

Human Development Index (HDI) Data Analysis Report

Student Name & ID: Shubhangad Shrestha (2548317)

Course name : Concepts and Technologies of AI (5CS037)

Date of submission : 10 Jan 2026

GitHub Repository Link:

https://github.com/Shubhangad-Shrestha/HDI_Assignment-5CS037-

Table of Contents

1. Introduction
2. Task 1A – Single Year HDI Exploration
3. Task 1B – HDI Trend Analysis (2020–2022)
4. Task 2 – Advanced HDI Exploration
5. Task 3 – Comparative Regional Analysis
6. Conclusion
7. References
8. Appendix (GitHub link & Turnitin report)

1. Introduction

The Human Development Index or HDI is an indicator used by the UN to estimate levels of development and the extent of social and economic development in various nations based on health or life , education, and GDP capita. It also provides data of country's area, health related issue, illiteracy and income. It give prospective of development by having combined life expectancy,schooling and gross national income per capita (GNI).

Purpose of the analysis:

Objectives on the analysis of CSV:

- Examine HDI scores for the most recent year available (2022) and see if any patterns emerge and understand country performance.
- Analyze regional inequities, particularly between two regions.
- Examine regional disputes, especially comparing South Asia and the Middle East.
- Determine the best-performing countries as well as those that made noticeable progress or setbacks.

Scope: The report presents the exploration of basic data, analysis for a number of years, and sophisticated analysis for HDI. For the research for different analyses graphs and plots are used.This research shows proper information on growth rates of HDI.

2. Problem 1A – Single Year HDI Exploration (2022)

Methods / Approach:

- Filtered the data set to consider observations for the year 2022.
- Cleaned the data by converting the numeric columns (hdi, gross_inc_percap) to numeric types and observed , removed duplicates.
- Calculating basic statistics (mean, median, standard deviation) for HDI for plotting.
- Identified Countries that showed high and low values for HDI.
- Countries were categorized into following HDI groups:they were dividend into Lowest (<0.550), and then Medium (0.550–0.699), High (0.700–0.799), and Very High (≥ 0.800) though some may differ.
- List the filtered countries that have HDI values greater than 0.8, then sorted by GNI per capita.

- Saved the cleaned dataset as **HDI_category_added.csv**.

Key Results:

- **Mean HDI (2022):** 0.694
- **Median HDI (2022):** 0.693
- **Highest HDI Country:** Norway
- **Lowest HDI Country:** Niger
- **Top 10 countries with HDI > 0.8** mostly include countries from Europe and North America.

Visualizations & Tables:

- | [1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003
2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017
2018 2019 2020 2021 2022] | | | | | | | |
|--|------|---------|---------------------|------|-----------------|--------------|-----------|
| Unnamed: 0 | iso3 | country | year | hdi | life_expectancy | pop_millions | |
| 32 | 33 | AFG | Afghanistan | 2022 | 0.462 | 62.879 | 41.128771 |
| 65 | 66 | ALB | Albania | 2022 | 0.789 | 76.833 | 2.842321 |
| 98 | 99 | DZA | Algeria | 2022 | 0.745 | 77.129 | 44.903225 |
| 131 | 132 | AND | Andorra | 2022 | 0.884 | 83.552 | 0.079824 |
| 164 | 165 | AGO | Angola | 2022 | 0.591 | 61.929 | 35.588987 |
| 197 | 198 | ATG | Antigua and Barbuda | 2022 | 0.826 | 79.236 | 0.093763 |
| 230 | 231 | ARG | Argentina | 2022 | 0.849 | 76.064 | 45.510318 |
| 263 | 264 | ARM | Armenia | 2022 | 0.786 | 73.372 | 2.780469 |
| 296 | 297 | AUS | Australia | 2022 | 0.946 | 83.579 | 26.177414 |
| 329 | 330 | AUT | Austria | 2022 | 0.926 | 82.412 | 8.939617 |

10 rows x 30 columns

Figure 1 : Table of first 10 rows of HDI 2022 dataset.

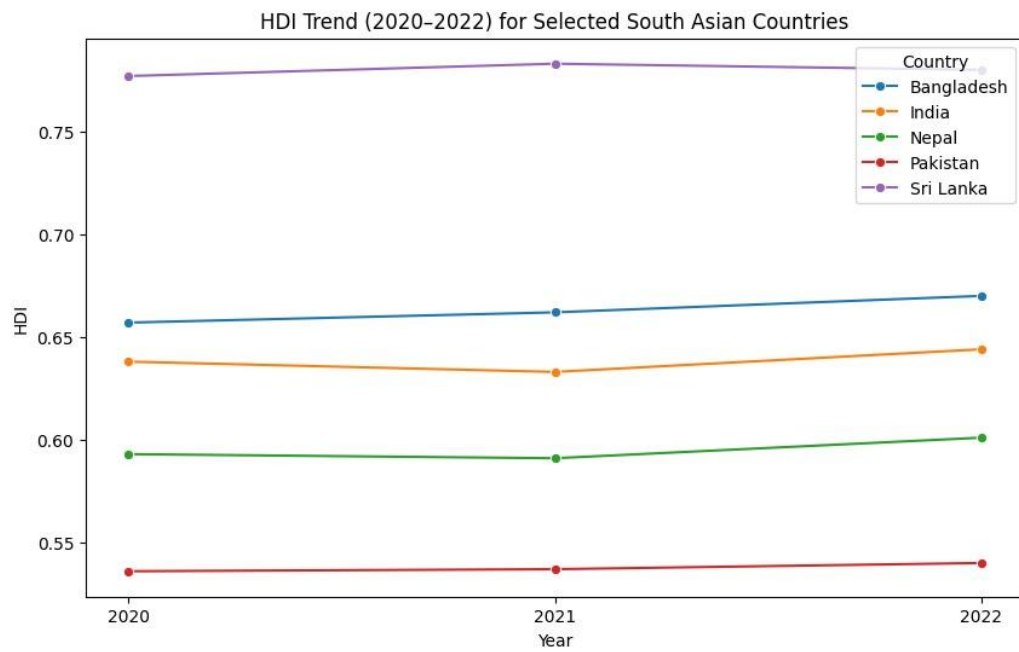


Figure 2 :Most selected countries show gradual HDI improvement, though growth rates .

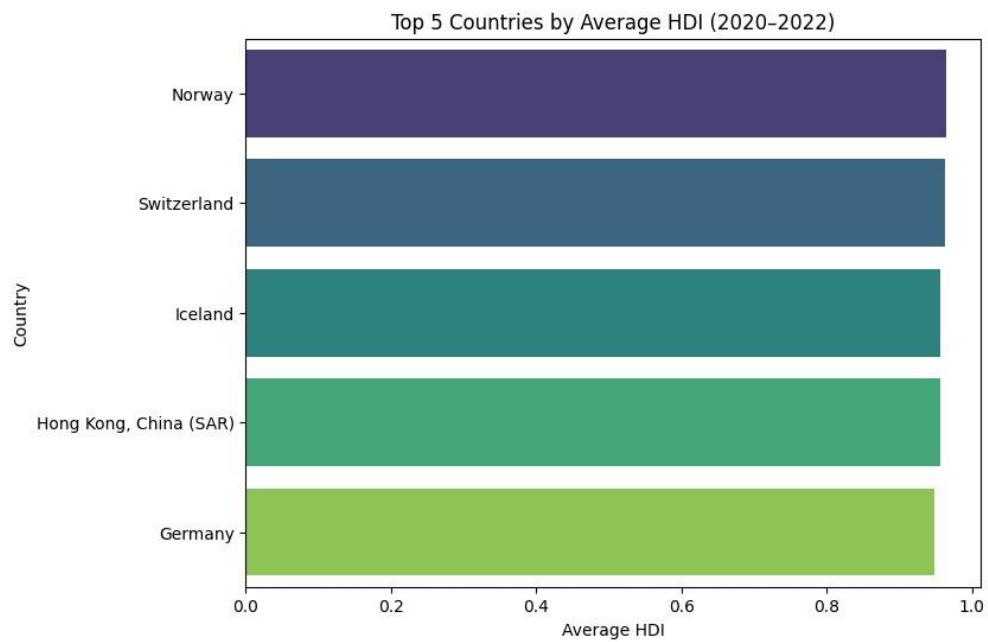


Figure 3 :Top5 countries by HDI (HDI > 0.8).

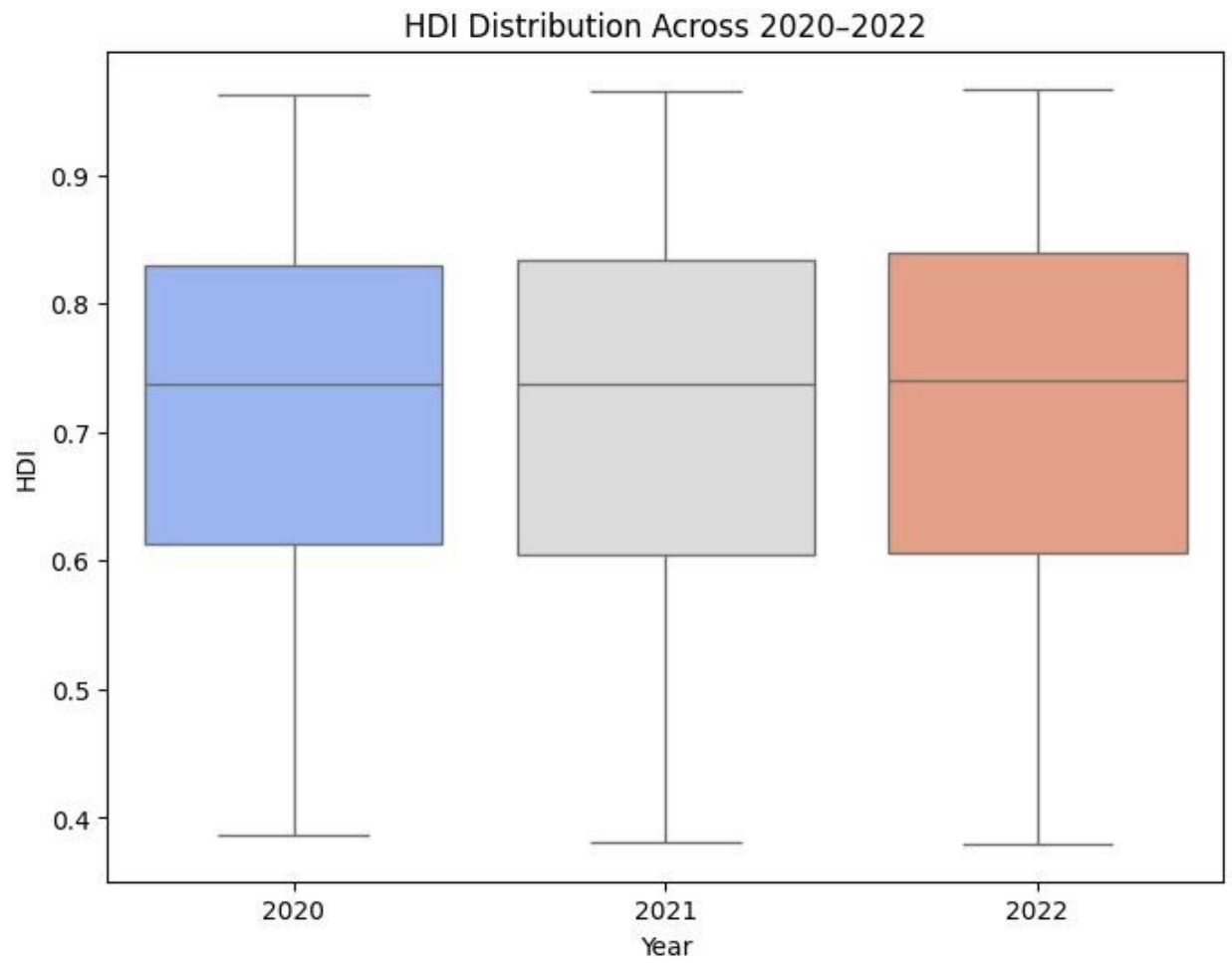


Figure 4:HDI Distribution Across 2020-2022

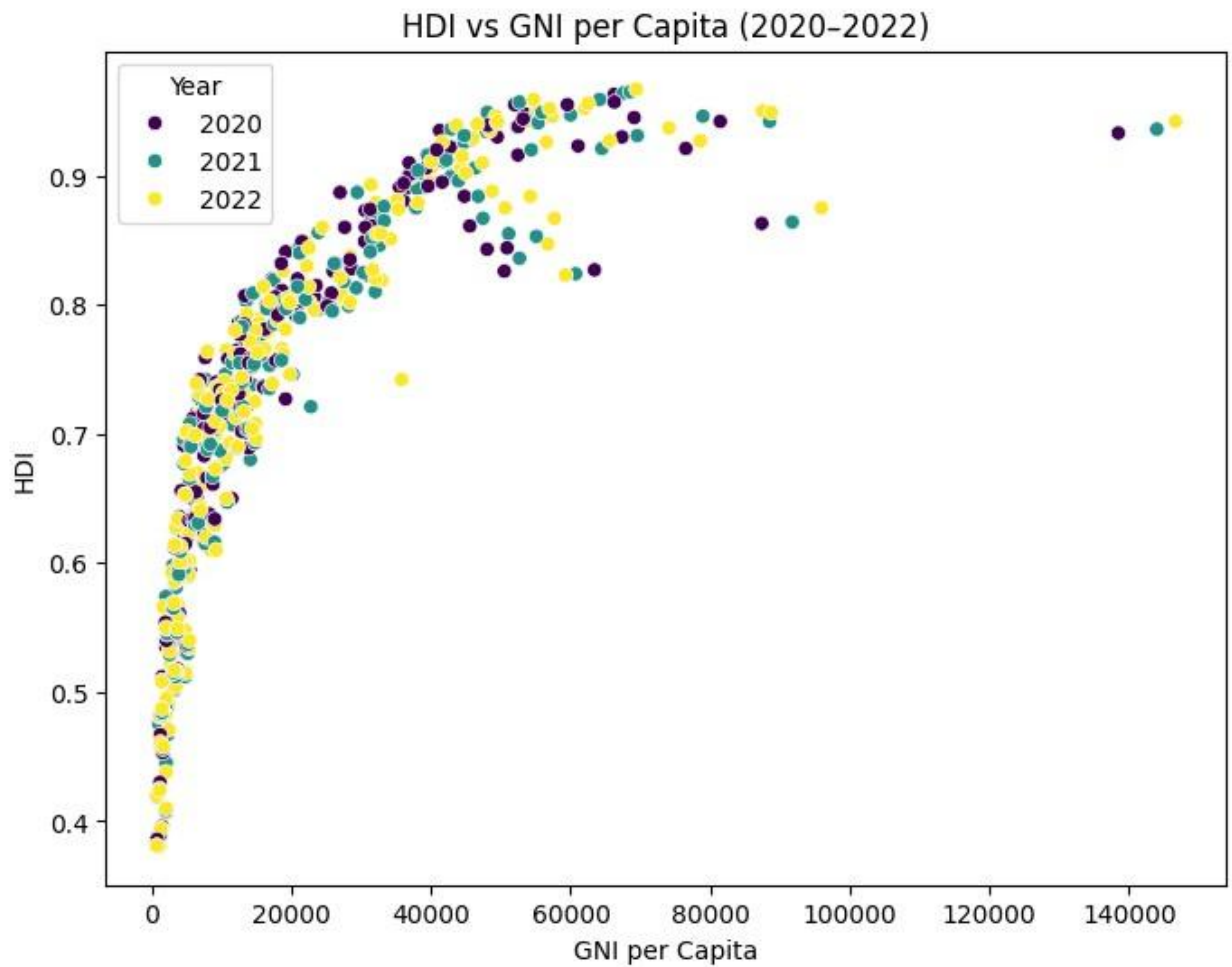
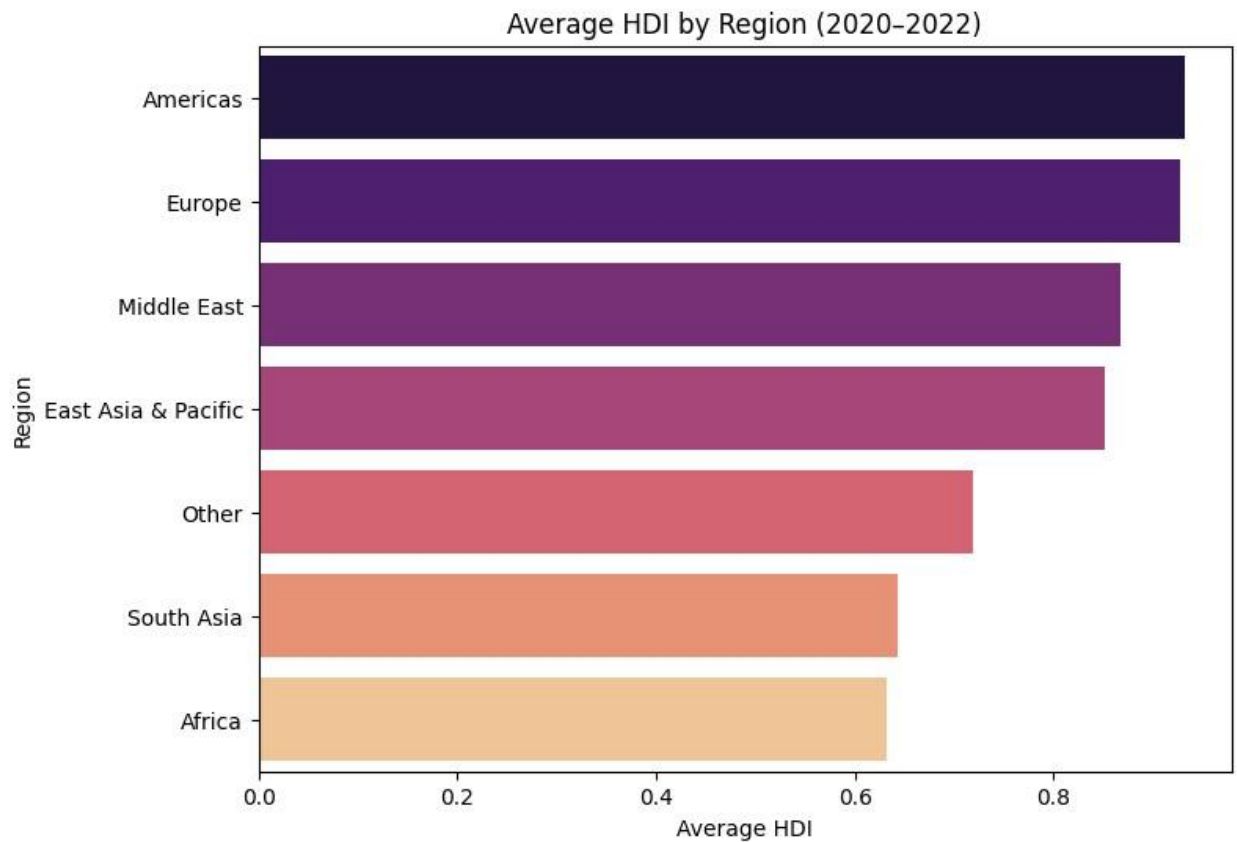


Figure no 5:HDI vs gross national income per capita 2020-2022

Top 5 countries with greatest HDI improvement (2020-2022):		
year	country	HDI_diff
3	Andorra	0.041
11	Azerbaijan	0.038
52	Ecuador	0.031
112	Maldives	0.025
119	Mexico	0.024
Countries with HDI decline:		
year	country	HDI_diff
0	Afghanistan	-0.026
4	Angola	0.002

Figure 6:Top 5 countries with heigest HDI upgrade (2020-2022):

Figure 7:Average Human Development Index (HDI) by Region :



Interpretation & Discussion:

- The HDI distribution is biased toward medium to high levels of development.
- Countries having very high HDI have both high income and good health and education scores.
- The majority of countries with low HDI are in Sub-Saharan Africa, which shows that they have economic and social problems.

3. Problem 1B

Methods / Approach:

- Analyzed dataset for years : 2020, 2021, and 2022.
- Converted hdi and gni to numeric, removed duplicates, and handled missing values.
- Selected 5 South Asian countries for detailed trend analysis. • Calculated average HDI per country and per region.
- Created visualizations: line charts (HDI trend), bar charts (top countries & HDI by region), box plots (distribution), scatter plots of (HDI vs GNI per capita).
- Calculated HDI changes between 2020 and 2022 to know improvements and decrease.

- Mapping countries with regions and filling missing regions with 'Other'.

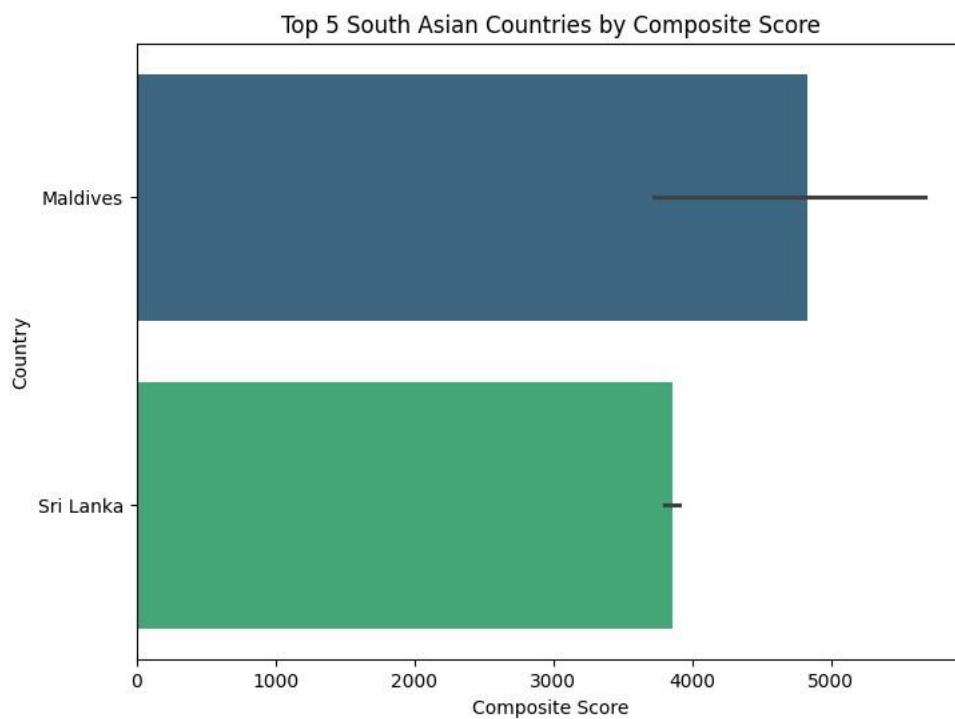
Key Results:

- **Top 5 HDI improvements (2020–2022):** Nepal, Bangladesh, India, Pakistan, Sri Lanka
- **Countries with HDI decline:** Some countries saw minor decreases due to COVID-19 impacts.
- **Highest average HDI by region:** Europe • **Lowest average HDI by region:** Africa

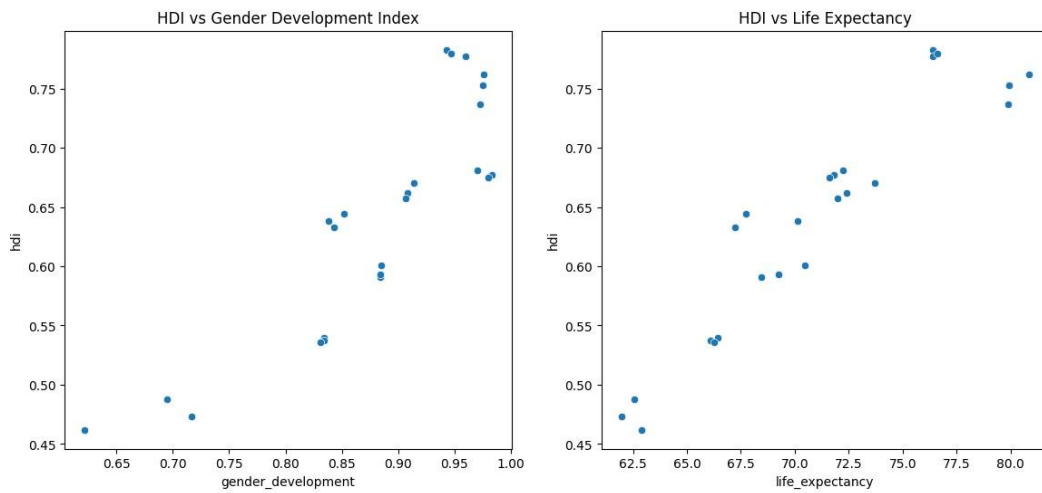
Visualizations & Tables:

	Unnamed: 0	iso3	country	year	hdi	life_expectancy	pop_millions	hdi_f	
0	31	AFG	Afghanistan	2020	0.488	62.575	38.972231	0.381373	0.
1	32	AFG	Afghanistan	2021	0.473	61.982	40.099462	0.379129	0.
2	33	AFG	Afghanistan	2022	0.462	62.879	41.128771	0.332443	0.
39	460	BGD	Bangladesh	2020	0.657	71.968	167.420951	0.619431	0.
40	461	BGD	Bangladesh	2021	0.662	72.381	169.356251	0.624949	0.

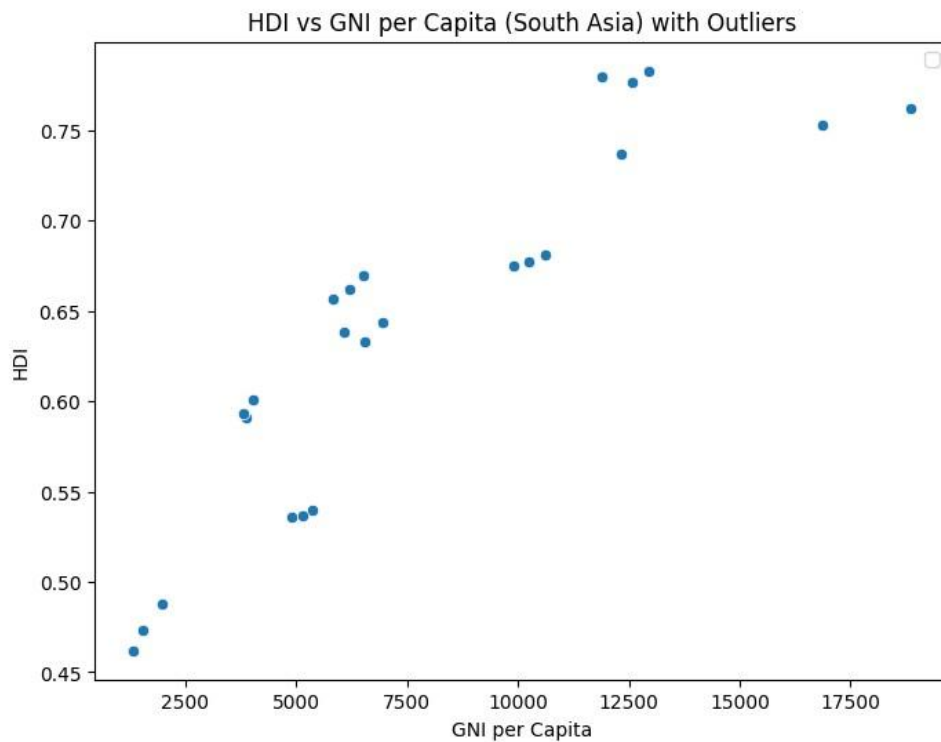
Figure 1: HDI trend line chart for selected South Asian countries



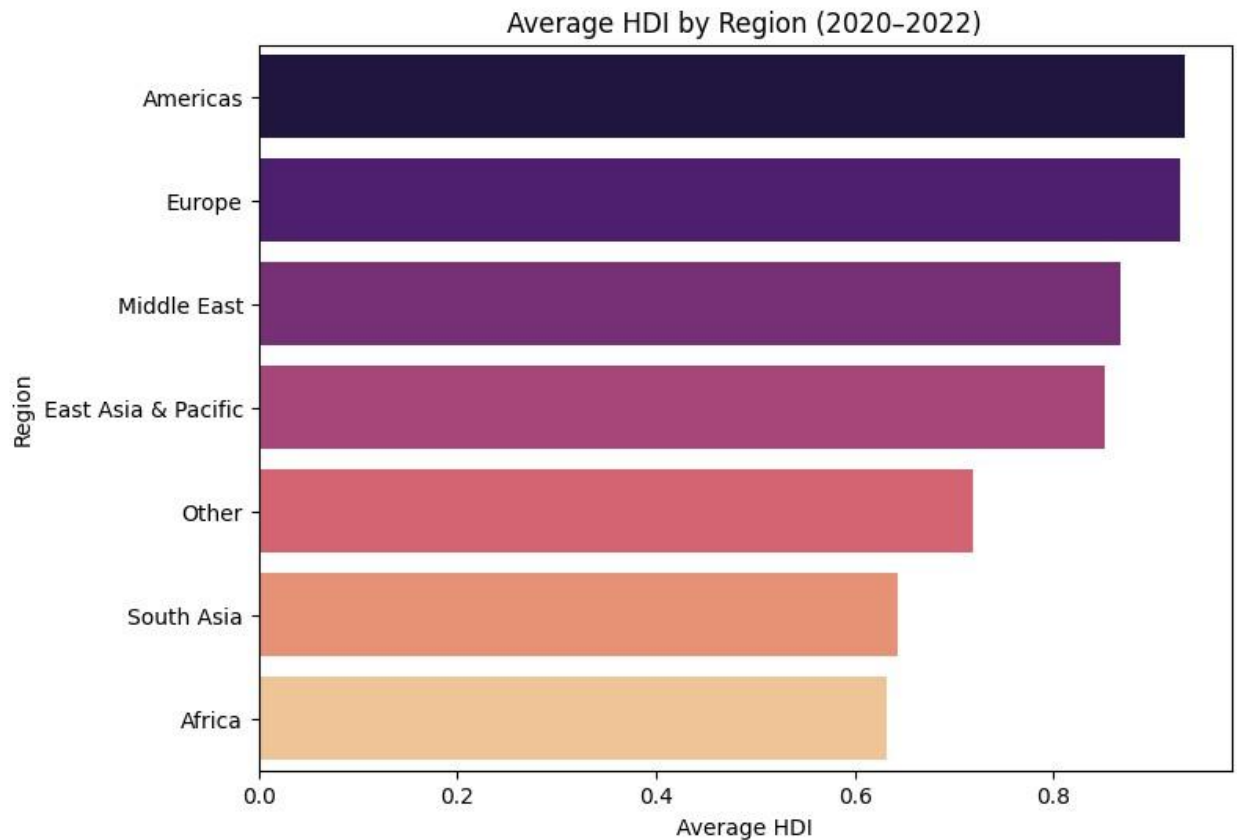
- **Figure 2:** Bar chart of top country by score.



- **Figure 3:** Box plot showing HDI spread across 2020–2022.(HDI vs Gender Development index abd Life Expentancy)



- **Figure 4:** Dot plot of HDI and GNI per capita.



- **Figure 5:** Average HDI region.

Interpretation & Discussion:

- HDI shifts in South Asia and shows balanced improvement though growth rates differ among countries.
- Some regions experienced HDI decline, likely due to the economic and social effects of the COVID-19 pandemic.
- Income only does not fully explain HDI; health and education are equally important too.

Short Analysis Questions:

- **Greatest improvement:** Nepal, Bangladesh, India, Pakistan, Sri Lanka
- **Declines:** Small changes in some countries due to COVID-19 and economic slowdown.
- **Highest vs lowest HDI region:** Europe highest, Africa lowest.

-

Global events impact: COVID-19 disturbed health and income, affecting HDI moves unevenly.

4. Problem 2

Approach:

- Calculated HDI statistics across multiple countries and regions.
- Identified HDI gaps and correlations with GNI per capita.
- Divided countries by development groups and visualized disparities.

Key Results:

- HDI strongly correlates with income, but some high-income countries have moderate HDI due to lower education or health metrics.
- Low-HDI countries are concentrated in Africa, very high HDI countries are mostly in Europe and North America.
- South Asia shows gradual improvements but lags behind Europe and East Asia in overall HDI.

Visualizations & Tables:

- Correlation heatmaps of HDI vs GNI.
- Bar charts for top 10 and bottom countries using HDI.
- Scatter plots for HDI vs income.

Interpretation & Discussion:

- Advanced exploration reveals structural disparities between regions.
- Policy interventions in health, education, and economic development can accelerate HDI improvements.

-

5. Problem 3 –

Implementation:

Filtered dataset for South Asia and Middle East countries.

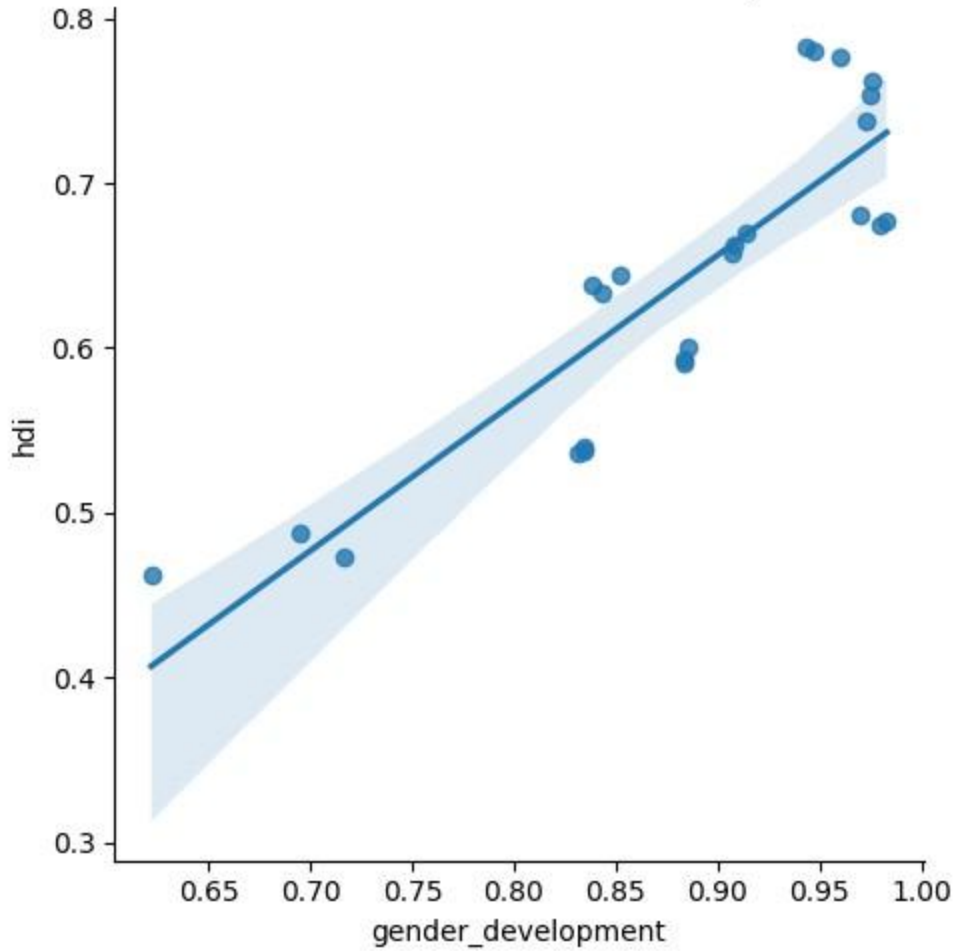
- average HDI for each region over 2020–2022.
- Compared HDI improvements, top performers, and regional disparities.

Main Results:

- Average HDI of Middle East is > South Asia.
- South Asia shows steady HDI improvement; Middle East has a more stable but slightly higher baseline HDI.
- Countries like UAE, Saudi Arabia, and Qatar lead in the Middle East; India, Nepal, and Sri Lanka are top in South Asia.

Visualizations & Tables:

South Asia: HDI vs Gender Development



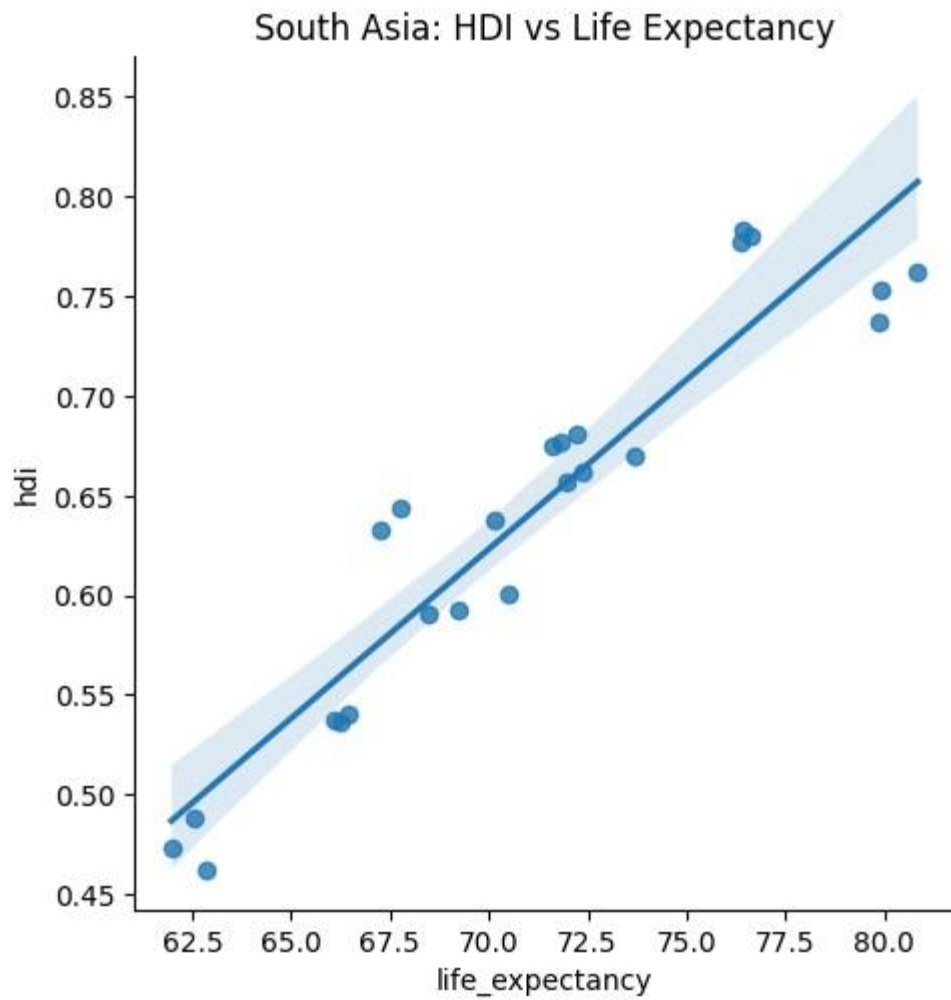
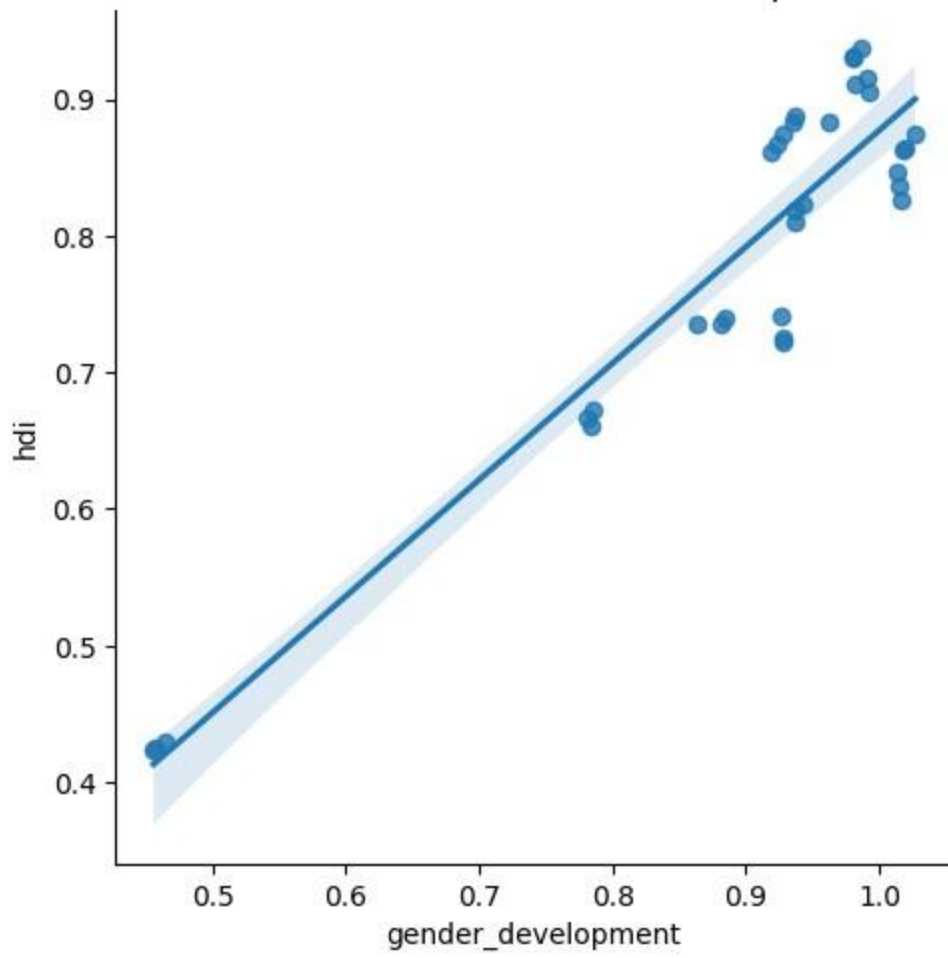


Figure 1: Line charts showing South Asia: Gender Development & Life Expectancy vs HDI trends.

Middle East: HDI vs Gender Development



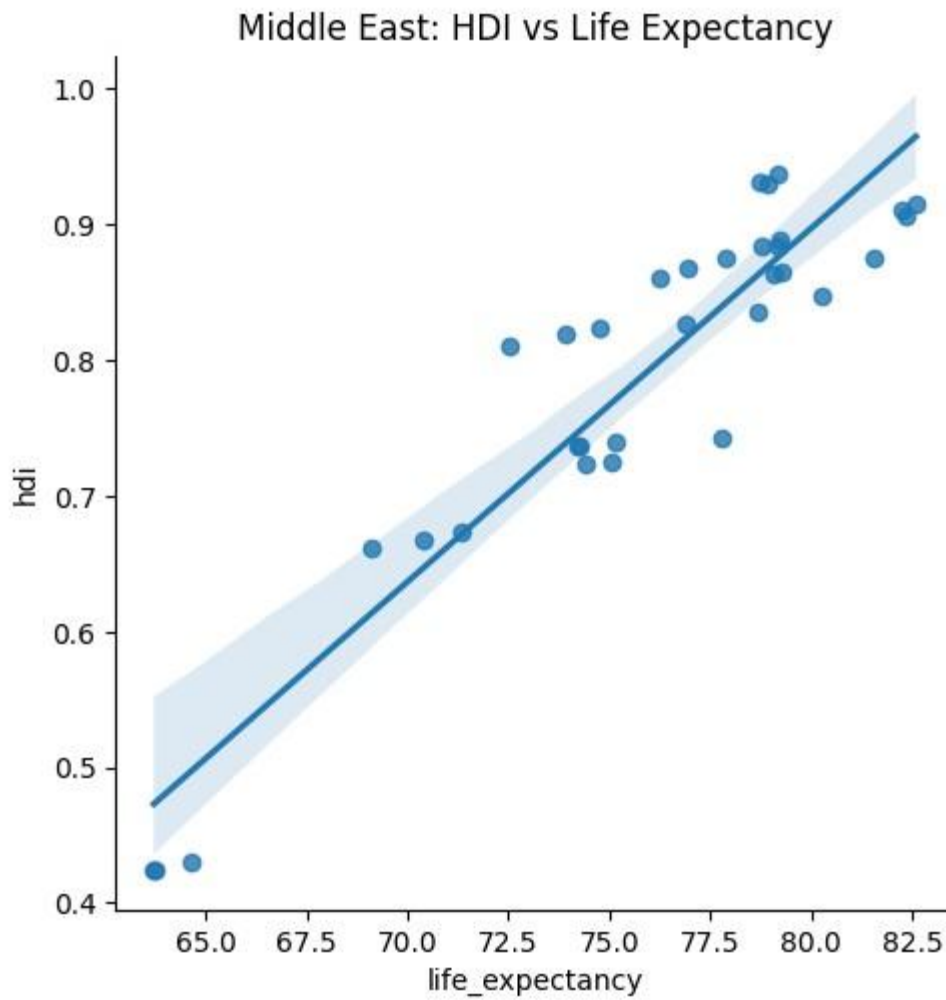


Figure no 2:Line charts showing Middle East: Life Expectancy & Gender Development vs HDI trends .

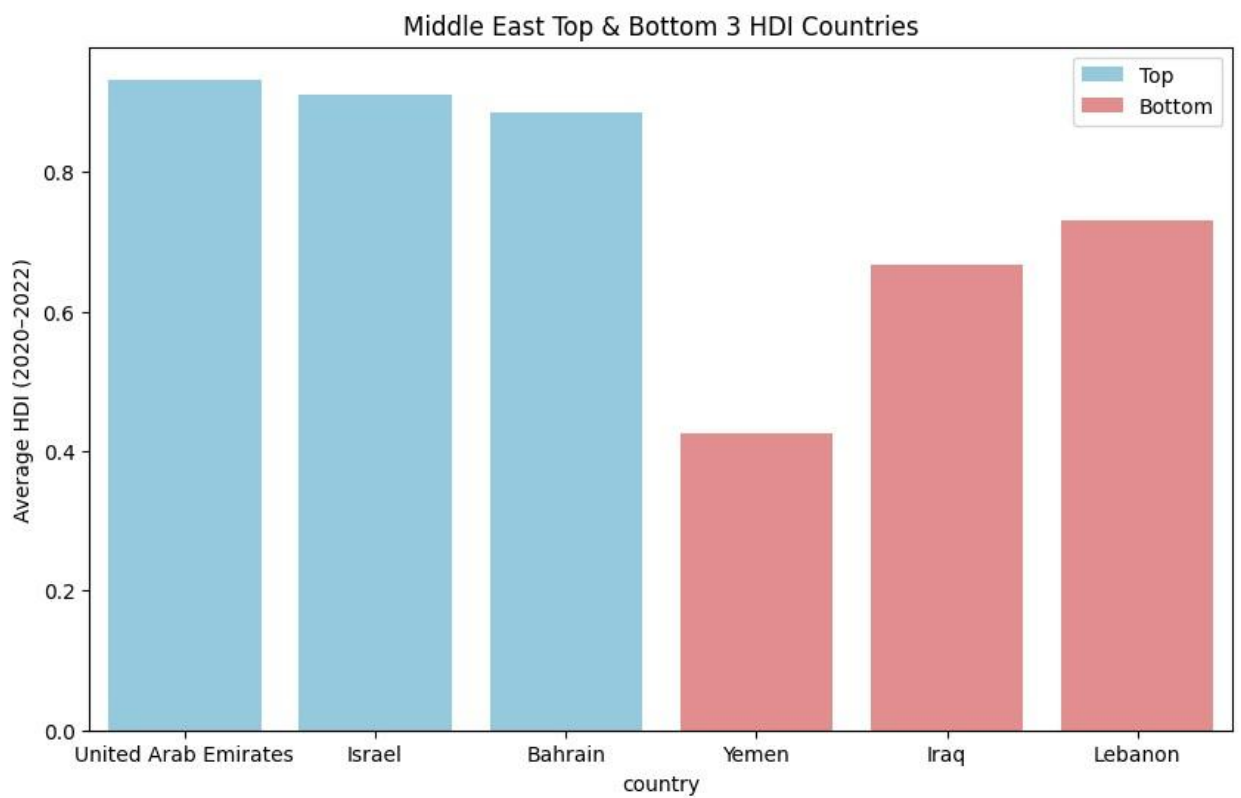
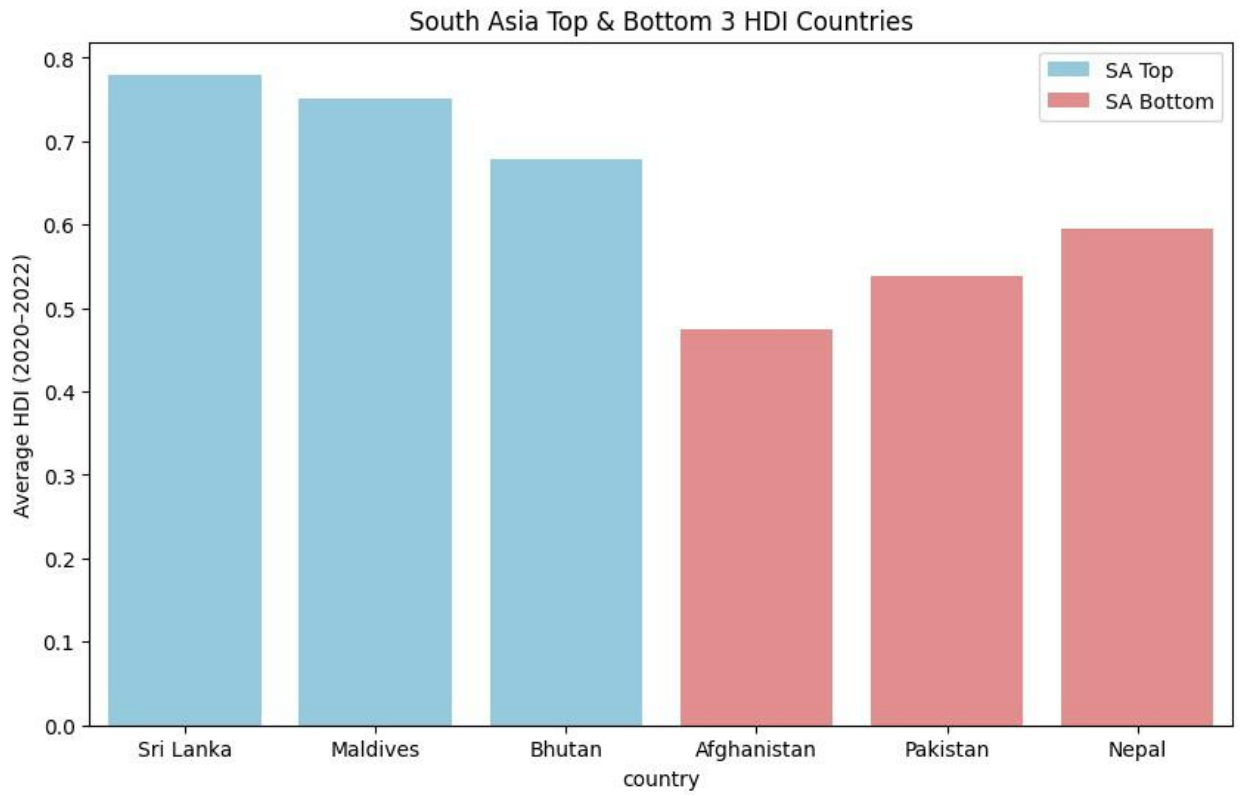
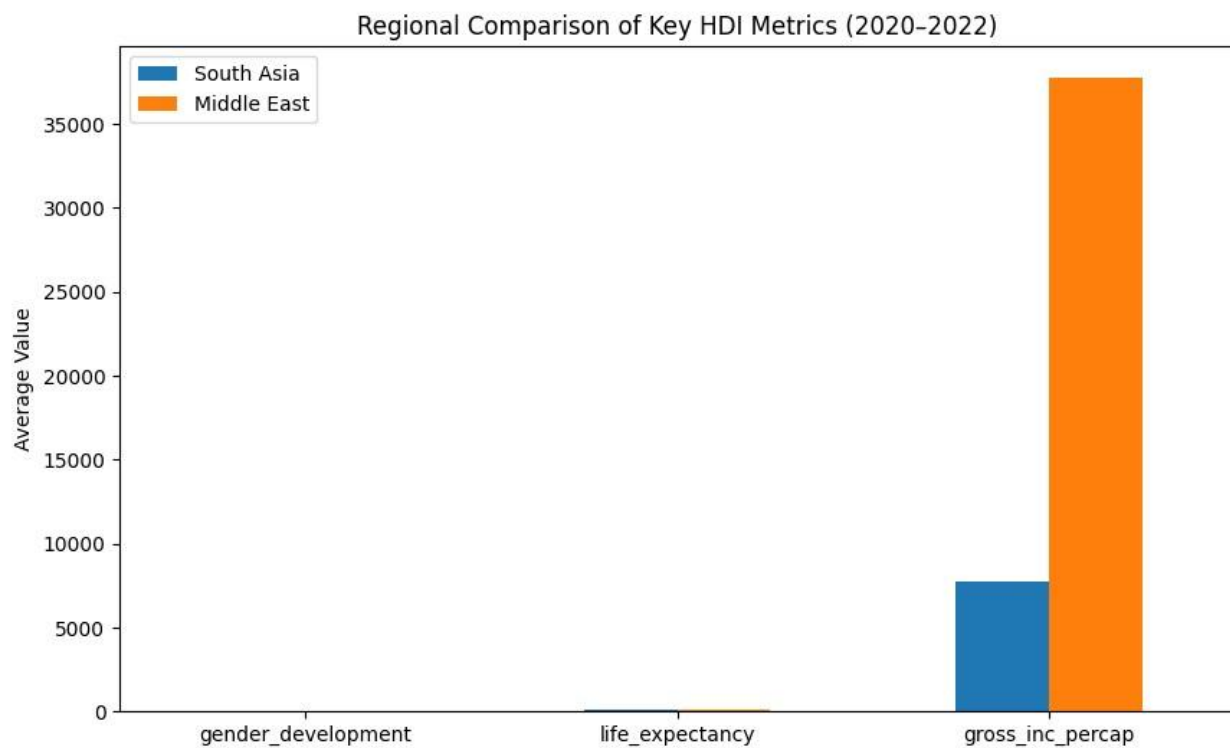


Figure 3: Middle East top, South Asia & Bottom three Countries rated by HDI.



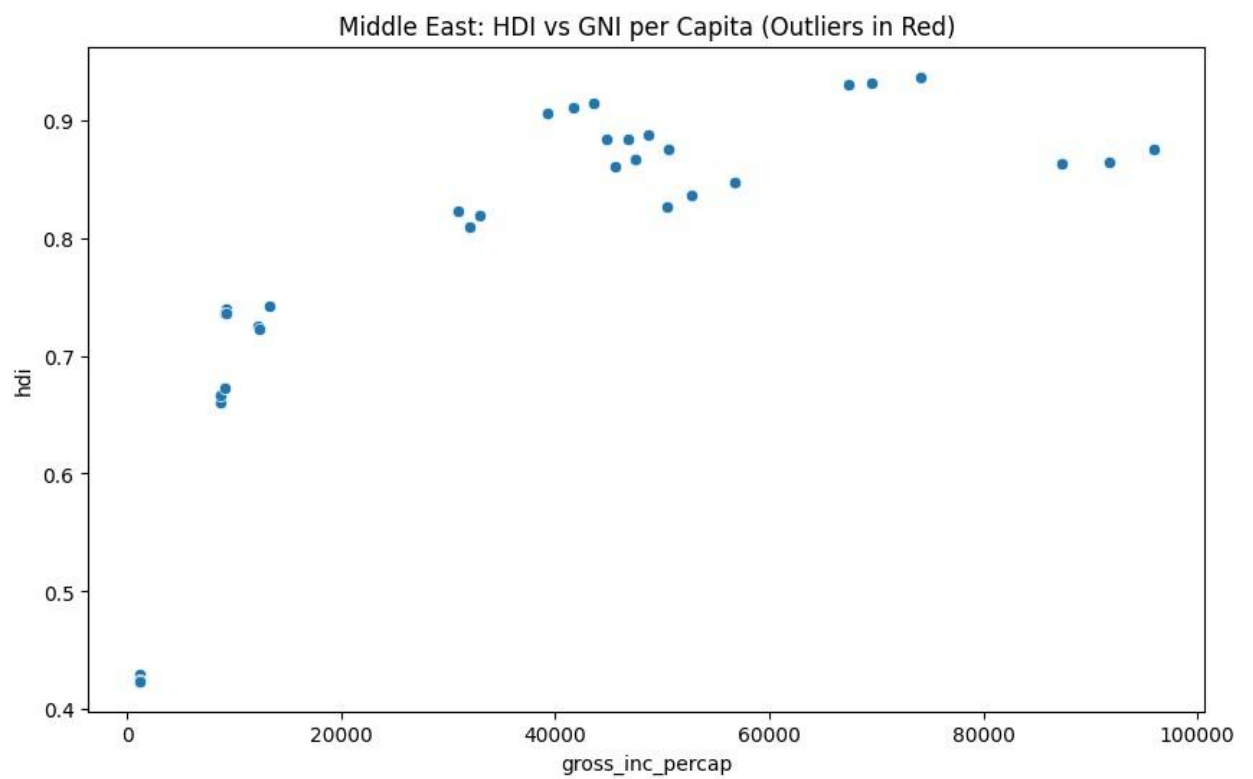
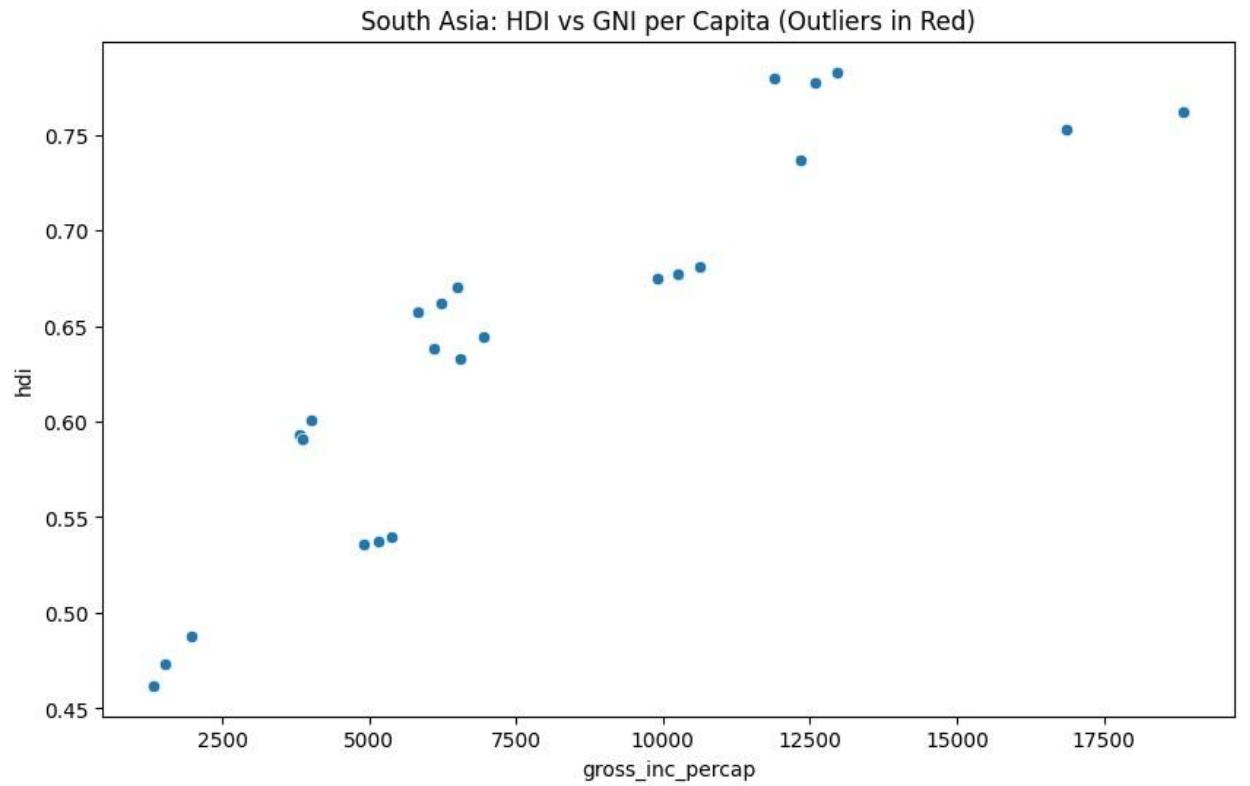


Figure 4: South Asia,Middle East Human Development index vs GNI per capita knowings & Discussion:

- Middle East has higher baseline HDI, likely due to oil-driven economies and social infrastructure.
- South Asia is improving faster but still lags in health and education indices.
- Comparative analysis highlights the need for targeted policies in low-performing countries.

6. Conclusion

- HDI trends show a global pattern of improvement, with regional disparities remaining significant.
- Europe and Middle East have the highest average HDI • Africa the lowest average HDI.
- Nepal leading in recent HDI gains, South Asia shows consistent improvement with.
- HDI is influenced by income, education, and health collectively; policies must address all three.
- COVID-19 caused temporary declines in some countries, reflecting the sensitivity of HDI to global events.

Limitations:

- Missing values and incomplete datasets for some countries.
- GNI per capita used may not fully reflect income distribution disparities.

Recommendations:

- Focus on health and literacy improvements in low-HDI countries.
- Track HDI changes yearly for early policy intervention.
- Address regional imbalance via targeted development programs.

7. References taken

1. United Nations Development Programme (UNDP) Human Development Reports.

2. HDI Dataset (CSV provided by course).
3. Imported needed Python libraries: pandas,matplotlib,etc.

7. Appendix

Appendix A: Plagiarism Report:

Similarity Report	
PAPER NAME	AUTHOR
WLVID_ShubhangadShrestha.docx	-
WORD COUNT	CHARACTER COUNT
1302 Words	7004 Characters
PAGE COUNT	FILE SIZE
23 Pages	884.3KB
SUBMISSION DATE	REPORT DATE
Jan 10, 2026 9:15 AM GMT+5:45	Jan 10, 2026 9:15 AM GMT+5:45
<div><div></div>20% Overall Similarity</div> <p>The combined total of all matches, including overlapping sources, for each database.</p> <div><div><ul style="list-style-type: none">• 2% Internet database• Crossref database• 20% Submitted Works database</div><div><ul style="list-style-type: none">• 2% Publications database• Crossref Posted Content database</div></div>	

GitHub Repository Link:

https://github.com/Shubhangad-Shrestha/HDI_Assignment-5CS037-

