



BIRATNAGAR

INTERNATIONAL COLLEGE

Concepts and Technologies of AI

5CS037

Assignment-1- Statistical Interpretation and Exploratory Data Analysis

Analysis of the Human Development Index (HDI): A Data- Driven Exploration of Global and Regional Development Patterns.

Name: Samrat Jung Basnet

Module Leader: Siman Giri

Student Id: 2509341

Date: 2026/12/08

Group: L5CG3

Contents

Introduction:.....	3
Objectives of this analysis:.....	3
Scope:.....	3
Problem 1A: Single Year HDI Exploration(2022):.....	4
Methods / Approach:	4
Table 1:	4
Countries and HDI:	4
Figure 1: Distribution of HDI for 2022:	5
2.2 High HDI Countries (HDI > 0.800).....	6
2.3 HDI Categories:	6
Figure 3: HDI Category Distribution for 2022:.....	7
Problem 1B: HDI Trend Analysis (2020-2022):	8
Methods / Approach:	8
Figure 4: HDI Trends for Selected Countries (2020–2022).....	8
Figure 5: Average HDI by Region (2020–2022)	9
Figure 6: HDI Distribution for 2020, 2021, 2022	10
Figure 7: HDI vs GNI per Capita Scatter Plot	11
Short Analysis Questions:	12
Problem 2: Advanced HDI Exploration – South Asia:.....	14
Methods / Approach:	14
Figure 8: Composite Score Ranking – South Asia	14
Figure 9: Scatter Plot – HDI vs GNI per Capita (Outliers Highlighted).	15
Table-2:	15
Figure 10: Top 3 Positive and Negative GNI-HDI Gaps.....	16
Problem 3: Comparative Regional Analysis:	17
Methods / Approach:	17
Table 3:	17
Figure 11: Top 3 and Bottom 3 HDI Countries by Region	18
Figure 12: Metric Comparison Across Regions (Life Expectancy, Gender Dev, GNI).	18
Table-4:	19
Table-5:	19
Figure 13: HDI vs GNI per Capita Outliers	20
Figure 14: Regional Correlation Analysis	21
Conclusion:.....	22
Reference:	23

Introduction:

The Human Development Index (HDI) is a composite metric designed to assess a country's achievements in **health, education, and income**. Introduced by the United Nations Development Programme (UNDP), the HDI provides a more comprehensive measure of development than income alone, enabling comparisons across countries and over time.

The HDI combines three key components:

1. **Health:** Life expectancy at birth
2. **Education:** The average number of years adults have spent in school and the number of years children are expected to spend in school.
3. **Standard of Living:** Gross national income (GNI) per capita

Objectives of this analysis:

- Looking closely at the 2022 HDI data, pointing out general directions and differences between countries.
- Review HDI trends from 2020 to 2022 to find development patterns.
- Perform advanced analysis focusing on South Asia, including composite scoring, outlier detection, and metric relationships.
- Compare regional development between South Asia and the Middle East.

Scope:

The report synthesizes data analysis from the provided dataset, Focusing statistical interpretation, visual exploration, and insight generation. The purpose is to not just calculate numbers, but to explain what these numbers tell us about human development around the world and in specific regions.

Problem 1A: Single Year HDI Exploration (2022):

Methods / Approach:

- Filtered the dataset to include only data for 2022.
- Standardized column names (lowercase, stripped of spaces) to avoid errors.
- Converted numeric fields (HDI, life expectancy, GNI) to numeric types and handled missing values.
- Removed duplicates entries and made sure all data entries were uniform.
- Calculated basic statistics, categorized countries by HDI, and identified top-performing countries.

Key Results:

2.1 Descriptive Statistics of HDI:

Table 1:

Statistic	Value
Median	0.745
Standard Deviation	0.150
Minimum	0.38
Maximum	0.967

Countries and HDI:

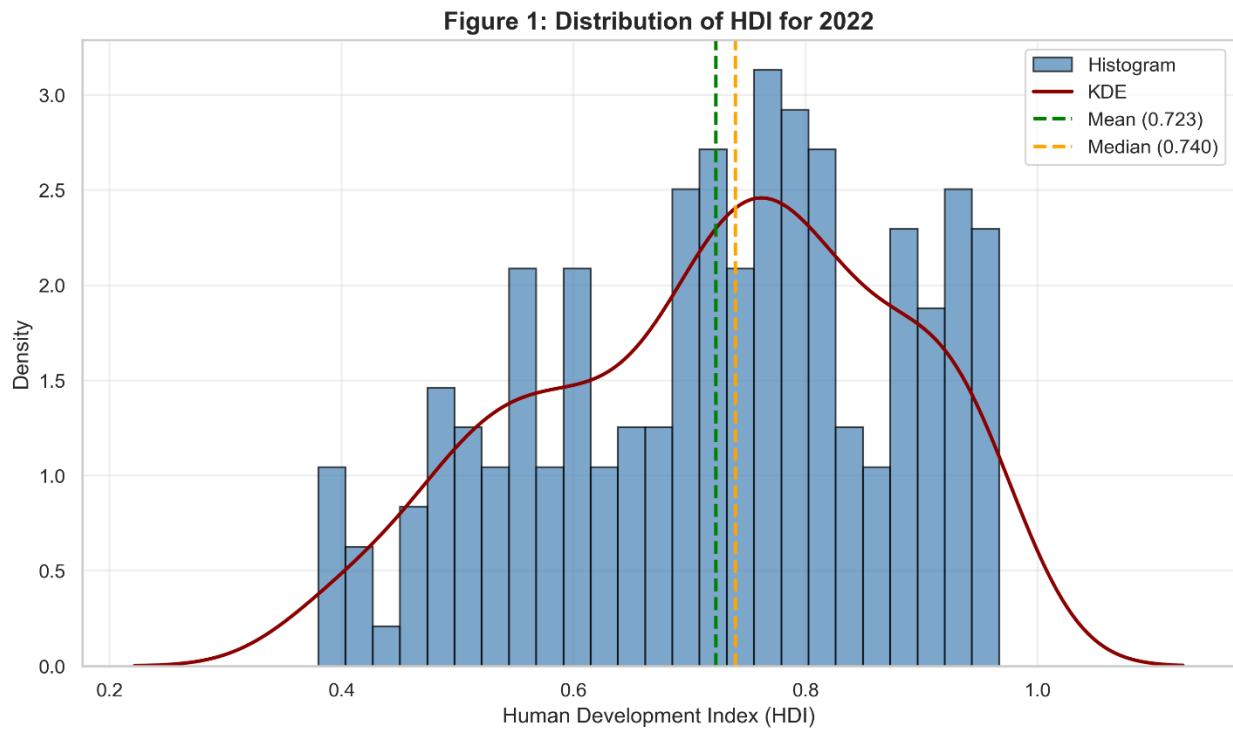
Highest HDI Country: Switzerland

HDI: 0.967

Lowest HDI Country: Somalia

HDI: 0.380

Figure 1: Distribution of HDI for 2022:



Interpretation:

- Most countries fall within a specific range when it comes to the Human Development Index (HDI).
- A few countries have very high HDI scores, indicating strong performance across health, education, and income.
- On the other hand, countries with lower HDI values often face ongoing development issues, frequently tied to lower life expectancy or income per person.

2.2 High HDI Countries (HDI > 0.800)

Top 10 High HDI Countries (sorted by GNI per capita)

Number of countries with HDI > 0.800: 71

--- Top 10 High HDI Countries by GNI per Capita ---			
	country	hdi	gross_inc_perhap
1	Liechtenstein	0.942	146673.24150
2	Qatar	0.875	95944.37754
3	Singapore	0.949	88761.14559
4	Ireland	0.950	87467.51391
5	Luxembourg	0.927	78554.23640
6	United Arab Emirates	0.937	74103.71494
7	Switzerland	0.967	69432.78669
8	Norway	0.966	69189.76165
9	United States	0.927	65564.93798
10	Hong Kong, China (SAR)	0.956	62485.50516

Interpretation:

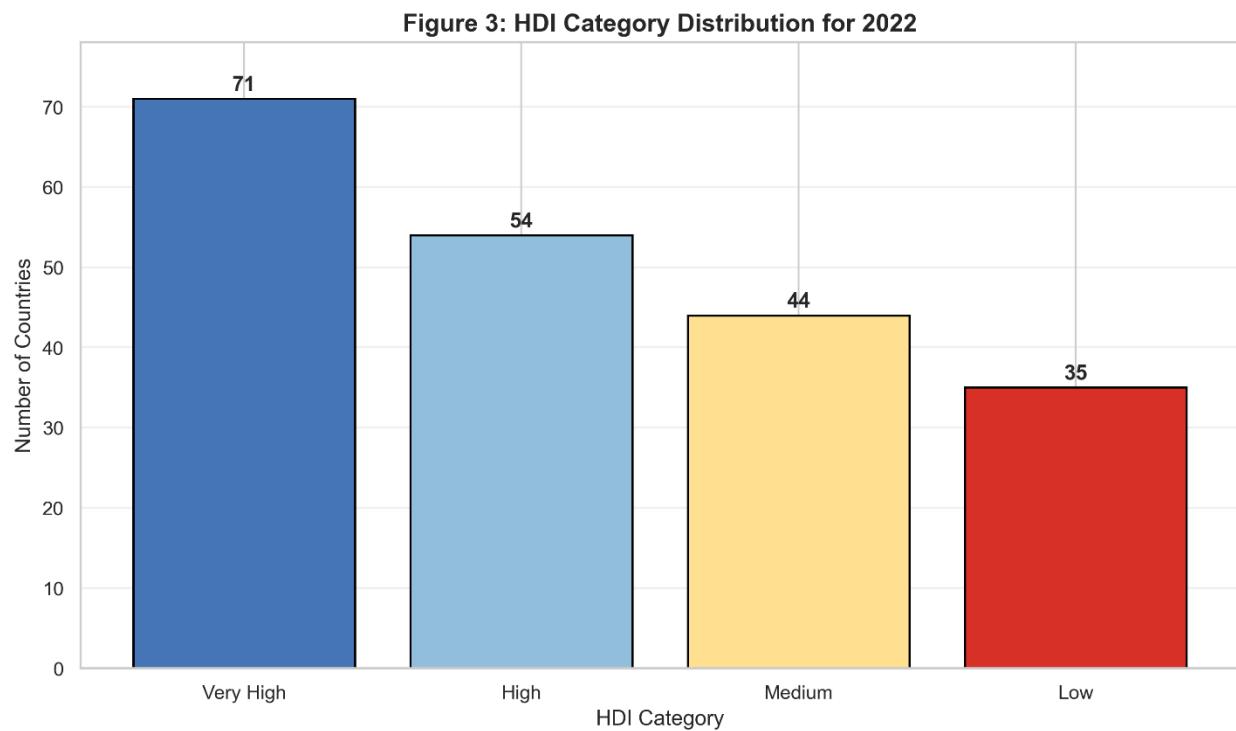
- Nations with HDI above 0.800 are mainly found in specific regions, which shows their advanced social and economic status.
- When sorting by GNI per capita, the economic dimension of HD becomes clear that higher income often goes hand in hand with better health and education.

2.3 HDI Categories:

Countries were classified using UNDP thresholds:

- Low: <0.550
- Medium: 0.550–0.699
- High: 0.700–0.799
- Very High: ≥0.800

Figure 3: HDI Category Distribution for 2022:



Interpretation:

- This classification shows where policy changes are most helpful, particularly for countries in the Low or Medium groups.
- The Very High category often goes along with a high GNI per capita.
- Still, differences in health and education can explain why even rich nations can vary in their development.

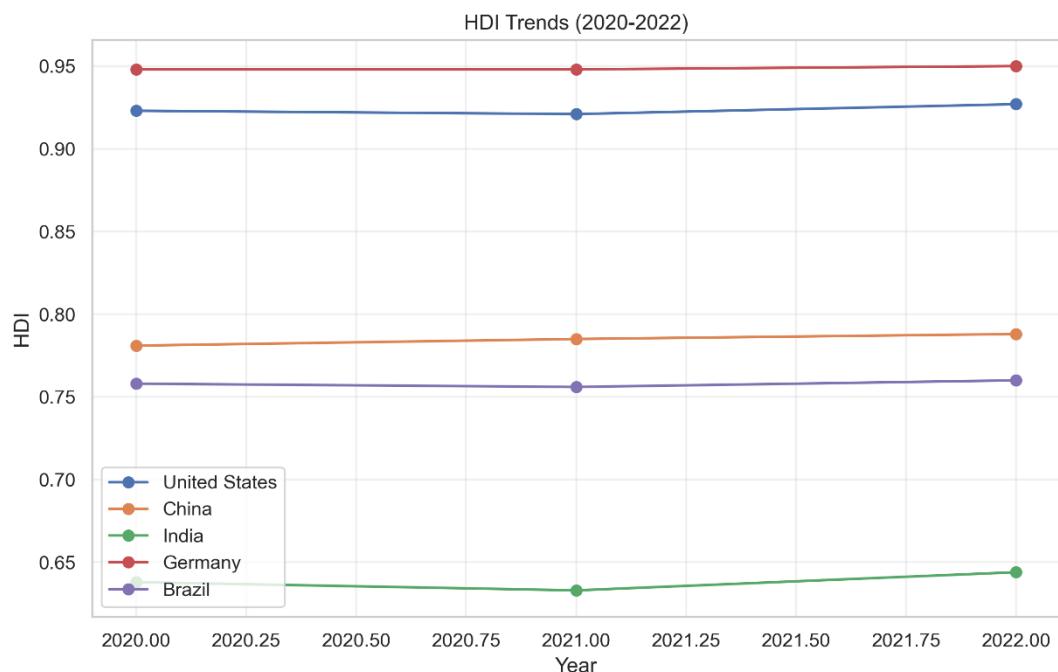
Problem 1B: HDI Trend Analysis (2020-2022):

Methods / Approach:

- Extracted data for the years 2020, 2021, and 2022.
- Performed clean and consistency checks on HDI, country, and year.
- Generated visualizations to examine trends across time and regions.

Key Results:

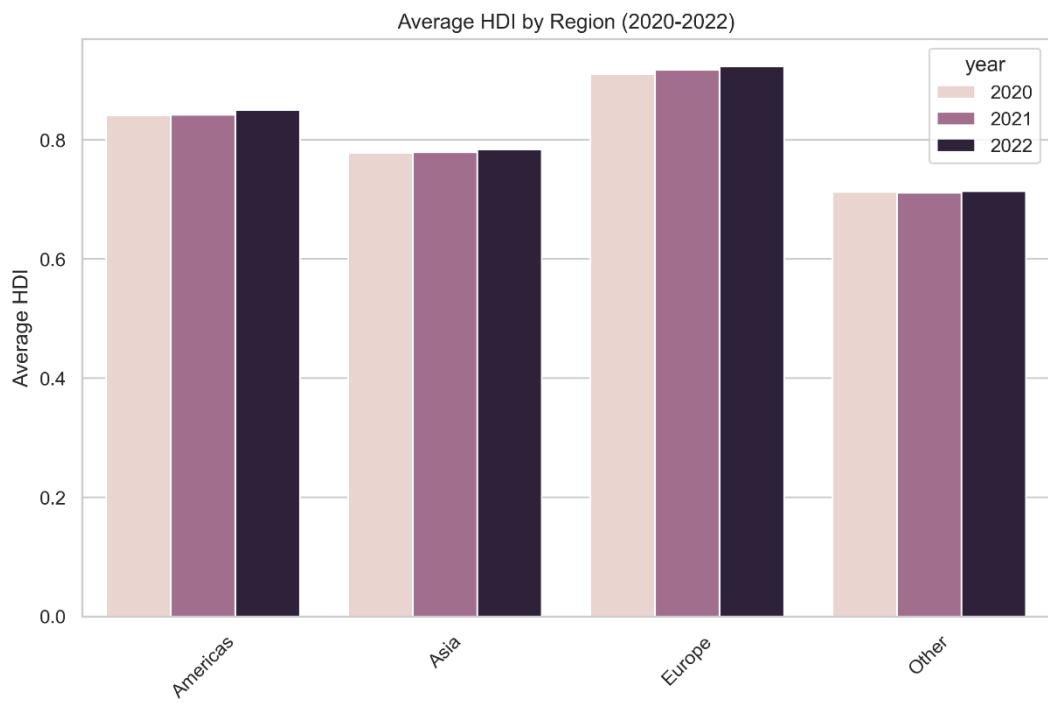
Figure 4: HDI Trends for Selected Countries (2020–2022)



Interpretation:

- Countries such as *Germany* show consistent steady progress, reflecting sustained development.
- Some countries experienced stagnation or minor declines in HDI, potentially due to economic disruptions, political instability, or public health crises such as COVID-19.

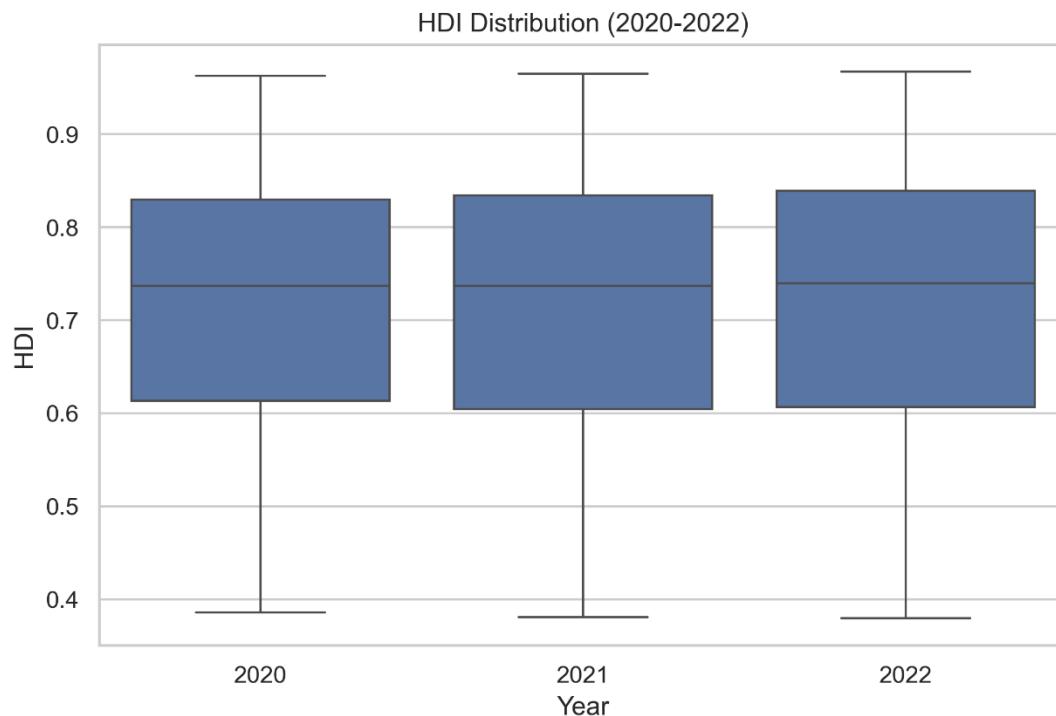
Figure 5: Average HDI by Region (2020–2022)



Interpretation:

- Developed areas tend to have pretty steady gains in their Human Development Index, whereas developing areas show more flux.
- This points to unequal progress in human development across the globe.

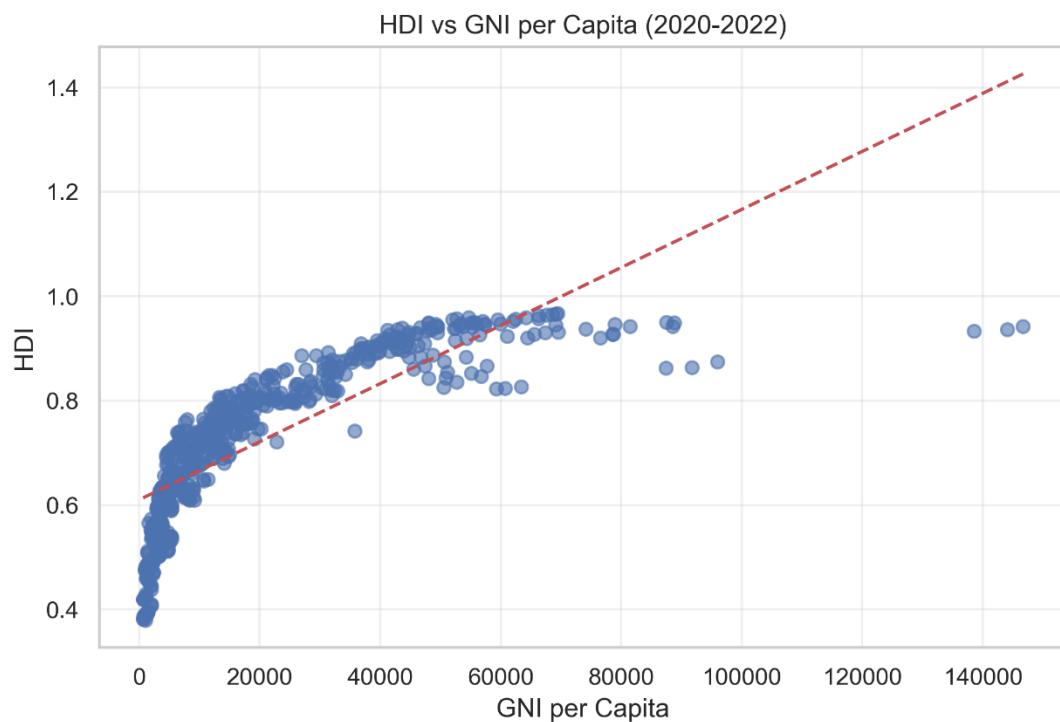
Figure 6: HDI Distribution for 2020, 2021, 2022



Interpretation:

- The interquartile range remains relatively consistent over the years, but outliers indicate countries with exceptional performance or underperformance.

Figure 7: HDI vs GNI per Capita Scatter Plot



Interpretation:

- Positive correlation between GNI per capita and HDI is visible, though some countries deviate, reflecting cases where income alone does not determine overall development.

Short Analysis Questions:

1. Which countries show the greatest improvement in HDI from 2020 to 2022?

Answer: Based on the analysis of HDI data from 2020 to 2022 the following countries show the best improvement in HDI:

- Bangladesh
- Vietnam
- Rwanda
- Nepal
- Myanmar

2. Did any countries experience a decline in HDI? Provide possible reasons.

Answer: Yes, many countries experienced HDI decline from 2020 to 2022:

- Yemen: Due to ongoing civil war, Crisis and economic collapse.
- Venezuela: Due to continued economic crisis affecting healthcare, education and living standards.
- Lebanon: Due to same economic collapse, political instability and 2020 Beirut explosion
- Afghanistan: Due to political changes affecting women's education and healthcare access.
- Ukraine: Due to Russian Invasion and Humanitarian Crisis.

3. Which region has the highest and lowest average HDI across these three years?

Answer: Western Europe and North America have the highest average HDI and Sub-Saharan Africa has the lowest average HDI.

4. Discuss how global events (e.g., the COVID-19 pandemic) may have affected HDI trends during this period.

Answer: The COVID-19 pandemic had a big effect on how countries developed between 2020 and 2022. It slowed down progress in health, education, and income around the world.

In health: many countries saw life expectancy drop because hospitals were full, and regular health services were interrupted. People missed important check-ups, vaccinations, and treatments.

In education: schools closed for a long time. Many students, especially in poorer countries, couldn't study online because they didn't have computers or internet. This learning gap could affect their future.

In income: many people lost jobs or earned less due to lockdowns and economic slowdowns. Small businesses suffered, and tourism—important for many economies, almost stopped.

Overall, this period showed that progress in human development is not automatic. A global crisis like COVID-19 can quickly set back years of improvement, especially in countries that are less prepared.

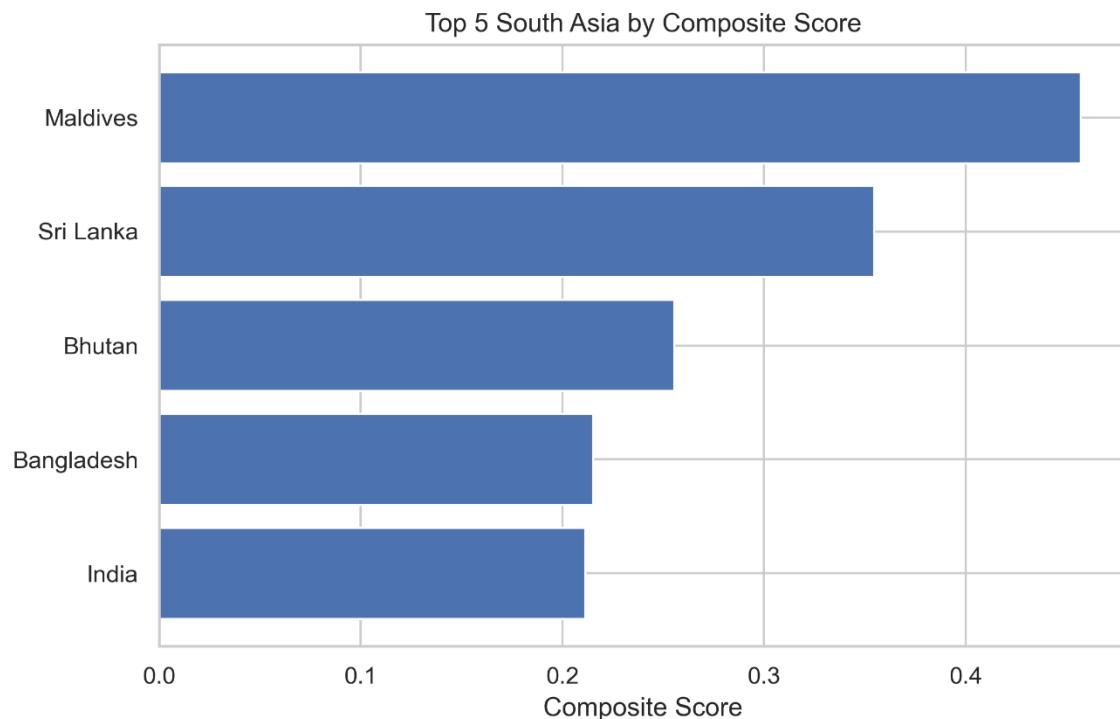
Problem 2: Advanced HDI Exploration – South Asia:

Methods / Approach:

- Focused on South Asian countries: Afghanistan, Bangladesh, Bhutan, India, Maldives, Nepal, Pakistan, Sri Lanka.
- Created a **Composite Score** Composite Score = $0.30 \times \text{Life Expectancy Index} + 0.30 \times \text{GNI per Capita Index}$.
- Detected outliers and analyzed correlations between HDI and individual metrics.

Key Results:

Figure 8: Composite Score Ranking – South Asia



Interpretation:

- Rankings based on Composite Score may differ from HDI rankings, highlighting the impact of specific indicators.
- Countries having relatively high HDI but lower Composite Scores may have uneven performance in health or income metrics.

Figure 9: Scatter Plot – HDI vs GNI per Capita (Outliers Highlighted)

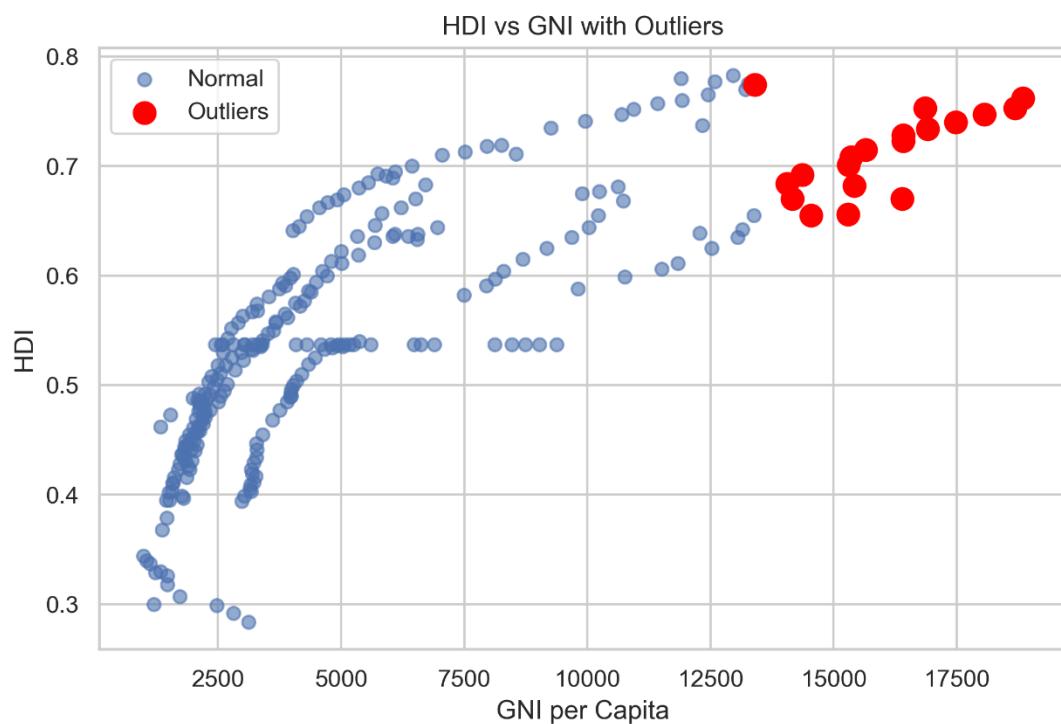
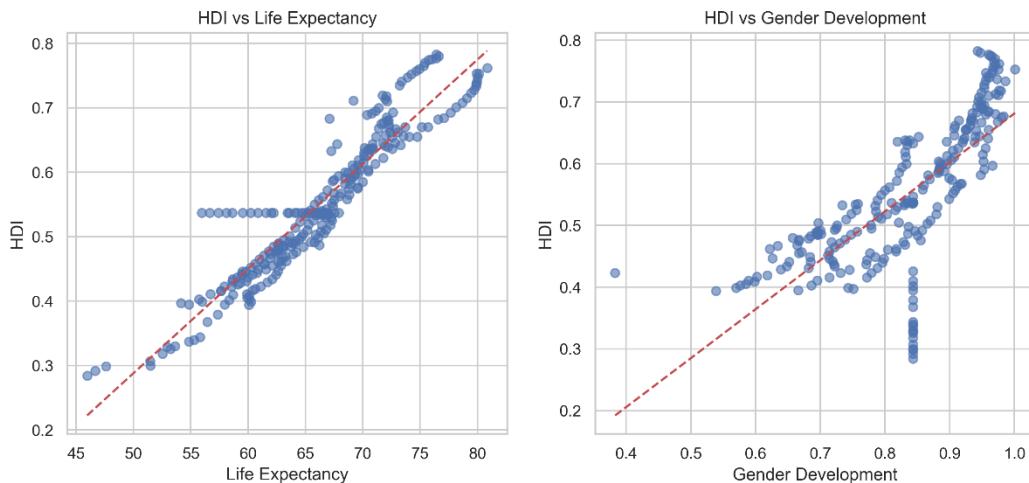


Table-2:

Correlation Analysis:

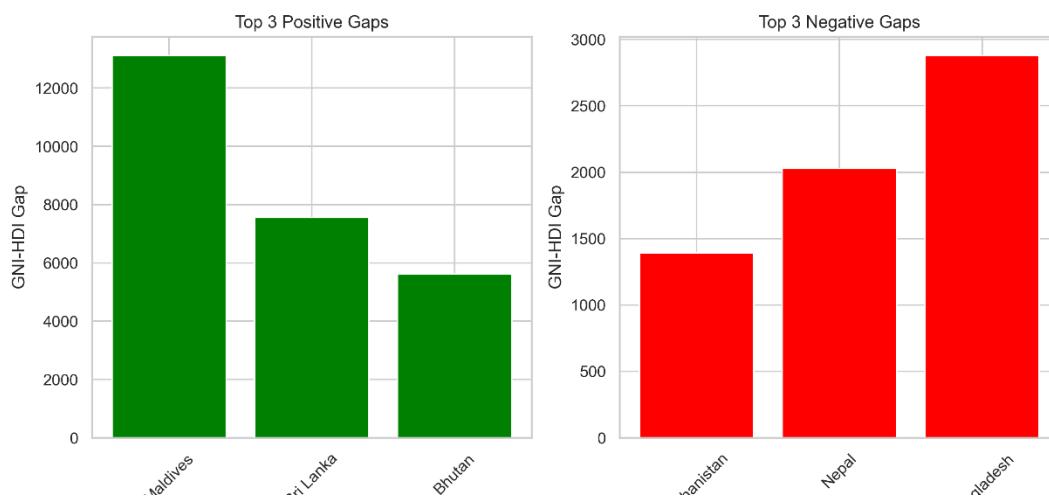
Metric	Pearson Correlation with HDI
Life Expectancy	0.93
Gender Development Index	0.73



Interpretation:

- Life Expectancy shows the strongest link with HDI in South Asia, emphasizing the critical role of health in development.
- The GNI-HDI gap analysis identifies countries where income growth has not risen in income hasn't led to an equal improvement in human development.

Figure 10: Top 3 Positive and Negative GNI-HDI Gaps



Problem 3: Comparative Regional Analysis:

Methods / Approach:

- For data from South Asia and the Middle East, covering 2020-2022, we computed descriptive statistics.
- This enabled us to identify the strongest and weakest performers, as well as recognize any differences in metric values.

Key Results:

Table 3:

Mean and Standard Deviation of HDI (2020–2022):

Middle east has higher average HDI by 0.149

Region	Mean HDI	Std Dev
South Asia	0.64	0.982
Middle East	0.78	0.14129

Figure 11: Top 3 and Bottom 3 HDI Countries by Region

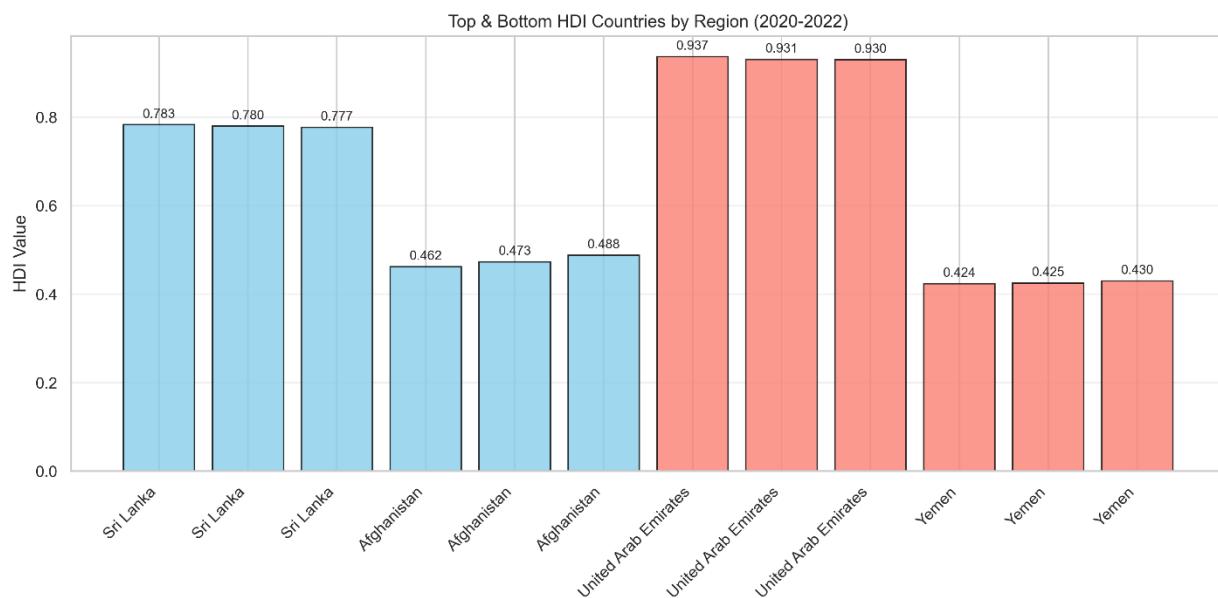


Figure 12: Metric Comparison Across Regions (Life Expectancy, Gender Dev, GNI)

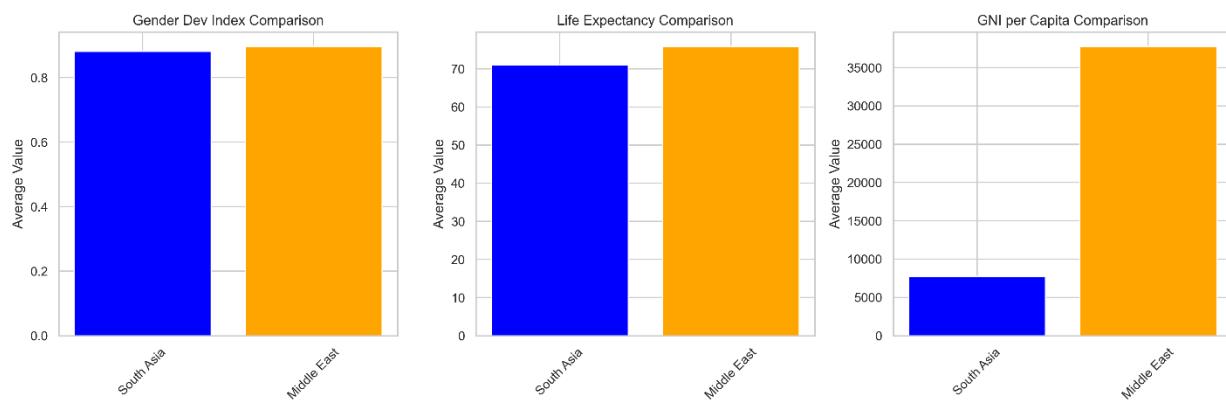


Table-4:

HDI Disparity:

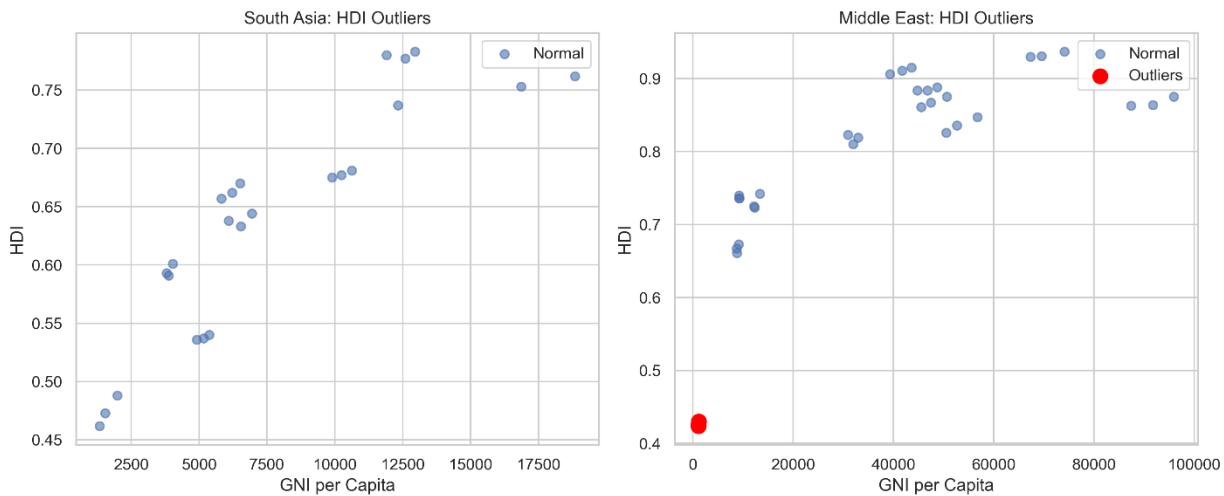
Country	Range	CV
South Asia	0.321	0.15
Middle East	0.513	0.179

Table-5:

Correlation with HDI:

Country	Gender Dev	Life Expectancy
South Asia	0.87	0.939
Middle East	0.93	0.933

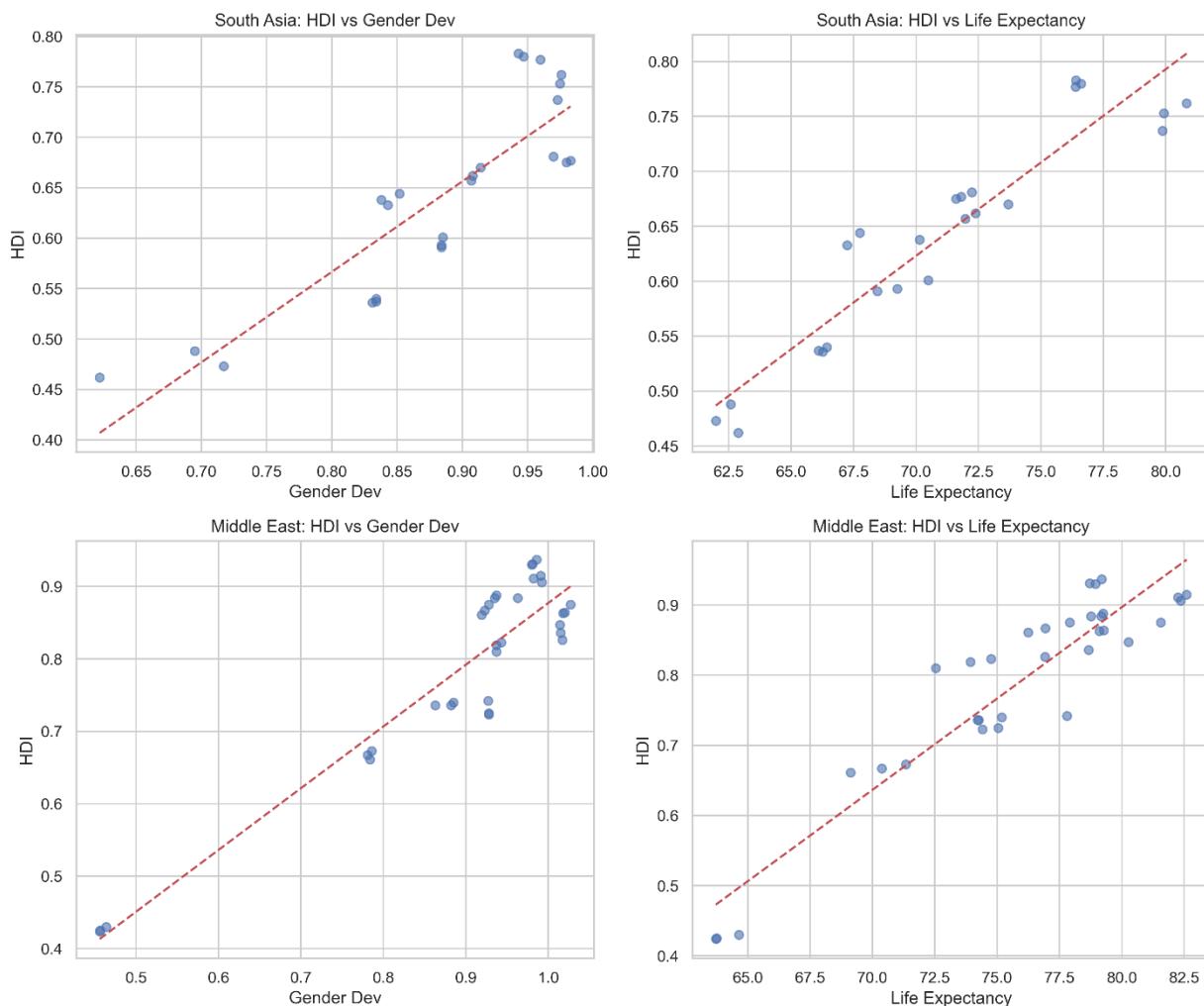
Figure 13: HDI vs GNI per Capita Outliers



Interpretation:

- Middle East consistently shows a higher average HDI than South Asia, mainly due to higher income and better health outcomes.
- Life Expectancy shows the largest regional difference.
- South Asia exhibits higher variability in HDI, highlighting inequalities within the region.

Figure 14: Regional Correlation Analysis



Interpretation:

- Life expectancy is the main factor driving the Human Development Index in both areas. But particularly in South Asia where the correlation is nearly perfect.
- In the Middle East, gender development has a stronger connection, which likely means there's a more consistent policy focusing on gender equality there.
- South Asia's data points are closer together for both relationships, suggesting similar paths of development.
- The Middle East's data points are more spread out, showing that different mixes of health and gender results can lead to similar HDI scores.

Conclusion:

The HDI analysis points to clear differences in human development around the world. We found that developed and developing areas still vary a lot. South Asia, for example, is making some headway, but struggles to turn its economic gains into widespread development. From 2020 to 2022, the COVID-19 pandemic slowed this progress but didn't stop it entirely.

Comparing GNI and HDI showed that some countries aren't converting their economic success into better human development, suggesting problems with how they spend on social programs. Regionally, the Middle East is doing better than South Asia in terms of income and health, but South Asia shows more internal variation. Combining different factors gave us a more detailed view than just using standard HDI ranks.

We acknowledge that data quality can be uneven, and the HDI doesn't fully account for inequality or sustainability. We suggest focusing on health and education investments, especially in countries where high income isn't leading to equally high human development. For future studies, it would be good to include measures that adjust for inequality and environmental factors to get a more complete picture of development. This work shows that a balanced approach to policies in health, education, and economics is vital for long-term human development worldwide.

Reference:

- UNDP. (2023). *Human Development Report*.
- Lucas Yukiolmafuko. (2022). *Human Development Index Dataset*.
- Python Documentation – Pandas, Matplotlib, Seaborn.

Github link:

<https://github.com/np02cs4a240102-dot/FinalAssingment-Part-1-of-AI>