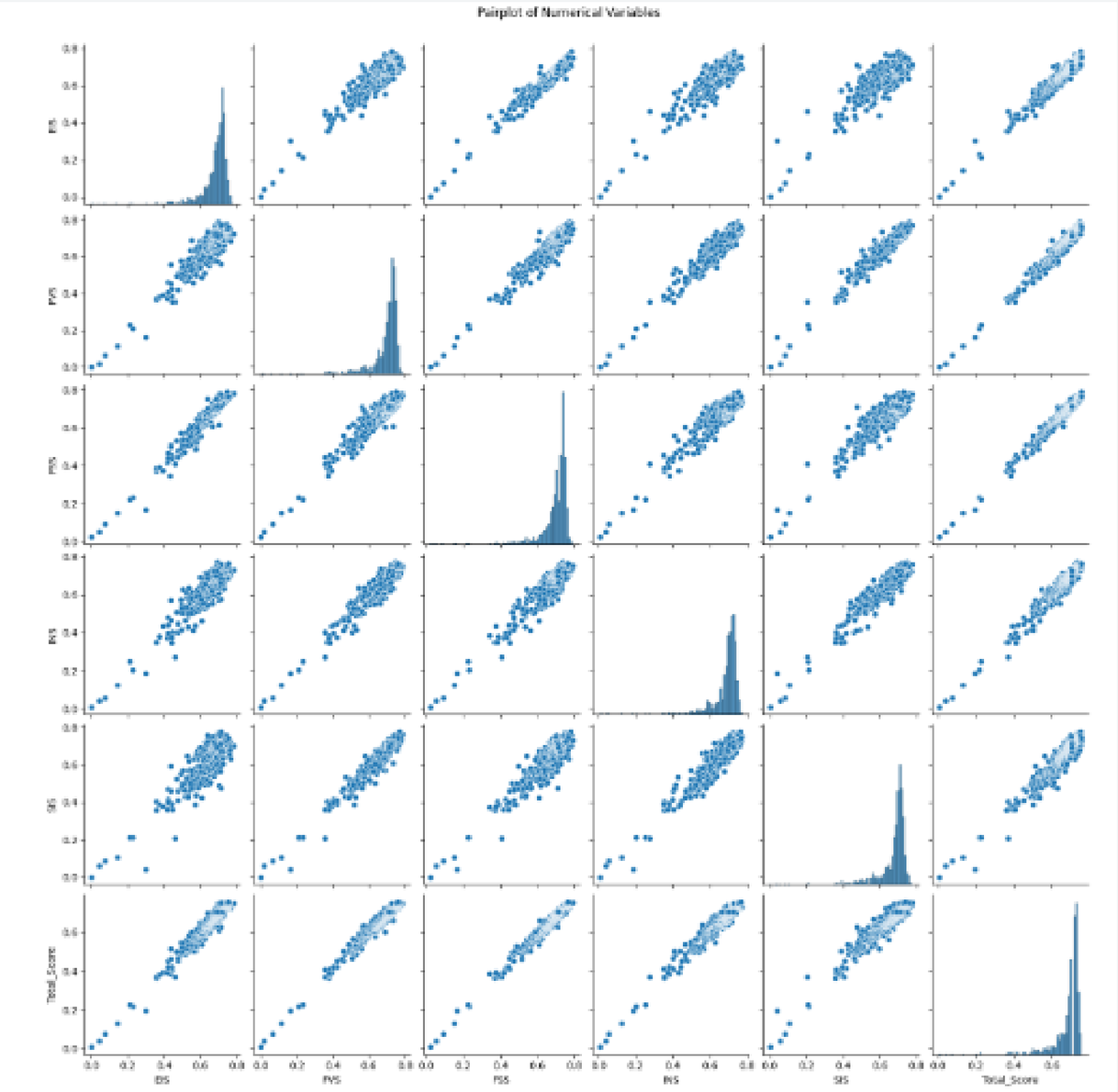


Through this bar graph, we can conclude that Mean value for the Top 500 is more than 0.72 and the same value for the Bottom 500 is 0.65

Heatmap



From the above heatmap, we can conclude that Total_Score has the highest correlation between Environmental Impact Score (EIS) ,Financial Viability Score (FVS), Feasibility & Scalability Score (FSS), Innovation & Novelty Score (INS) & Social Impact Score (SIS) with value 0.97, 0.98, 0.99, 0.97 & 0.96 respectively.



Pairplot graph allow us to plot pairwise relationship between variables, which in our case is EIS, FVS, FSS, INS, SIS & Total_Score.

Pairplot Graph

Thank You