

PAPER NAME

**kapil adhikari.docx**

AUTHOR

**-**

WORD COUNT

**1067 Words**

CHARACTER COUNT

**5831 Characters**

PAGE COUNT

**11 Pages**

FILE SIZE

**625.8KB**

SUBMISSION DATE

**Jan 8, 2026 11:13 PM GMT+5:45**

REPORT DATE

**Jan 8, 2026 11:14 PM GMT+5:45**

### ● 17% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

- 2% Internet database
- 0% Publications database
- Crossref database
- Crossref Posted Content database
- 17% Submitted Works database



HERALD  
COLLEGE  
KATHMANDU



UNIVERSITY OF  
WOLVERHAMPTON

Herald College, Kathmandu

4 Concepts and Technologies of AI

5CS037

Assignment - I - Statistical Interpretation and Exploratory Data Analysis.

Student Name: - Kapil Adhikari

Student University Id: - 240078

Date: - 9<sup>th</sup> January 2026

## Contents

11	Introduction .....	3
	Problem-wise Analysis .....	4
2	Problem 1A – Single Year HDI Exploration .....	4
	Problem 1B – HDI Trend Analysis (2020–2022) .....	5
	Problem 2 – Advanced HDI Exploration .....	8
	Problem 3 – Comparative Regional Analysis: South Asia vs Middle East .....	9
	Conclusion .....	9
	• References: .....	11
	Appendix .....	11

# Introduction

The Human Development Index (HDI) is an index which is being created by the United Nations Development Program, which helps to ascertain the level of development of the nation, as well as the measure of the outcome of the growth of the Gross Domestic Product, instead of being concerned with the outcome of the Gross Domestic Product. In addition to that, other aspects include health aspects (Life Expectancy), the level of educational standards, together with the level of living standards.

The current report encompasses a statistical reason and an exploratory study concerning the specified HDI for the assessment of different events that have been occurring as well as those occurring in the current situation. The justifications for carrying out the statistical analysis on the specified data include assessment of the specified value for the latest year, assessment for the specified value for periods between 2020 and 2022, further comprehensive assessment concerning countries belonging to South Asia, and comparative assessment concerning different events for countries in the geographical regions of South Asia as well as Middle East. The assessment for this current report will be conducted in connection with this specified set.

# Problem-wise Analysis

## 6 Problem 1A :

### Single Year HDI Exploration

The data was then filtered to ensure it was relevant to the year 2022, which is as up-to-date as possible. The exploratory analysis of data was completed, which included checking the type of the data, managing missing observations, as well as summary statistics to gain insights regarding HDI.

Regions of Focus

!!!

The below summary statistics are ordered in an effort to extract information regarding mean, median, and standard deviation of the development stage in the world. With regard to increased comprehension regarding development stages of different countries based on HDI ranking, from most to least developed countries, it is essential to assess countries with highest and lowest HDI. Finally, countries are grouped based on categories that are accepted as HDI categories.

***	hdi	life_expectancy	gross_inc_percap
count	204.000000	204.000000	204.000000
mean	0.722887	71.791847	20722.658911
std	0.153029	7.758711	22105.824388
min	0.380000	52.997000	690.660758
25%	0.606500	65.946750	4800.833454
50%	0.739500	72.208000	12663.864770
75%	0.839500	77.753000	32042.765448
max	0.967000	84.820000	146673.241500

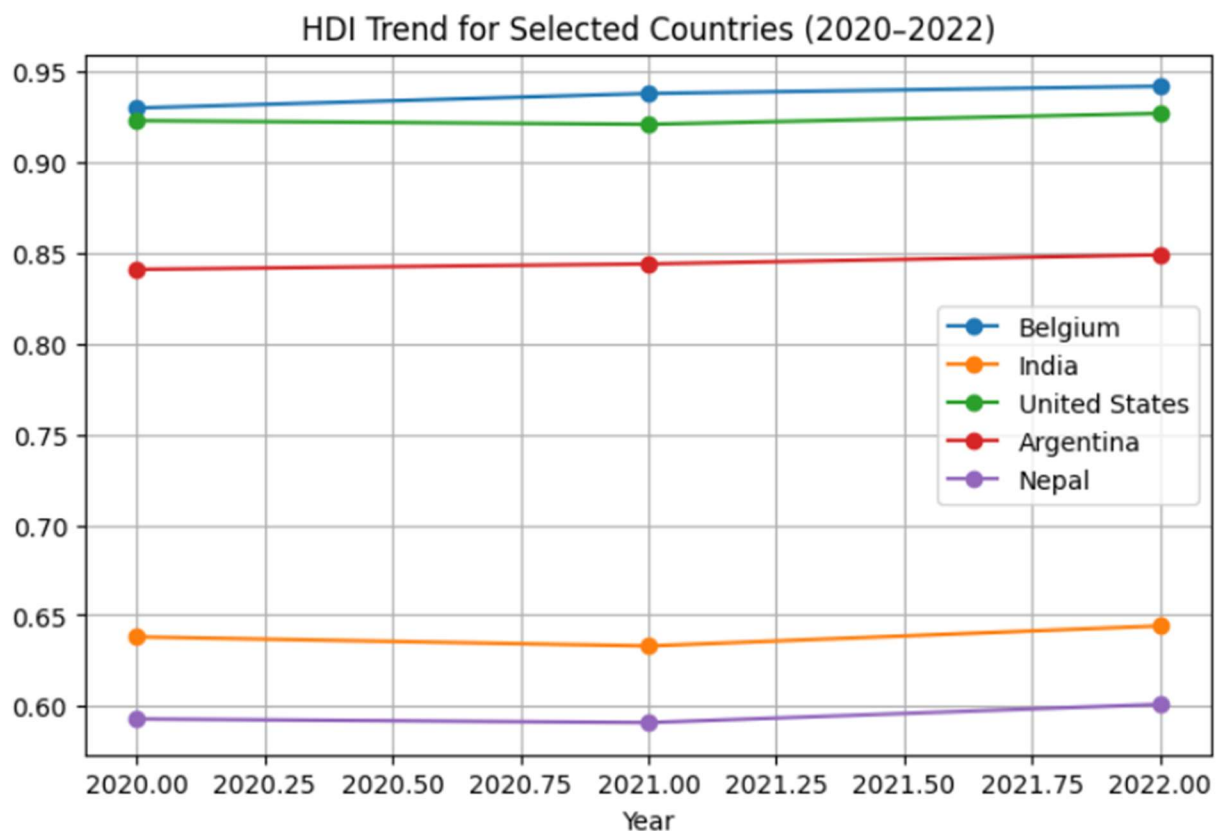
## 7 Problem 1B:

### HDI Trend Analysis (2020–2022)

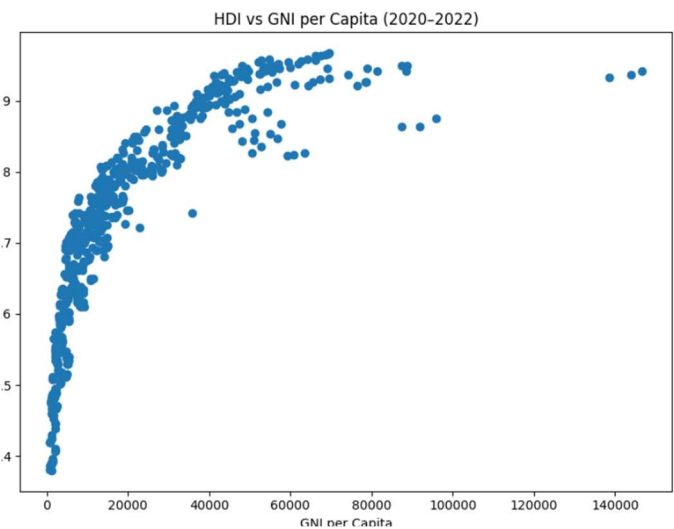
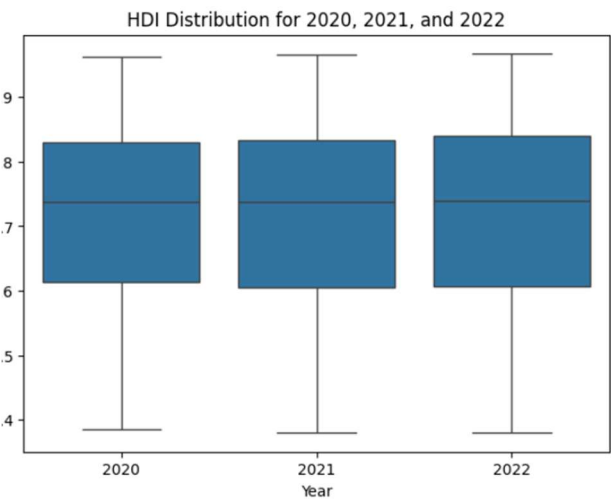
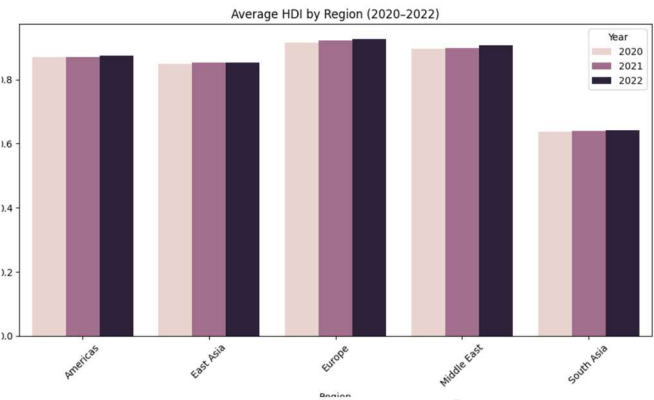
The data is sorted based on a query to retrieve the data for the year 2020, 2021, and 2022. The data is cleaned for consistency in the country name variable and in numeral variables for carrying out trend analysis.

#### Key Results

In a span of three years, some countries have demonstrated progress in HDI. Some other countries have had stagnation or a slight deterioration. The regions have portrayed some uneven developments.



Above is the line graph that depicts the trend of HDI for the selected countries from the years 2020 to 2022. The progress is visible in selected countries, although some countries are still struggling in the wake of the e-disruption in the process of recovery. The countries include India.



#### Task 4. <sup>1</sup> Short Analysis Questions

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

The countries that are showing noticeable improvement in HDI are India, Bangladesh, and Nepal. The improvement in HDI from 2020 to 2022 is possibly because of improved recovery in healthcare, education, and economic systems that were affected by COVID-19.

<sup>3</sup> 2. Did any countries experience a decline in HDI? Provide possible reasons.

Yes, some countries experienced a slight stagnation in HDI, specifically in the period 2020-2021. For instance, nations which experienced a slow pace in economic growth, as well as those whose medical systems are overstressed, experienced minimal progress.

<sup>1</sup> 3. Which region has the highest and lowest average HDI across these three years?

More developed regions, such as Europe, generally have the highest average HDI across 2020–2022, while less developed regions, such as parts of South Asia or Sub-Saharan Africa, tend to have lower average HDI values.

<sup>1</sup> 4. How may global events (e.g., the COVID-19 pandemic) have affected HDI trends during this period?

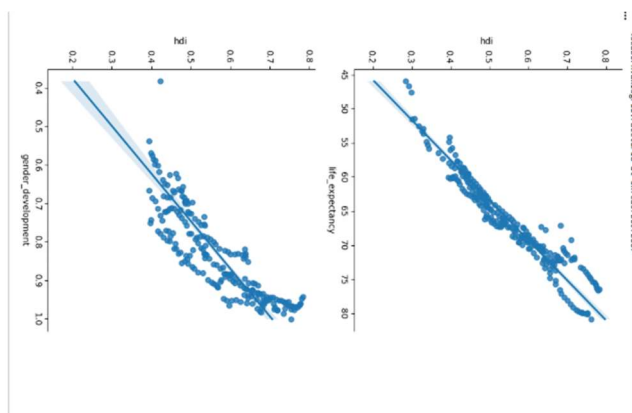
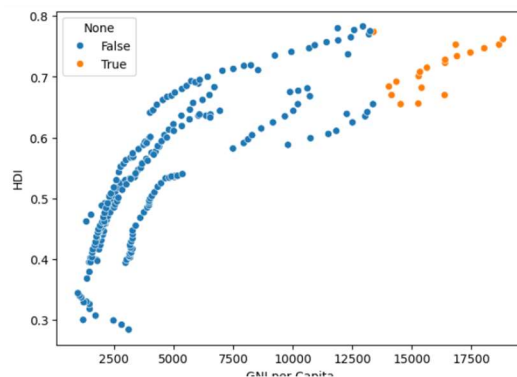
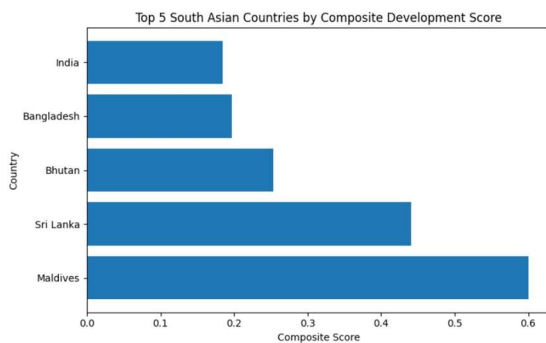
The HDI was adversely impacted by the COVID-19 pandemic in terms of lower life expectancy, disrupted educational processes, and hindered economic development. This resulted in a stagnation and/or decline in HDI levels for a certain period, followed by a gradual revival.



## Problem 2 – Advanced HDI Exploration

The report focused on south Asian countries using a region-specific subset of the dataset. It also produced a Composite Development Score using different development indices for correlation and outlier identification.

Variations existed between rankings of composite scores and scores of HDI. This implied that countries might rank differently if more attention is given to a few variables. Outlier tests identified countries whose development outcomes fall above and below expected regional levels.

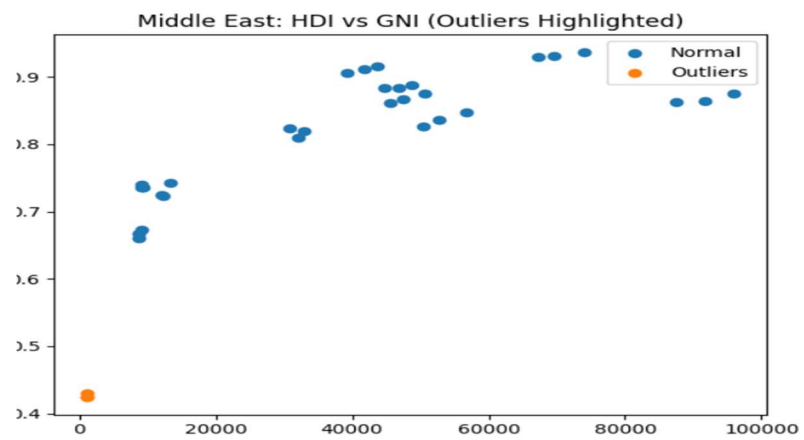
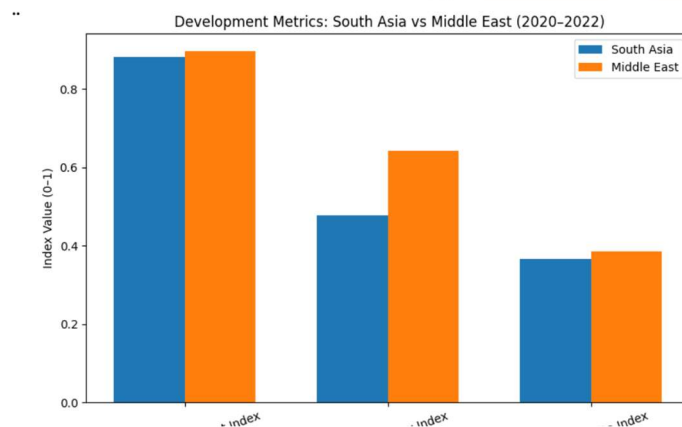
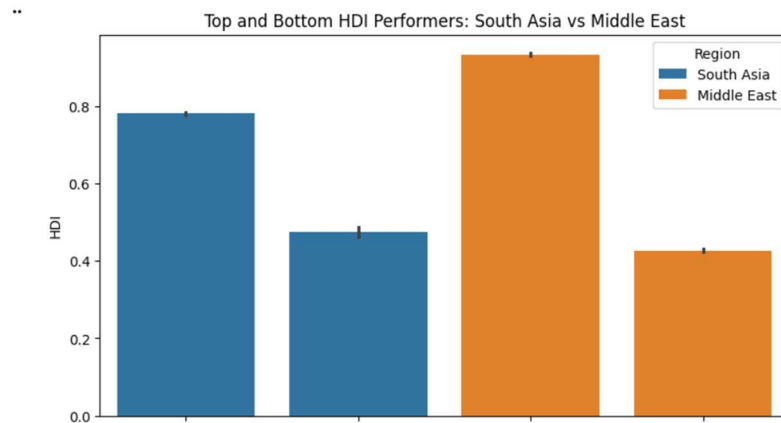


5

### Problem 3 – Comparative Regional Analysis: South Asia vs Middle East

The datasets for the years 2020 to 2022 were employed to carry out a comparison between the South Asia region and the Middle Eastern region using descriptive statistical analysis.

The value of the HDI has been revealed to be higher for the Middle East, while the range of the HDI is broader for the South Asia region. Inequality between the two regions is evident for the value and development aspect.



# Conclusion

In this situation, there exist regions with inequalities with regard to the aspect of human development, as mentioned above. However, there is an improvement with regard to the levels of HDI generally. There is inequality with regard to the various regions. South Asia region has lower HDI than the Middle East region. Inequality with regard to the outcomes of development is higher in the South Asia region.

An important variable that influences HDI includes health and income. This is represented through high correlation coefficients from the linear relation analysis carried out, including variables such as life expectancy at birth and economic variables. The implications from the results include emphasizing improvements with regard to health care, education, and economic development. Some variables considered for inclusion were life expectancy at birth and a variety of economic variables. The limitations include the fact that it only utilizes data obtained from secondary sources. This research has a time limit.

## References:

8 United Nations Development Program. (2022). *Human Development Index Dataset*.

## Appendix:

Link to Code:

## ● 17% Overall Similarity

Top sources found in the following databases:

- 2% Internet database
- 0% Publications database
- Crossref database
- Crossref Posted Content database
- 17% Submitted Works database

### TOP SOURCES

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	<b>University of Wolverhampton on 2026-01-07</b> Submitted works	4%
2	<b>University of Wolverhampton on 2026-01-08</b> Submitted works	4%
3	<b>University of Wolverhampton on 2026-01-07</b> Submitted works	2%
4	<b>University of Wolverhampton on 2026-01-07</b> Submitted works	1%
5	<b>University of Wolverhampton on 2025-12-07</b> Submitted works	1%
6	<b>University of Wolverhampton on 2026-01-06</b> Submitted works	<1%
7	<b>University of Wolverhampton on 2026-01-08</b> Submitted works	<1%
8	<b>link.springer.com</b> Internet	<1%

9	University of Wolverhampton on 2026-01-08	<1%
<hr/>		
10	University of Wolverhampton on 2026-01-08	<1%
<hr/>		
11	University of Wolverhampton on 2026-01-06	<1%