UNIVERSITY PARTNER

UNIVERSITY OF WOLVERHAMPTON

HERALD COLLEGE KATHMANDU

# Statistical Interpretation and Exploratory Data Analysis of Human Development Index (2020-2022).


Student Id : 2509467
Student Name : Unibha Neupane
Group : L5CG11
Module Leader : Siman Giri
Submitted on : 01/06/2026

# Table of Contents

# 1. Introduction

Human Development Index (HDI) is a composite indicator and is created with the aim of determining the overall development of a nation in terms of health, education and life quality. The report will include an exploratory data analysis of HDI data in terms of recent trends (2020-2022), regional comparisons, will also offer a complex analysis of the South Asian region, and will as well compare the South Asian continent and the Middle East. Statistical methods and data illustrations were also found to establish patterns and relationship within the human development indicators, and outliers.

# 2. Dataset Description

It was also purged by eliminating records that contained unspecified values within the key columns such as country, year and HDI. To avoid the possibility of counting the records twice, duplicate records were eliminated. Numeral variables including the HDI and GNI/per capita used were converted into numerical data type since they were to be analyzed statistically.

# 3. Data Preprocessing

The dataset was also purged by getting rid of rows that had missing values in key columns like country, year and HDI. To prevent the situation of counting the records twice, duplicate records were discarded. Applicable numerical variables such as the HDI and GNI/per capita were transformed to numeric data type because they would be analyzed statistically.

# 4. Problem 1A: Single Year HDI Exploration

In this section, the author examines how Human Development Index (HDI) is distributed in one year as a way of learning the general development trends around the world. The analysis is based on determining variation and inequality in level of HDI.

## 4.1 Methods / Approach

In this section, the breakdown of Human Development Index (HDI) values in a one year period will be studied to observe the general developments of the world. The analysis is aimed at defining the variation and inequality in the HDI levels among nations.

## 4.2 Key Results

The HDI scores are broad which means that there is a great disparity in human development across the nations. Although the countries get very high HDI scores, majority of countries are still in medium or low development levels.

## 4.3 Visualization

This part was not clearly placed under a graphical visualization as the objective of Problem 1A was to explore the values of HDI in one year (2022) in relation to descriptive statistics, rather than trend and comparative plots. The analysis is supported by the generation of tables that were presented in the Jupiter notebook like a summary statistic, the country having the highest and lowest HDI and categorizing the countries to the following groups HDI (Low, Medium, High and Very High). These tables are quite useful to demonstrate the fluctuation and dispersion of the levels of HDI between the countries of the selected year.

# 5. Problem 1B: HDI Visualization and Trend

## 5.1 HDI Trend for Selected Countries

The trends of HDI between the years 2020 and 2022 of five chosen countries (Nepal, India, China, the United States, and Germany) were plotted using a line chart.

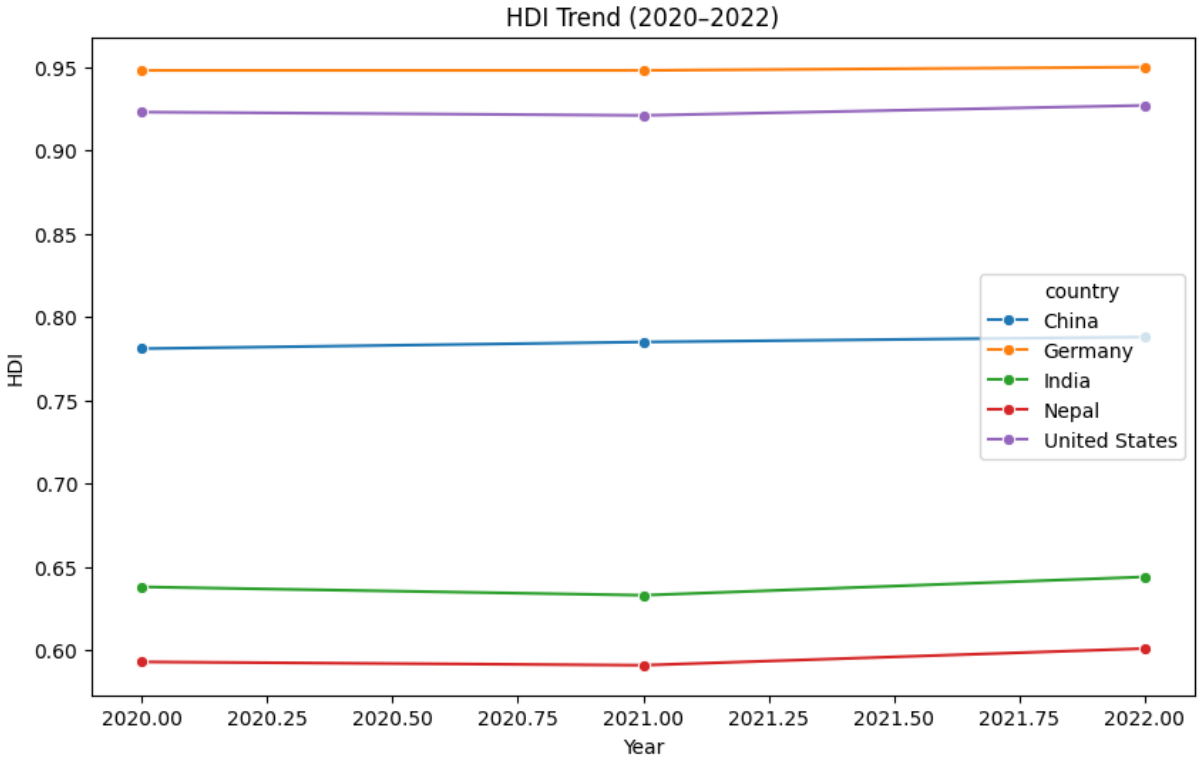


Figure 1: HDI trends for selected countries from 2020 to 2022 show gradual improvement in most countries, with developed nations maintaining higher HDI levels.

## 5.2 Average HDI by Region

The countries were divided into regions and average HDI per year was used as a value. The levels of HDI in the region were compared by a bar chart.

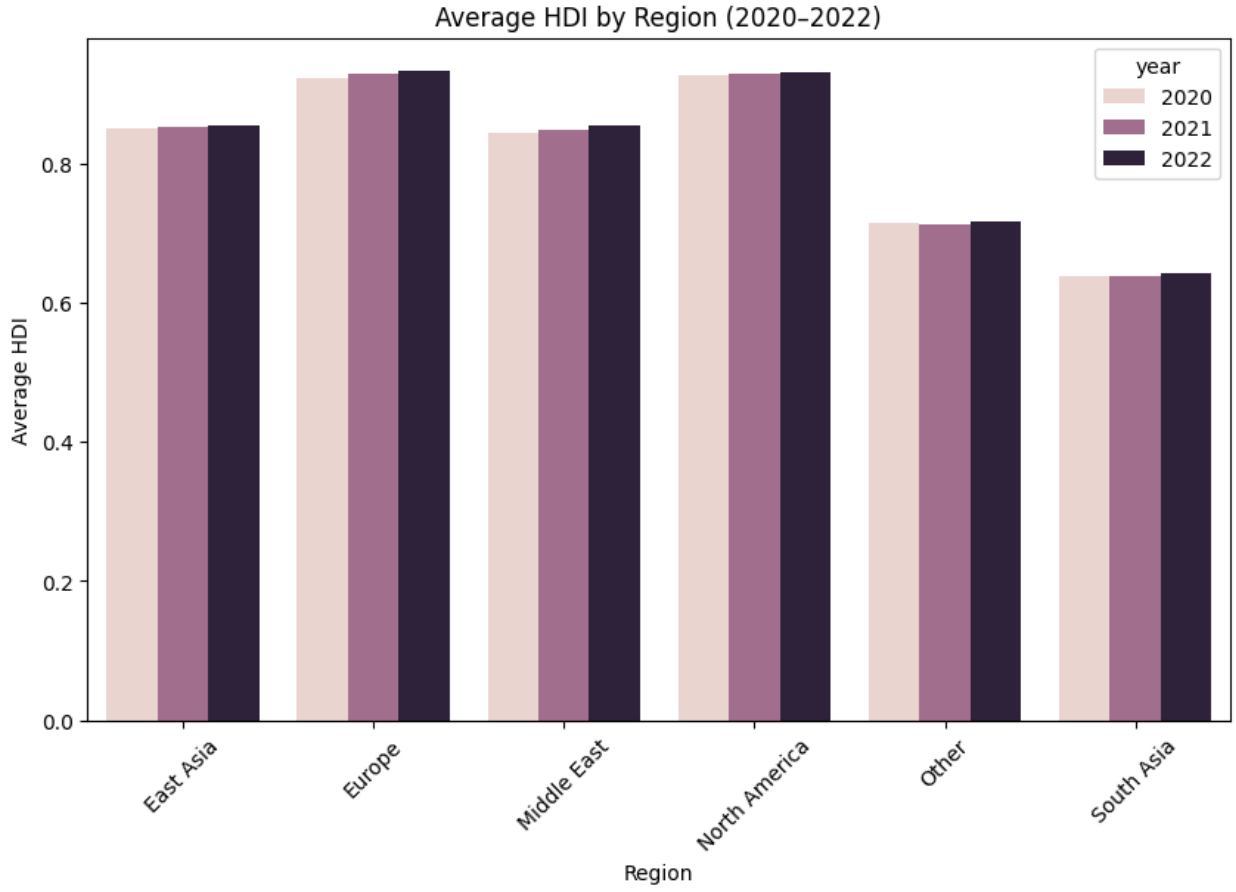


Figure 2 : Average HDI by region (2020–2022) highlights clear regional disparities, with developed regions outperforming South Asia.

## 5.3 HDI Distribution Across Years

The distribution of values of HDI in the three years was studied in the form of a box plot.

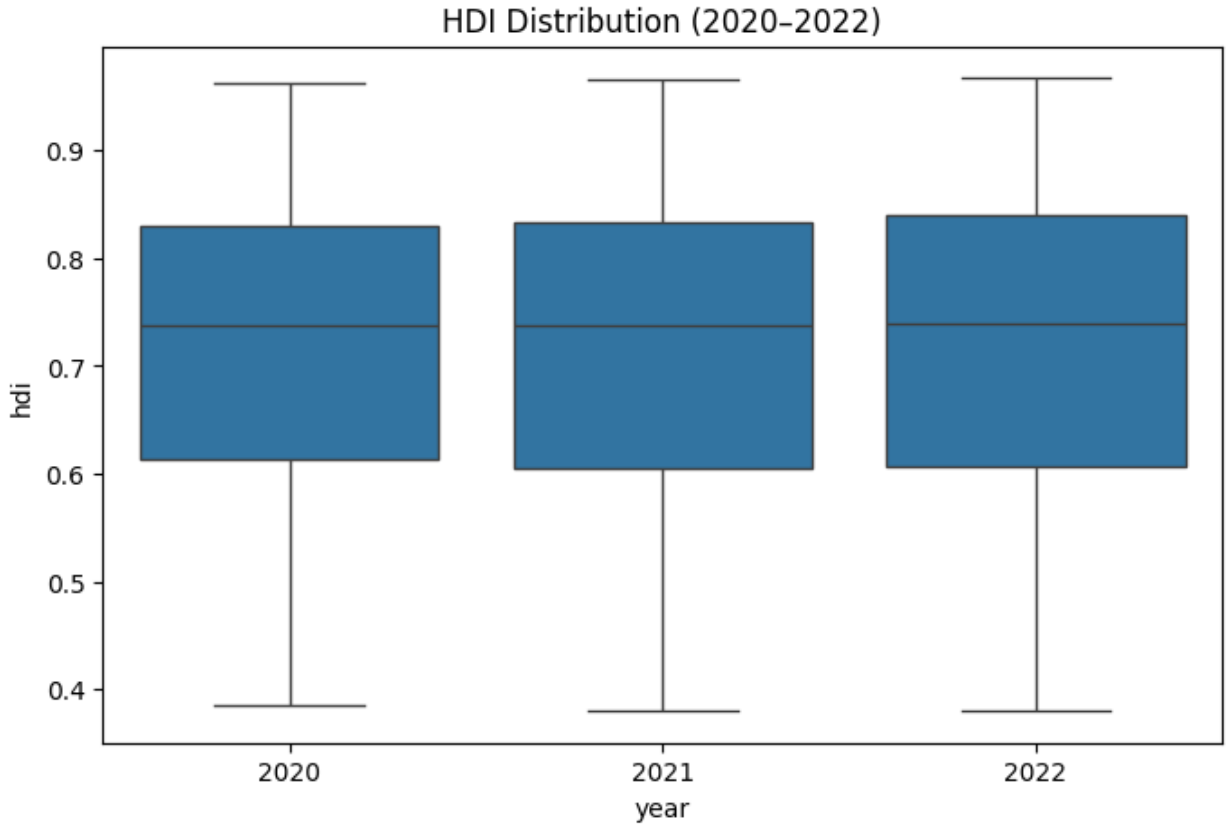


Figure 3: HDI distribution across years shows a slight upward shift in median HDI, indicating gradual global development improvement.

## 5.4 Relationship Between HDI and GNI per Capita

Such relationship between income and human development was on a scatter plot with regression line.

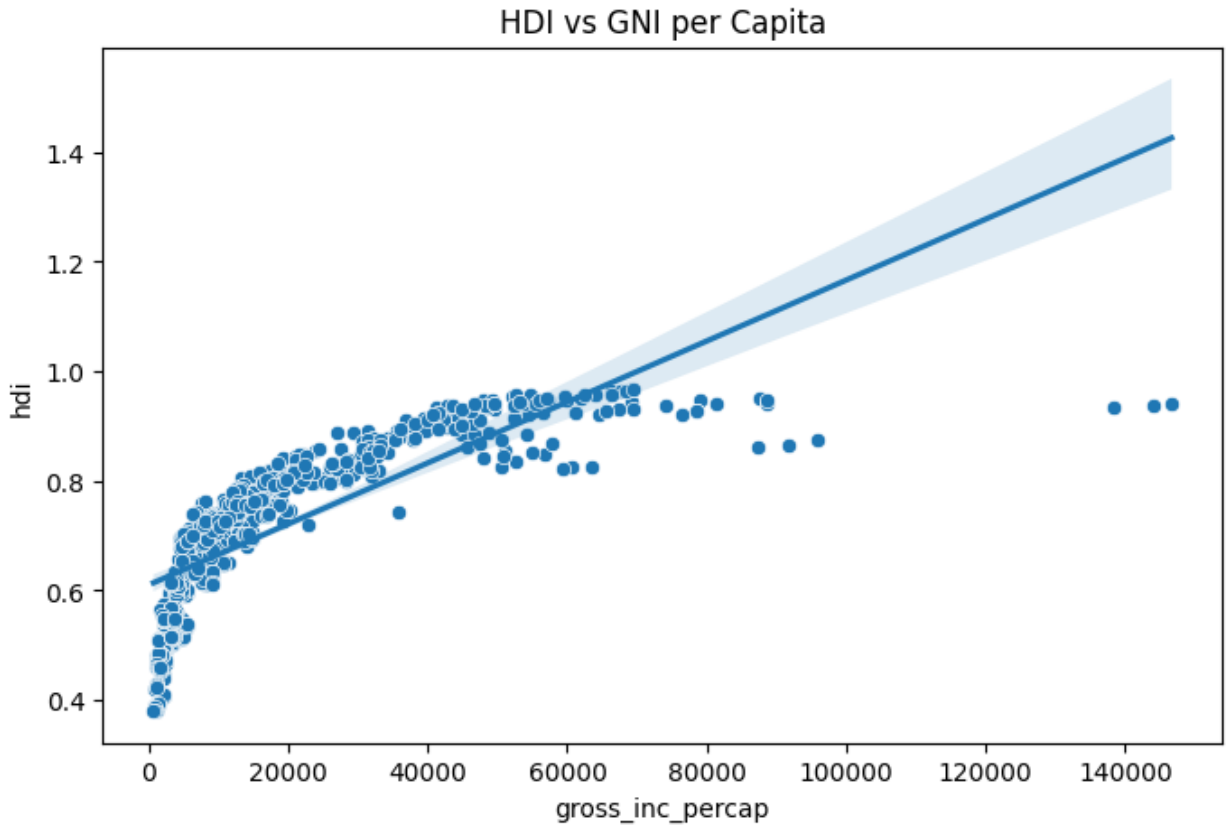


Figure 4 : The scatter plot shows a positive relationship between GNI per capita and HDI, suggesting that higher income generally supports better human development outcomes.

# 6. Problem 2: Advanced HDI Analysis (South Asia)

## 6.1 Composite Development Score

The South Asian countries were summed up to come with a composite development score basing on the chosen indicators. The visualization of the top five countries was ranked.

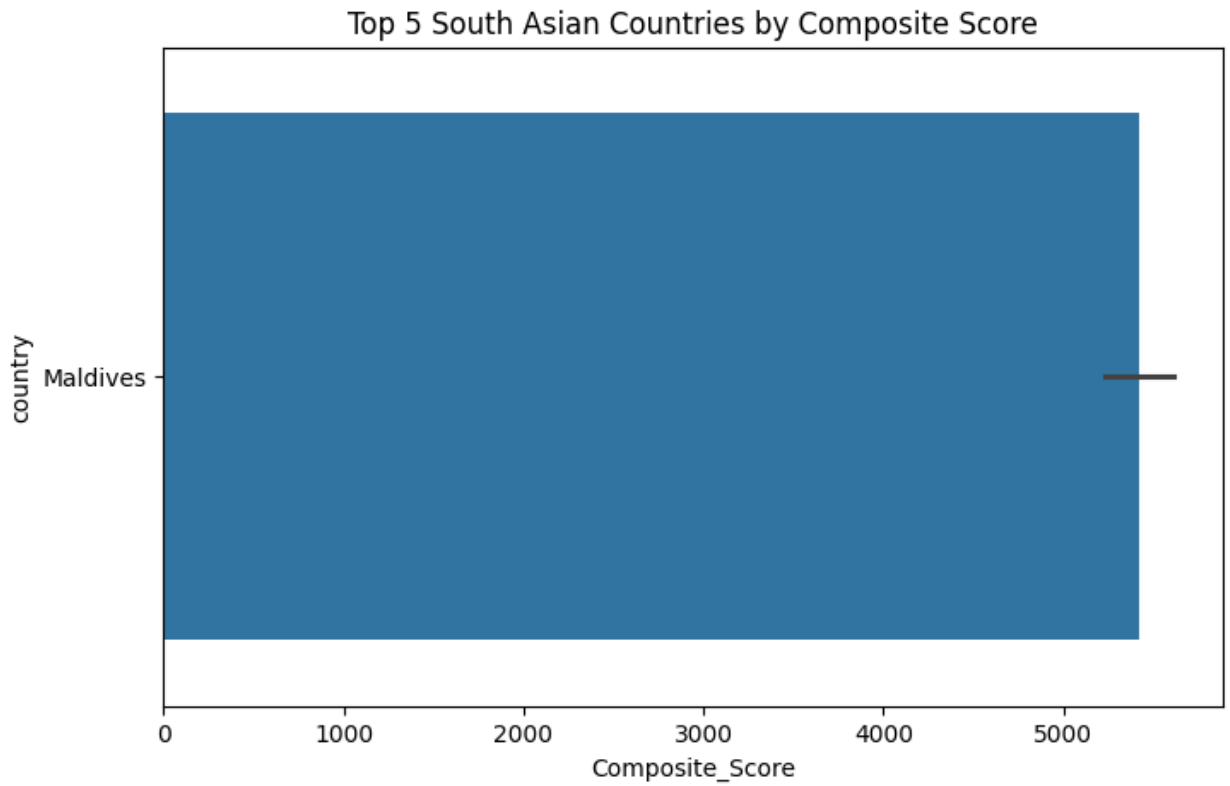


Figure 5: Composite development scores highlight differences in country performance beyond HDI alone.

## 6.2 GNI–HDI Gap Analysis

Disparity between GNI per capita and HDI was examined with the view of determining income-development discrepancies.

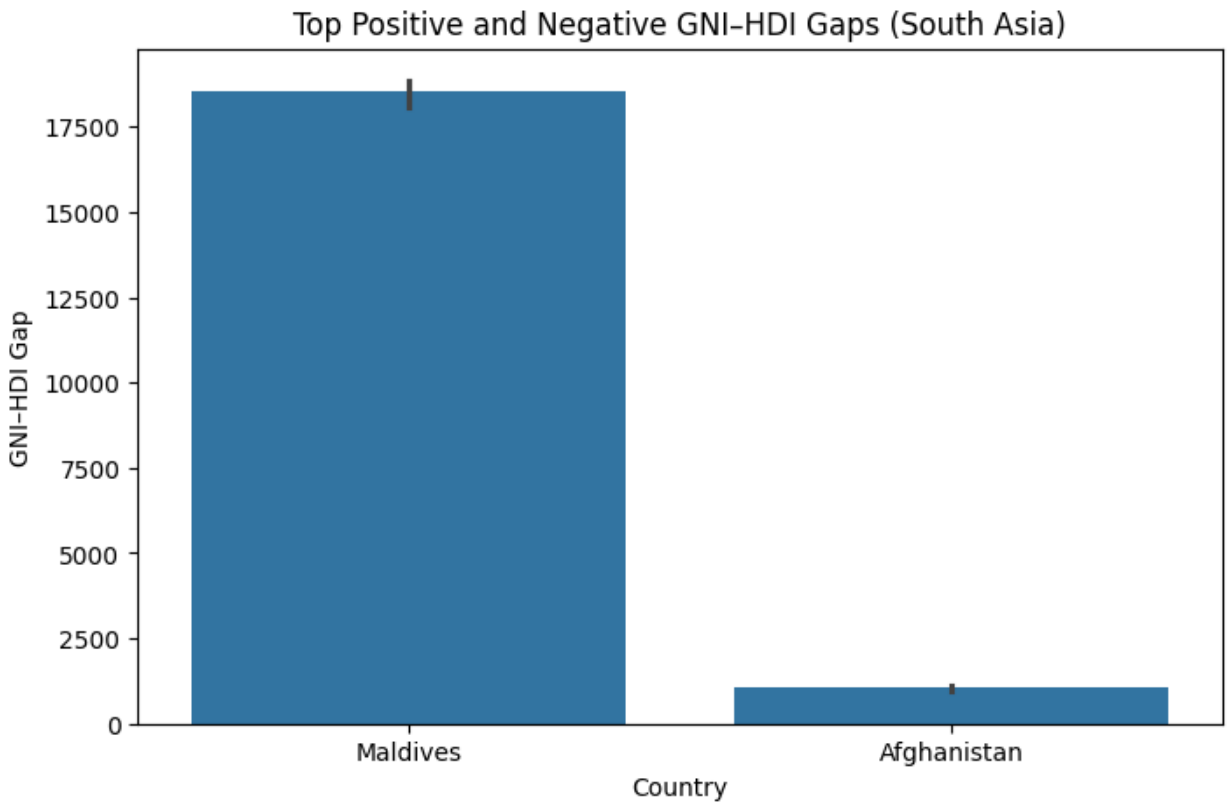


Figure 6: The GNI–HDI gap illustrates that higher income does not always translate into higher human development in South Asian countries.

## 6.3 Correlation Analysis

Scatter plots were used in the analysis of the relationship between HDI and other development indicators.

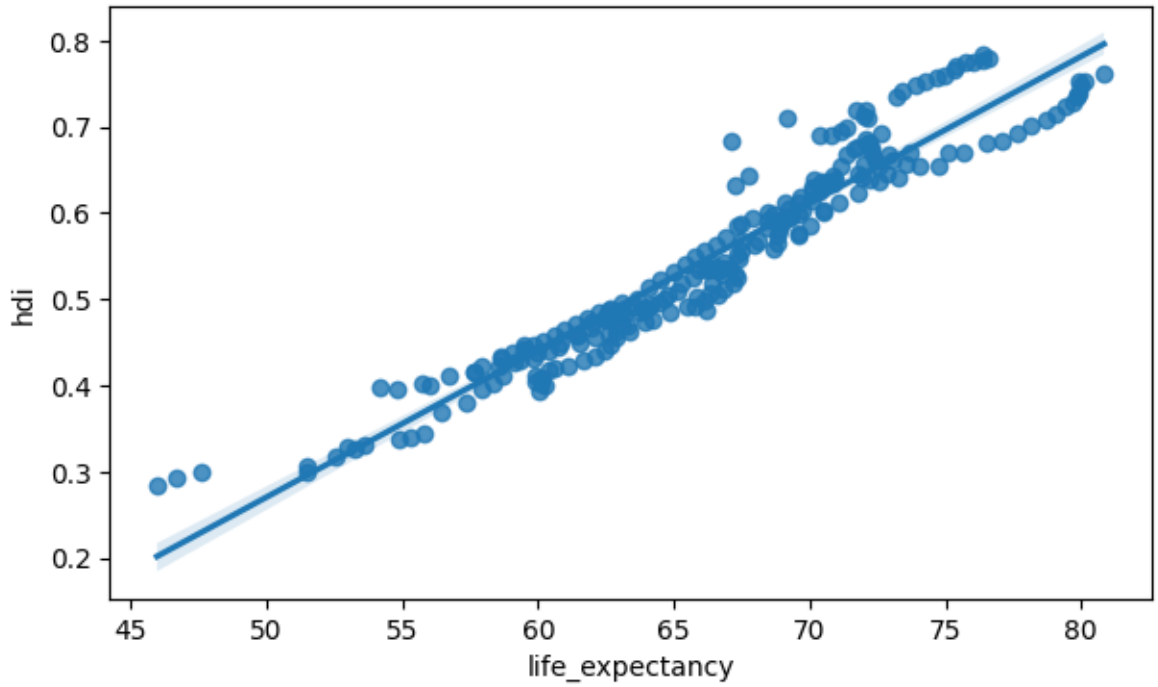


Figure 7: Life expectancy shows a strong positive correlation with HDI in South Asian countries.

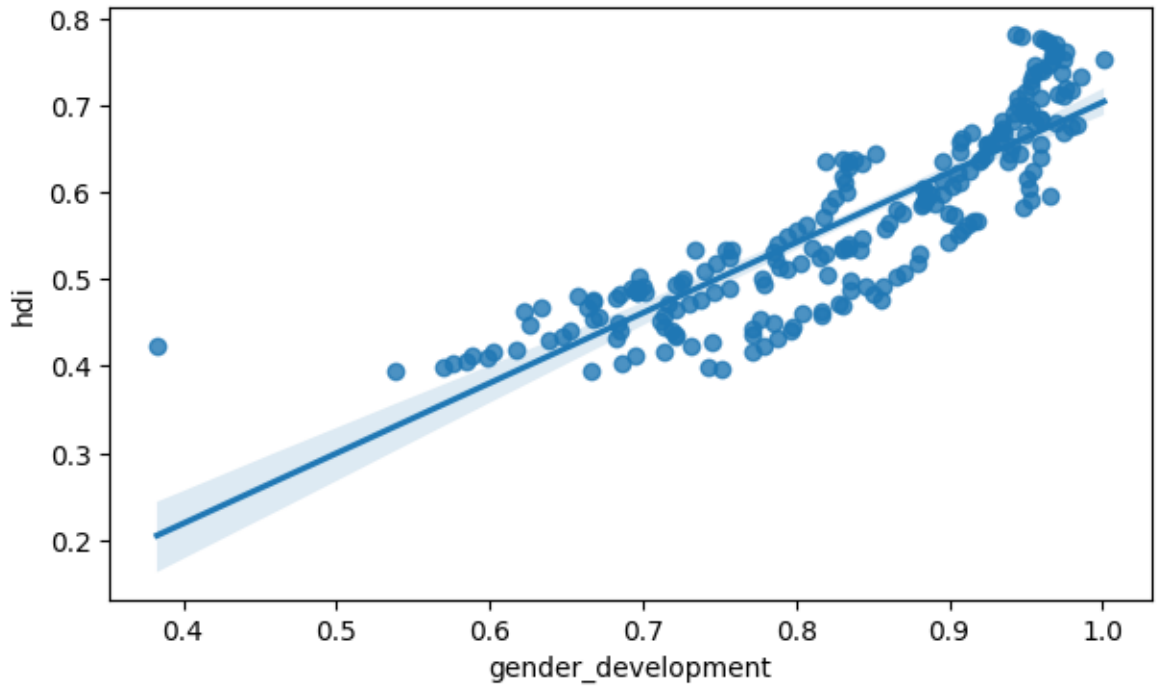


Figure 8: Gender development is positively associated with HDI, indicating the importance of gender equality in human development

# 7. Problem 3: South Asia vs Middle East Comparison

## 7.1 HDI Comparison Between Regions

The countries that were doing better and worse in HDI the most among the South Asians and the Middle Eastern countries were compared.

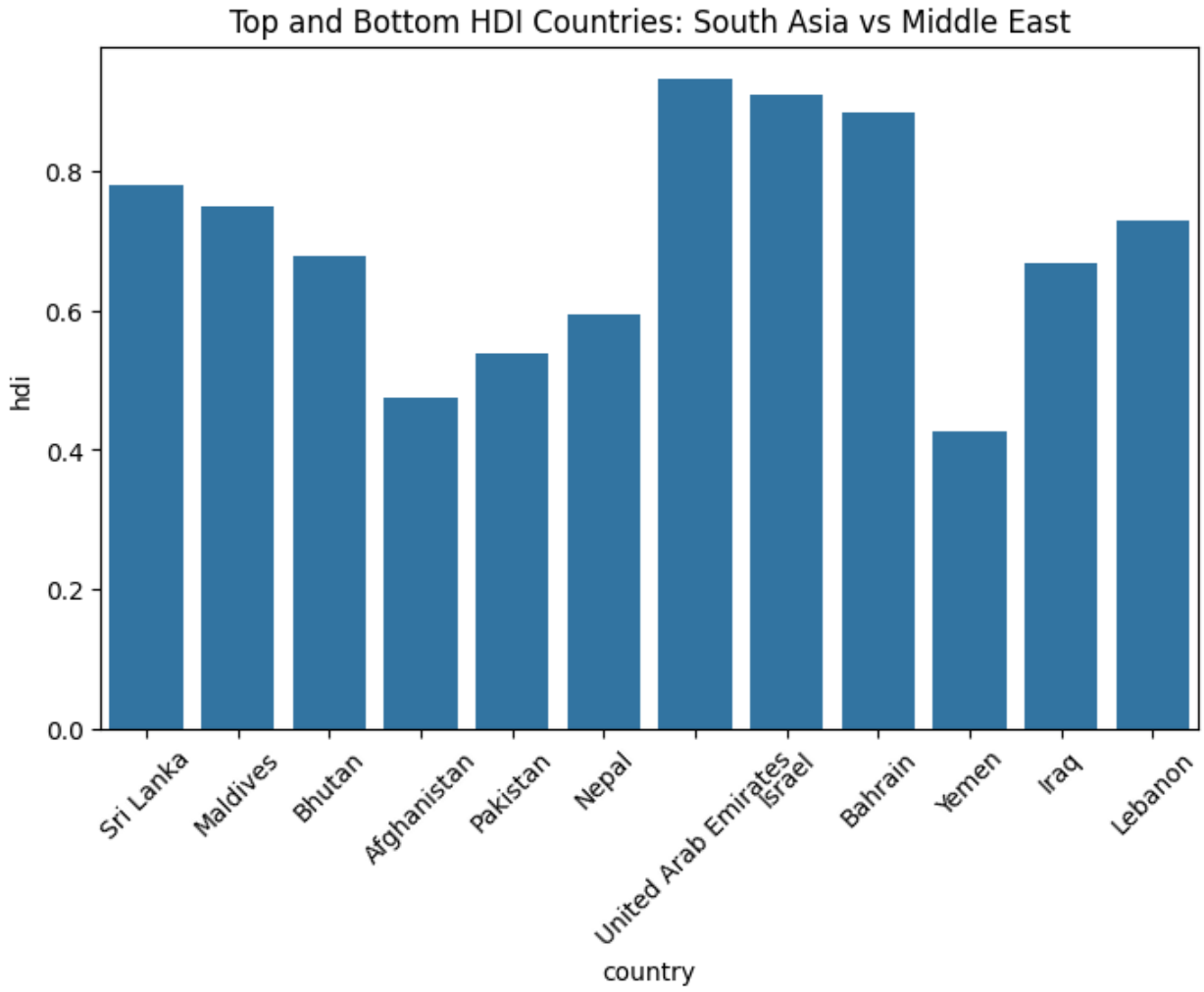


Figure 9: Comparison of HDI performance shows Middle Eastern countries generally achieving higher HDI values, while South Asian countries are more concentrated in the lower range.

## 7.2 Outlier Detection

The 1.5xIQR rule was used to identify outliers in the case of HDI and GNI per capita.

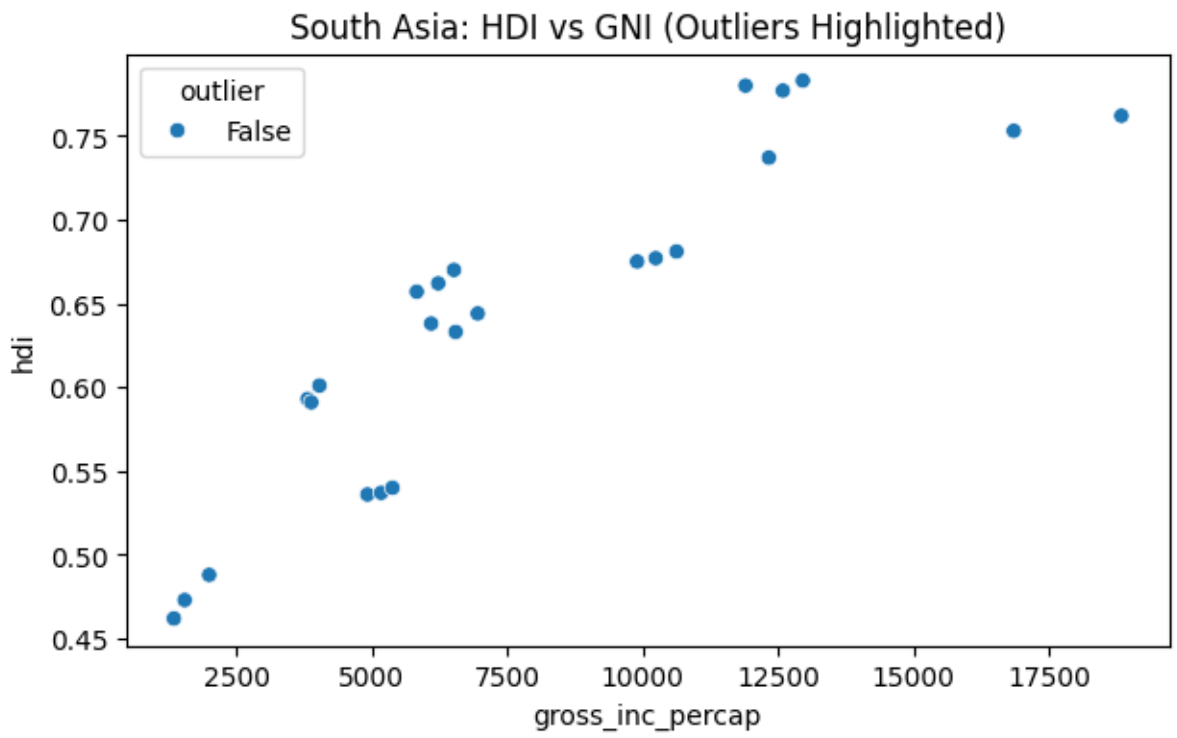


Figure 10: Income–HDI outliers in South Asia highlight countries with unusual development patterns.

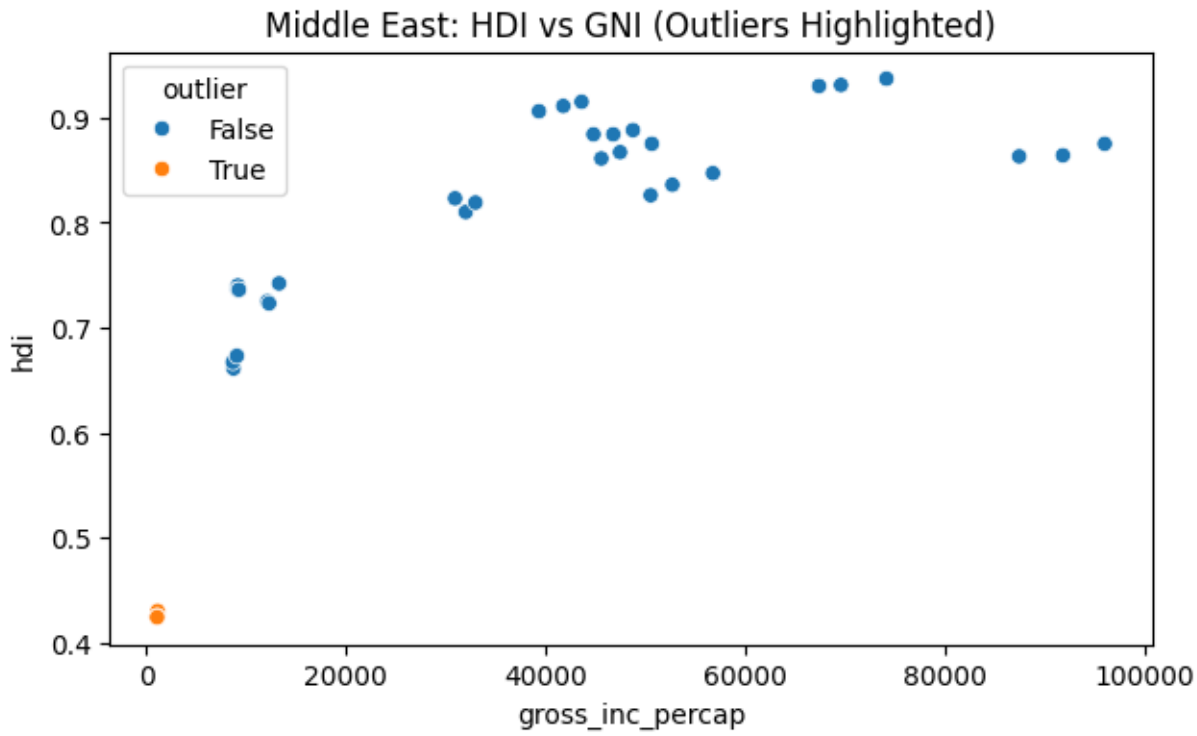


Figure 11: Middle Eastern outliers reveal countries with exceptionally high income but varying HDI levels.

## 7. Conclusion

As the analysis indicates, the HDI increased relatively between 2020 and 2022 but the differences between regions remain large. HPI is closely connected with income and the amount of life expectancy, and gender development is also of significant importance. Comparative and outlier analyses prove that economic wealth is not enough to guarantee even-loyal human progress.

# 8. Appendix

**Dataset Source:**

Human Development Report (UNDP)

**Tools Used:**

Python (Pandas, Matplotlib, Seaborn)

https://github.com/unibha/Assignment_01