Analysis of the Performance of the Environmental Indicators in India

Graduate Diploma in Economics
2023-24
Birkbeck, University of London

Smaranya Dey

Student Id: 13919411

Abstract

The environmental indicators are measures that provide practical ways to track the health of the environment. Since the climate crisis is an imminent threat of the present time, it is crucial to analyse the data related to the environment and climate to learn about the risks posed due to the same. This paper attempts to analyse the trend of the different environmental indicators, e.g., CO2 emissions, forest areas present in the country, electricity production from renewable sources etc for India over the past years. It plans to conduct a comparative analysis between India and the rest of the world of the mentioned environmental indicators. Additionally, the project presents a correlation analysis to check the linear association between the said indicators present and a regression analysis to gauge the impact of one or two such indicators. The project will use the data on environmental indicators obtained from the World Bank Database (link).

Word Count

~3000

(excluding the Bibliography and the Appendix)

Programming Language Used

R

I. Introduction

The environmental indicators are essential tools for tracking environmental progress, supporting policy evaluation and providing public information. For example, the World Bank's 'The Little Green Data Book' was published in 2017 with key environmental data for over 200 countries organised under the sections of agriculture, forestry, biodiversity, oceans, energy, emission, pollution water and sanitation [1]. These indicators are means to explain the complex patterns of diverse aspects of the health of the environment. In 2015, at the United Nations Framework Convention on Climate Change Conference, the parties adopted the Paris Agreement to limit the global temperature increase to well below 2^0 C above pre-industrial levels. The climate crisis poses a major impact in India as well as all over the globe. Since the climate crisis is an imminent threat of the present time, it is crucial to analyse the data related to the environment and climate to learn about the risks posed due to the same, which can influence the direction of national-level policymaking.

This project is structured into five main sections to provide a comprehensive exploration of the topic. The introduction and the background section prepare the stage by outlining the significance and context of the research problem. The data section provides the source and explanation of the data considered for the project. The analysis section delineates the methodology employed and analytical techniques utilized with the presentation of results with visual and numerical representations. Finally, the project concludes with a summary of findings and suggestions for future research directions. The bibliography and appendix section contain the materials, papers and concepts referred to due to the successful completion of the project.

II. Background

In 2015, all the member countries of the United Nations created 17 goals, also known as Sustainable Development Goals (SDG), covering a range of social, economic and environmental issues in a global partnership to provide a shared blueprint for peace and prosperity for people and the planet to aim for alleviating poverty, protecting the planet and ensuring prosperity[2]. In 2023, the SDG progress report suggests that mobilizing climate finance and bringing climate change into national policies are two beneficial steps for making progress on all SDG targets[3]. India has been ranked 112^{th} among the 166 countries for its performance in 2023, towards achieving the SDGs, as per the Centre for Science and Environment (CSE). It is intriguing to check India's performance on certain environmental indicators over time and the comparison of the same among the South Asian countries. This project attempts to analyse the trend of the different environmental indicators, e.g., CO2 emissions, forest areas present in the country, electricity production from renewable sources etc for India over the past years. It plans to conduct a comparative analysis between India and the rest of the world in terms of the mentioned environmental indicators. Additionally, the project presents a correlation analysis to check the linear association between the said indicators present and a regression analysis to gauge the impact of one or two such indicators. The project will use the data on environmental indicators obtained from the World Bank data repository [4]. The environmental indicators include CO2 emissions, the percentage of forest area in the country, the percentage of the population with access to electricity, the percentage of resources for freshwater withdrawal, the percentage of electricity production from renewable sources etc.

III. Data

In this project, I utilised the data on environmental indicators from the World Bank Data Repository. The collection encompasses 266 countries from 1990 to 2020, out of which I specifically focused on India, and made comparisons with other South Asian countries and global averages. The available indicators were,

- **CO2 emissions**, measured in metric tonnes per capita, can be tied to *SDG 9: Industry, Innovation and Infrastructure*. The data is available in the data repository of the World Bank[5]. In this analysis, the focus is on India, for which the data ranges from the year 1990 to 2020.
- **Forest areas**, measured as the percentage of land areas, can be used as one of the metrics to measure *SDG 15: Life on Land*. This dataset is available in the data repository of the World Bank[6]. For India, the data ranges from 1990 to 2021.
- **The percentage of the population with access to electricity** is available from 1993 to 2021 for India. This indicator can be tied to *SDG 7: Affordable and Clean Energy* and is available in the data repository of the World Bank[7].
- **Total annual freshwater withdrawals**, measured in percentage of internal resources, range from 1975 to the year 2020 for India. The indicator is tied to *SDG 6: Clean Water and Sanitation* and is available in the data repository of the World Bank[8].
- **The percentage of electricity production from renewable sources** is available from 1971 to 2015 in the data repository of the World Bank[9]. This indicator is a potential metric for *SDG 7: Affordable and Clean Energy*.
- **The percentage of the population using safely managed sanitation services** ranges from 2000 to 2022 in the data repository of the World Bank[10].

Since the data availability may vary for different countries, I maintained the same period of the indicators data while comparing among the South Asian countries, which will be explained in the Analysis section.

IV. Analysis

This section contains summary statistics for each of these indicators of India are provided and compared with the other South Asian Countries which are Afghanistan (AFG), Bangladesh (BGD), Bhutan (BTN), India (IND), Sri Lanka (LKA), Maldives (MDV), Nepal (NPL) and Pakistan (PAK). Additionally, the line charts are showcased to understand the trend of these indicators.

A. Exploratory analysis of the variables

- CO2 emissions (metric tonnes per capita)

The data on CO2 emissions is available for India as well as for other South Asian countries from 1990 to 2020.

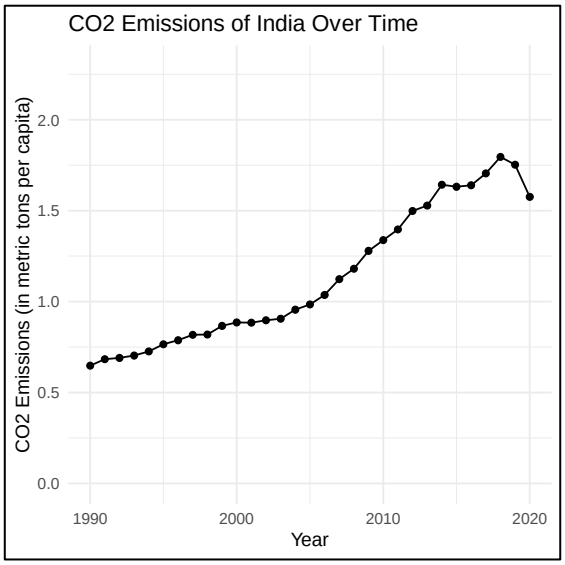

Figure 1: Line Chart for CO2 emissions (metric tonnes per capita) of India from 1990 to 2020

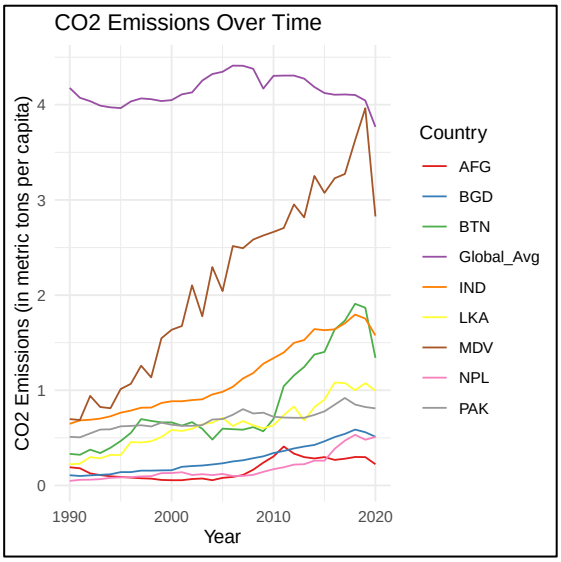

Figure 2: Multiple Line Chart for CO2 emissions (metric tonnes per capita) of South Asian countries and global average from 1990 to 2020

Figure 1 shows a line chart which depicts the CO2 emissions of India in metric tonnes per capita during 1990 and 2020. It shows an increasing trend in emissions per capita from 1990 to 2018, which can be attributed primarily to a rapidly growing economy, an upsurge in energy consumption from fossil fuels and policies prioritising industrialisation. As described in Figure 1, a decreasing trend can be seen in 2019 onwards. The country's focus on solar and wind power as renewable energy sources may be the reason for such a downward trend.

Figure 2 shows a comparison of CO2 emissions in metric tonnes per capita among the South Asian countries and the global average during 1990-2020. Naturally, the global average of the CO2 emissions is higher compared to the emissions of the countries of South Asia. The emissions of Maldives (MDV) are significantly higher compared to the other seven South Asian countries, though it started around a similar level to India in 1990. Maldives (MDV), Bhutan (BTN), Sri Lanka (LKA), Bangladesh (BGD) and Afghanistan (AFG) have a similar downward trend to India (IND) in their CO2 emissions from 2019.

Country	*Mean*	*Median*	*Min*	*Max*	*Q1*	*Q3*
AFG	0.17	0.13	0.05	0.41	0.08	0.28
BGD	0.28	0.23	0.10	0.59	0.16	0.40
BTN	0.85	0.66	0.32	1.91	0.56	1.20
IND	1.13	0.98	0.65	1.80	0.82	1.51
LKA	0.64	0.63	0.22	1.08	0.46	0.78
MDV	2.13	2.29	0.69	3.96	1.20	2.82
NPL	0.18	0.12	0.05	0.53	0.10	0.22
PAK	0.69	0.70	0.51	0.92	0.63	0.76
Global_Avg	4.15	4.11	3.77	4.41	4.05	4.29

Table 1: Summary Statistics of CO2 emissions over 1990-2020 for SA countries and global average

Similar insights can be drawn additionally from Table 1. Similar to Figure 2, the values for global average emissions are quite above the country's values. Among the eight counties in South Asia, Maldives (MDV) has the highest average emissions followed by India (IND) and Bhutan (BTN). Furthermore, the distribution of emissions within each country varies, as indicated by differences between median and mean values, suggestive of potential skewness or outliers in certain cases.

- Forest area (% of land)

The data for forest areas as a percentage of total land is available from 1990 to 2021 for India as well as other South Asian countries.

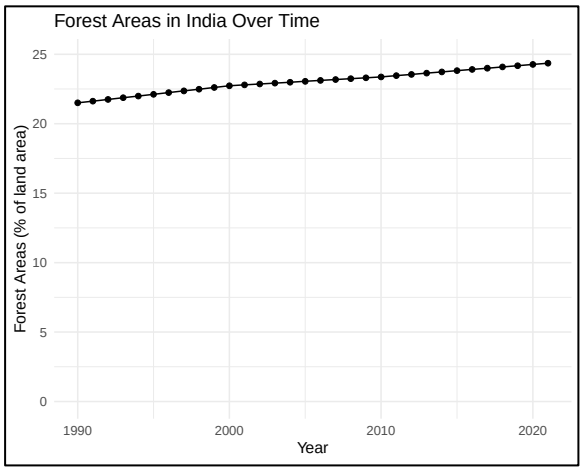

Figure 3: Line Chart for forest areas (% of lands) of India from 1990 to 2021

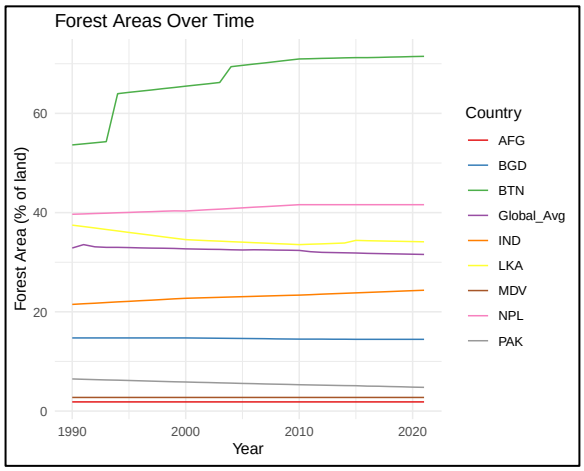

Figure 4: Multiple Line Chart for forest areas (% of land) of South Asian countries and global average from 1990 to 2021

Figure 3 depicts a line chart of forest areas as a percentage of land in India during 1990 and 2021. The percentage of forest areas varies between 20% to 25% during the course of the years of interest and exhibits a slight overall increase over the years. This suggests that the proportion of land covered by forests has been relatively stable over the period with a gradual growth of such areas. From Figure 4, similar insights can be drawn for most of the countries in South Asia as well as for the global average forest lands, where each one of them shows a steady proportion of the same over the years. Interestingly, among the South Asian countries, Bhutan (BTN) enjoys the highest proportion of forest areas, even more than the global average, with a consistent upward trend over time. Notably, there were periods of accelerated growth observed, from 1993 to 1994 (54% to 63%) and again from 2003 to 2004 (66% to 69%).

Country	*Mean*	*Median*	*Min*	*Max*	*Q1*	*Q3*
AFG	1.85	1.85	1.85	1.85	1.85	1.85
BGD	14.62	14.62	14.47	14.75	14.48	14.75
BTN	66.94	69.80	53.65	71.50	64.92	71.14
IND	23.04	23.09	21.50	24.36	22.46	23.67
LKA	34.71	34.28	33.55	37.48	33.93	35.21
MDV	2.73	2.73	2.73	2.73	2.73	2.73
NPL	40.88	41.02	39.66	41.59	40.25	41.59

PAK	5.58	5.55	4.78	6.47	5.17	5.99
Global_Avg	32.42	32.50	31.56	33.55	31.93	32.83

Table 2: Summary Statistics of forest areas over 1990-2021 for SA countries and global average

Table 2 offers a comprehensive overview of the summary statistics of the forest areas for the South Asian countries and average global forest areas over 1990-2021 with mean, median, minimum, maximum, first and third quartile values. Bhutan (BTN) exhibits the highest mean with 66.94%, which is even higher than the mean of the global average forest areas, 32.42%, explaining the trend found in Figure 4, followed by Nepal (NPL) with a mean of 40.88% and Sri Lanka (LKA) with a mean of 34.71%.

- Access to electricity (% of population)

The data on the percentage of the population that has access to electricity in India is available from 1993 to 2021, with a yearly average of 72.74%. By 2021, 99.57% of the population of the country has access to electricity.

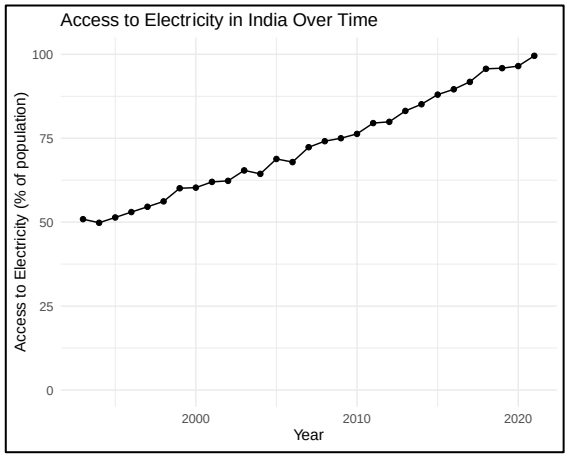

Figure 5: Line Chart for % of population with electricity access in India from 1993 to 2021

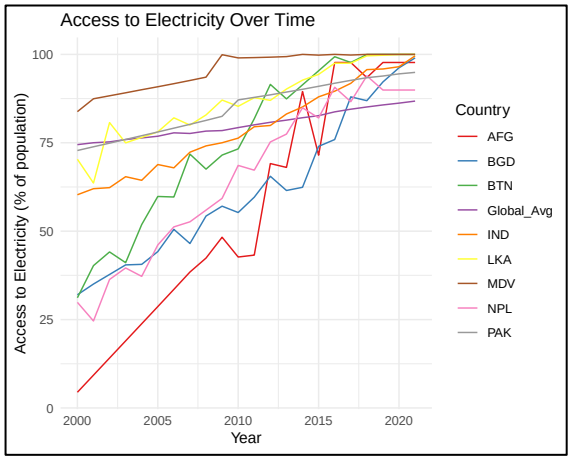

Figure 6: Multiple Line Chart for % population with access to electricity of South Asian countries and global average from 2000 to 2021

In Figure 5, a line chart describes the percentage of the population with electricity access in India from 1993 to 2021. During this time frame, there has been a steady increase in electricity access, going from 50% in 1993 to almost 100% (99.57%) in 2021. This remarkable trend indicates the effort of the policy decisions to expand access to electricity services. Figure 6, which is a multiple-line chart of the percentage population with access to electricity in South Asian countries and globally, shows an upward trend for each of the countries from 2000 to 2021. In the year 2021, Bhutan (BTN), Sri Lanka (LKA) and Maldives (MDV) have 100% of their population owning electricity, whereas the global average stands at 86%. Table 3 shows that the mean percentage population of Pakistan (PAK), Maldives (MDV) and Sri Lanka (LKA) are respectively 85.04%, 95.62% and 86.72% over the years, which are greater than the mean of global average percentage of population with access to electricity (80.20%).

Country	*Mean*	*Median*	*Min*	*Max*	*Q1*	*Q3*
AFG	55.82	45.75	4.45	97.70	29.89	92.45
BGD	61.59	58.34	32.00	98.99	44.80	75.44

BTN	75.30	77.53	31.15	100.00	59.71	97.11
IND	78.80	77.91	60.29	99.57	68.14	89.19
LKA	86.72	87.05	63.60	100.00	80.18	96.71
MDV	95.62	99.15	83.80	100.00	91.07	99.98
NPL	64.96	67.93	24.60	93.92	47.36	86.19
PAK	85.04	87.48	72.82	94.92	78.33	91.62
Global_Avg	80.20	79.69	74.49	86.78	77.05	83.48

Table 3: Summary Statistics of % population with access to electricity over 2000-2021 for SA countries and global average

- Annual freshwater withdrawal

The data for annual freshwater withdrawals as a percentage of internal resources of India is available from 1975 to 2020.

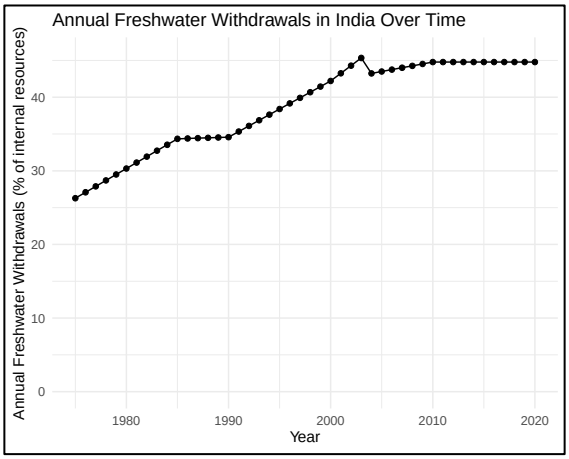

Figure 7: Line Chart for % of freshwater withdrawal for India over 1975-2020

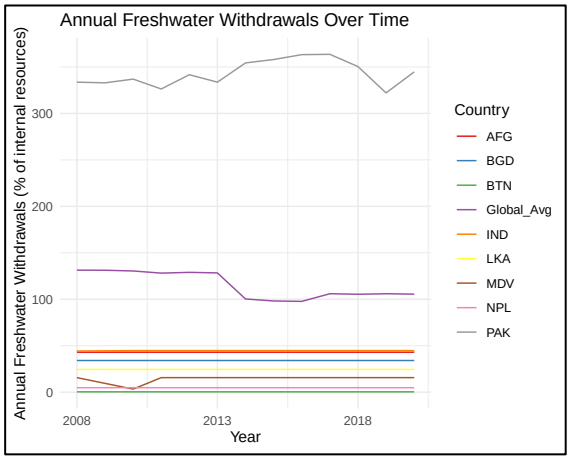

Figure 8: Multiple Line Chart for % of freshwater withdrawal for India over 1962-2020

The line chart in Figure 7 shows the percentage of freshwater withdrawal relative to total internal resources in India from 1975 to 2020. It has an overall steady upward trend apart from 2003 to 2004 when the annual freshwater withdrawal reduced from 45% to 43%. From Figure 8, Pakistan's (PAK) annual withdrawal is the highest, which is around 300%, and greater than that of the global average. India (IND), Sri Lanka (LKA), Bangladesh (BGD), Afghanistan (AFG) and Bhutan (BTN) have a steady withdrawal percentage over the years from 2008 to 2020, where Bhutan (BTN) is at the lowest.

Country	*Mean*	*Median*	*Min*	*Max*	*Q1*	*Q3*
AFG	55.82	45.75	4.45	97.70	29.89	92.45
BGD	61.59	58.34	32.00	98.99	44.80	75.44
BTN	75.30	77.53	31.15	100.00	59.71	97.11
IND	78.80	77.91	60.29	99.57	68.14	89.19
LKA	86.72	87.05	63.60	100.00	80.18	96.71
MDV	95.62	99.15	83.80	100.00	91.07	99.98
NPL	64.96	67.93	24.60	93.92	47.36	86.19
PAK	85.04	87.48	72.82	94.92	78.33	91.62
Global_Avg	80.20	79.69	74.49	86.78	77.05	83.48

Table 4: Summary Statistics of % freshwater withdrawal relative to internal resources over 2008-2020 for SA countries and global average

- Electricity from renewable sources

The data for the percentage of electricity production from renewable sources is available from 1985 to 2015. The data for the same of other South Asian countries are not available in the World Bank repository. The data for India post-2015 are also missing due to unknown reasons.

In Figure 9, the line chart of the production of electricity from renewable sources (% of total) in India shows an upward trend from the year 1998, though the overall percentage of electricity production from renewable sources remains less than 6%.

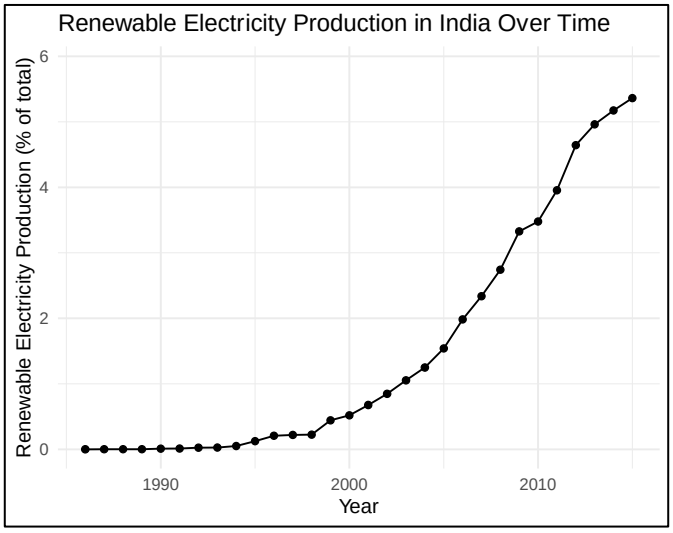

Figure 9: Line Chart for % of electricity production from renewable sources for India over 1986-2015

- Access to safely managed sanitation service (% of population)

The data for the percentage of the population in India using safely managed sanitation services is available from 2000 to 2022. Among the South Asian countries, the data on the same is available for Bangladesh (BGD), Bhutan (BTN), India (IND) and Nepal (NPL) from 2000 to 2022.

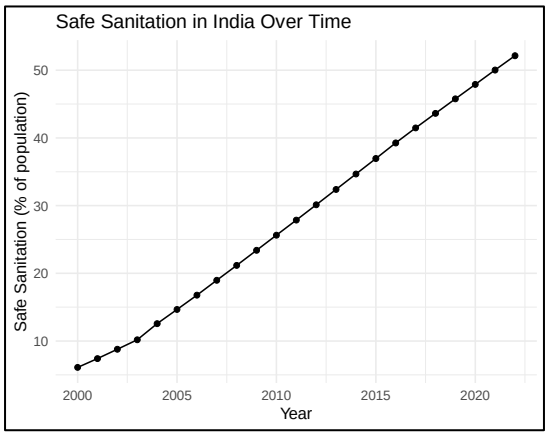

Figure 10: Line Chart for % of population with safely managed sanitation in India over 2000-2022

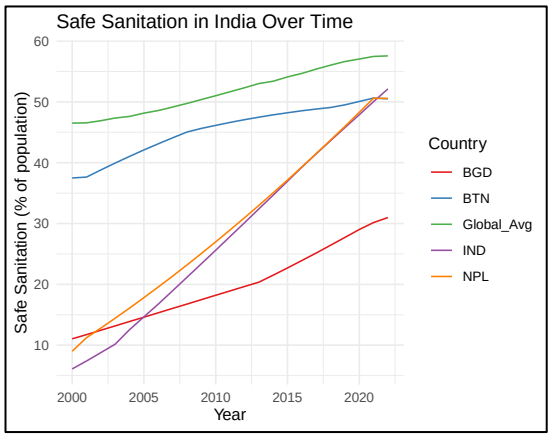

Figure 11: Multiple Line Chart for % of population with safely managed sanitation in South Asian countries and global average over 2000-2022

In the line chart demonstrating the percentage of the population with safely managed sanitation services in India as shown in Figure 10, there is an explicit upward trend from 2000 to 2022. A comparable increasing pattern can be seen for the global average population with safe sanitation services and the same for the South Asian countries. For example, in India, less than 10% of the population had access to properly managed sanitation services in 2000 and it steadily increased since then, with more than 50% of the population availing of safe sanitation services in 2022. Similarly, Nepal (NPL) had around 10% of their population who could use safe sanitation which rose to almost 50% in 2022, whereas Bangladesh (BGD) also had around 10% of the population who had access to safe sanitation services, however, the growth of the same is around 30% in 2022.

Country	*Mean*	*Median*	*Min*	*Max*	*Q1*	*Q3*
BGD	19.88	18.92	11.05	30.98	14.96	24.54
BTN	45.46	46.63	37.48	50.62	42.61	48.68
IND	28.16	27.86	6.10	52.14	15.72	40.36
NPL	29.68	28.94	9.00	50.61	18.69	40.40
Global_Avg	51.80	51.67	46.52	57.56	48.37	55.03

Table 5: Summary Statistics of % population with safe sanitation access over 2008-2020 for SA countries and global average

B. Correlation among the variables

In Figure 12, a heatmap or correlation matrix is presented to represent the pairwise correlations between the variables corresponding to India. The variables in Figure 12, which are displayed in both the x and y axis are:

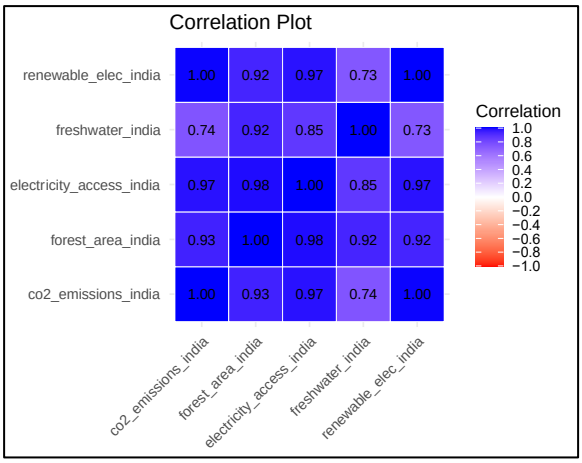

Figure 12: Correlation plot of the variables

From Figure 12, most of the variables have a strong positive correlation (pairwise) with the co2_emissions_india variable, whereas freshwater_india has a moderately strong positive correlation with the same. Renewable_elec_india has the perfect positive linear association with co2_emissions_india.

C. Regression Analysis

In this part, we are regressing the logarithm of CO2 emissions on the logarithm of forest areas and the logarithm of access to electricity in India. The details of the model are:

- Dependent variable: logarithm of CO2 emissions (metric tonnes per capita)
- Independent variables: logarithm of forest areas (% of land areas) and logarithm of access to electricity (% of population)
- Number of observations: 28

The regression equation, with 96% adjusted R^2, we obtained from the analysis is given below (details of the models are shown in Figure 13):

$$\log(CO2\ emissions\ India) = 5.92 + (-4.78) * \log(forest\ areas\ India) + 2.17 * \log(population\ with\ electricity\ access)$$

```
Call:
lm(formula = log_co2_emissions ~ log_forest_areas + log_electricity_access,
    data = reg_df)

Residuals:
     Min       1Q   Median       3Q      Max
-0.10962 -0.04264  0.01031  0.04433  0.09756

Coefficients:
                      Estimate Std. Error t value Pr(>|t|)
(Intercept)            10.2122     8.6844   1.176   0.2507
log_forest_areas       -6.3929     3.4086  -1.876   0.0724 .
log_electricity_access  2.3492     0.4794   4.901 4.82e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.05855 on 25 degrees of freedom
Multiple R-squared:  0.9673,    Adjusted R-squared:  0.9647
F-statistic: 369.9 on 2 and 25 DF,  p-value: < 2.2e-16
```

Figure 13: Regression of log of CO2 emissions on log of forest areas and log of electricity access in India

From the results of the regression, the following can be interpreted:

- If forest areas of India grow by 1%, per capita CO2 emissions will decrease by 6.39%
- If the percentage of population with access to electricity grows by 1%, per capita CO2 emissions will increase by 2.34%
- Adjusted R^2 is 96%, which means 96% of the variance in CO2 emissions (log) is explained by the considered independent variables
- The F-statistics is 369.9 with degrees of freedom 2 and 25, and the p-value is less than 0.05. Hence, the slope coefficients are jointly significant
- Among the two individual coefficients, the p-value of log_electricity_access is less than 0.05, which is significant

V. Summary and Conclusion

In conclusion, the project sheds light on India's performance on environmental indicators compared to the global average and other South Asian countries over the years. While India has almost 100% of the population with access to electricity, there are scopes to increase renewable energy production or the forest areas of the country. Further analysis can concentrate on the comparative study among the higher, medium and lower-income countries for insights and effective change-making. As the project shows, constant monitoring, evaluation and calls to action are salient steps for India to adapt. Considering some careful changes in the developmental policy planning may be helpful to reach the collective SDGs.

VI. Bibliography

1. The Little Green Data Book, The World Bank 2017
2. Sustainable Development, The Department of Economics and Social Affairs, the United Nations, 2015 Paris Agreement
3. *Sustainable Development Goal interactions through a climate lens: a global analysis*, Adis Dzebo, Zoha Shawoo, SEI Report, February 2023
4. The World Bank data repository for India: link
5. CO2 emissions (metric tonnes per capita) – India, world Bank Data, Climate Watch Historical GHG Emissions, 2023, License Attribution-Non Commercial 4.0 International (CC BY-NC 4.0), link
6. Forest Area (% of land area) – India, World Bank Data, Food and Agriculture Organisation, electronic files and website, License CC BY-4.0, link
7. Access to electricity (% of population) – India, IEA, IRENA, UNSD, World Bank, WHO. 2023. Tracking SDG 7: The Energy Progress Report. World Bank, Washington DC, License: Creative Commons Attribution—Non Commercial 3.0 IGO (CC BY-NC 3.0 IGO), link
8. Annual freshwater withdrawals, total (% of internal resources) – India, Food and Agriculture Organization, AQUASTAT data, License CC BY-4.0, link
9. Electricity production from renewable sources, excluding hydroelectric (% of total) – India, IEA Statistics © OECD/IEA 2014, link
10. People using safely managed sanitation services (% of population) – India, WHO/UNICEF Joint Monitoring Programme (JMP) for Water Supply, Sanitation and Hygiene, link

VII. Appendix

- Details of the data considered for this project:
  - The data on **CO2 emissions**, stemming from the burning of fossil fuels and the manufacture of cement, is measured in metric tonnes per capita. This indicator can be tied to the *SDG 9: Industry, Innovation and Infrastructure*. The data is sourced from Climate Watch Historical GHG Emissions (1990-2020) and is available in the data repository of World Bank[5]. In this analysis, the focus is on India, for which the data ranges from the year 1990 to 2020.
  - The data on **forest areas**, measured as percentage of land areas. Here, the forest areas mean the natural or planted stands of trees, excluding the trees in agricultural production system and trees in urban parks or gardens. This indicator can be used as one of the metrics to measure *SDG 15: Life on Land*. This dataset is sourced from Food and Agriculture Organisation and is available in the data repository of the World Bank[6]. For India, the data ranges from 1990 to 2021.

- The data on the **percentage of population with access to electricity** is available from 1993 to 2021 for India. This indicator can be tied to the *SDG 7: Affordable and Clean Energy*. The data is collected from the industry, national surveys and international sources and is available in the data repository of the World Bank[7].
- The **total annual freshwater withdrawals**, measured in percentage of internal resources, ranges from 1975 to the year 2020 for India. The indicator is tied to the *SDG 6: Clean Water and Sanitation*. The data is collected from the Food and Agriculture organisation and is available in the data repository of the World Bank[8]. It is crucial to note that withdrawals can exceed 100 percent of total renewable resources where extraction from non-renewable aquifers or desalination plants is considerable or where there is significant water reuse.
- The **percentage electricity production from renewable sources** which include geothermal, solar, tides, wind, biomass, and biofuels sources (excluding hydroelectric) is available from 1971 to 2015, though the initial years (from 1971 to 1985) have zero entries. This indicator is a potential metrics for *SDG 7: Affordable and Clean Energy*. The data is collected from IEA Statistics and is available in the data repository of the World Bank[9].
- The **percentage of the population using safely managed sanitation services** ranges from 2000 to 2022. This indicator is a potential metrics for *SDG 6: Clean Water and Sanitation*. The data is sourced from WHO and UNICEF joint monitoring programme for water supply, sanitation and hygiene and is available in the data repository of the World Bank[10].

- In the correlation analysis, the variables used are:
  - co2_emissions_india: CO2 emissions (metric tonnes per capita)
  - forest_area_india: Forest area (% of land area)
  - electricity_access_india: Access to electricity (% of population)
  - freshwater_india: Annual freshwater withdrawal (% of internal resources)
  - renewable_elec_india: Electricity production from renewable sources (% of total)
  - safe_sanitation_india: People with access to safe sanitation services (% of population)

- R^2 is the proportion of the variance of y_t, the dependent variable explained by the regression. $R^2 = 1 - \frac{\sum \widehat{u_t^2}}{\sum (y_t - \bar{y})^2}$, where $\sum \widehat{u_t^2}$ is the unexplained part of the model. R^2 increases as we include variables in the model, hence adjusted R^2 is adjusted for the number of predictors in the model.

- The test statistics of joint hypothesis that all coefficients are equal to zero is $\frac{(RRSS - URSS)/m}{URSS/(N-k)} \sim F(m, N-k)$ where RRSS = restricted residual sum of squares and URSS = unrestricted residual sum of squares.

- The null hypothesis for the individual slope coefficients is, H_0: $\beta = 0$, and if the p-value of the test statistic is lower than 5%, significance level (or 1%), then we fail to accept the null hypothesis, and conclude that the coefficient is significant.