Research Report

SAP Sustainability & Multidimensional

Prepared by: Justin Kim, Joseph Lim, Paul Park
Date: Monday, February 24th, 2025
Github Link: https://github.com/paulpark6/SAP_Datathon

Table of Contents

Abstract

The COVID-19 pandemic has exacerbated global poverty, reinforcing its status as a wicked problem—a complex social challenge with no simple solution. To better understand poverty through a data-driven approach, we leverage global sustainability and socioeconomic indicators from the SAP dataset to construct a Multidimensional Poverty Index (MPI) that captures key aspects of poverty at the country level. Using this index, we apply unsupervised and supervised machine learning techniques to identify the most influential indicators contributing to poverty. Our findings suggest that indicators related to the following topics are the most significant determinants of poverty: environment, healthcare and education. Furthermore, our analysis shows Eritrea, Somalia, the Solomon Islands, South Sudan, and Equatorial Guinea as the most affected countries. Using the topics we found as most influential, we propose data-driven policy interventions towards these affected countries around the top three topics, addressing key indicators such as electricity access, clean water availability, and primary school completion rates. By aligning these recommendations with SAP's sustainability efforts, stakeholders are better equipped to drive meaningful and effective poverty reduction initiatives in the world's most vulnerable regions.

1.0 Introduction

Poverty is a wicked problem—a deeply complex and multifaceted issue with no single cause, simple solution, or universally agreed-upon definition. It extends beyond financial hardship, intertwining with education, health, infrastructure, and environmental conditions. These factors are deeply interconnected, shaping and reinforcing one another in a persistent cycle that is difficult to break. Since poverty does not follow a simple cause-and-effect relationship, traditional economic measures fail to capture its full complexity. According to the United Nations, they define poverty as more than a lack of income and resources, encompassing hunger, malnutrition, limited access to education and services, social exclusion, and lack of participation in decision-making [1].

The UN's Sustainable Development Goals Report in 2020 observes that the COVID-19 pandemic has significantly reversed decades of positive progress in poverty, health, and education. Among some findings, an estimated 71 million people have been pushed back into extreme poverty, the first increase in poverty since 1998 [2]. The pandemic has further exacerbated pre-existing inequalities as well, with the worst impact on the poorest and most vulnerable countries, including disproportionately harming women and children [3]. These disruptions pose essential challenges to achieving the Sustainable Development Goals by 2030. Overall, this highlights poverty's multidimensional nature and the need for a holistic approach to addressing it.

Rather than attempting to solve poverty outright, this project focuses on measuring and understanding it through a data-driven approach. By leveraging global sustainability and socioeconomic indicators, we aim to develop a MPI. This tool provides a holistic framework for comparing poverty levels across countries, highlighting key areas of deprivation and revealing patterns in its underlying causes. To further analyze these patterns, we employ both supervised and unsupervised machine learning methods to determine the most influential indicators of poverty.

For the unsupervised learning approach, we first cluster countries into two groups using the MPI, categorizing them as "lower" and "upper" class. We calculated the absolute difference in averages for each indicator between two groups. Then, we conducted t-tests for each feature to test whether or not the mean MPI of the "lower" and "upper" class is statistically significant. So each indicator has the mean MPI of "lower" and "upper" class for all countries. Features with statistically significant differences are considered "important." To refine our understanding, we analyze the magnitude of variability between clusters for the important features, allowing us to rank them based on their impact on poverty classification.

Building on these clusters, we then apply a supervised learning approach by framing the problem as a binary classification task. Each row in the dataset represents a country, with indicators from 2022 and 2023 serving as features, while the assigned clusters act as the target variable. To classify countries into the upper or lower class, we train a logistic regression and random forest model. By evaluating model performance, we identify the best-performing model and analyze its feature importance values to determine which indicators have the strongest predictive power.

By combining insights from both approaches, we construct a comprehensive depiction of the key indicators driving poverty. We expect that the model will perform well in terms of accuracy in indicating the key contributing indicators to poverty. These findings serve as a foundation for targeted policy recommendations, supporting more effective strategies for sustainable and equitable development.

2.0 Methodology

2.1 Data Exploration

2.1.1 Data Scope

When brainstorming approaches to constructing a MPI, we initially considered incorporating external data to establish a reference ranking of poverty. However, we ultimately decided against this approach. One key reason was the lack of transparency regarding the data and methodology used to create pre-existing rankings. Training our model based on these rankings could introduce bias and potentially skew our analysis toward the assumptions embedded in external datasets. Additionally, we could not ensure that the provided indicators fully captured all the factors considered in those rankings. To maintain objectivity and consistency, we chose to rely exclusively on the given dataset, ensuring that our evaluation of poverty-related indicators was derived solely from the data at hand.

2.1.2 Basic Data Cleaning:

We first checked the data for basic inconsistencies, developed a strategy for filling missing values for rates between **2000 and 2023**, and ultimately decided to use recent years for analysis. Initially, we considered splitting the data into **pre-COVID, during COVID, and post-COVID** periods. The COVID-19 pandemic has significantly reversed decades of progress in poverty reduction, emphasizing the urgency to re-evaluate the current poverty landscape. In light of this, we decided to cater our analysis on recent years, specifically years **2022 and 2023**, which allows us to assess the immediate aftermath and recovery efforts.

We have also made an assumption on countries, that all of them are independent across each other. To address missing values, we applied **forward filling**, where missing values were replaced using the last available data from previous years. If data was still missing after forward filling, we then applied **backward filling**, which fills gaps using future available values. This approach ensured that trends remained continuous while minimizing artificial data creation. One country name that we noticed in that data was titled **“Not classified”**, which had no data available for the entire period from **2000 to 2023.** As a result, we simply removed it from our dataset. Table 1 concisely shows the scenarios and the actions we took for each.

Table 1: Data Scenario with the Respective Action

Data Scenario	Action Taken
Missing values in certain years but existing historical/future values available	Forward fill, then backward fill
No data available for the entire period (2000-2023)	No filling applied (excluded from imputation)
Partially missing values across multiple years	Filled using available values from the dataset

2.1.3 Indicator Exploration

After handling missing data, we conducted an initial analysis to assess indicator availability for all the countries. Figure 1 presents a box plot illustrating the distribution of indicators across countries, highlighting key statistical points and the number of countries at each level. The dataset includes 88 indicators in total, and our analysis provides insight into how comprehensively different countries track poverty-related metrics.

Our findings reveal moderate variation in indicator usage, with most countries utilizing between 49 and 81 indicators. The average number of indicators per country is 63, closely aligning with the median of 64, suggesting a fairly balanced distribution. However, 25% of countries rely on fewer than 49 indicators, while another 25% exceed 81 indicators, indicating disparities in data collection capacity. Notably, one country stands out as an outlier, with only 11 indicators, suggesting significant data limitations. On the other end of the spectrum, two countries track 87 indicators, reflecting the most comprehensive data collection efforts.

Additionally, we identified a unique case where the indicator "Population, total" was recorded for only a single region: "Africa Eastern and Southern." This suggests potential inconsistencies or differences in data reporting practices. Overall, while most countries follow a relatively standardized approach to indicator adoption, some exhibit notable gaps or extensive tracking efforts.

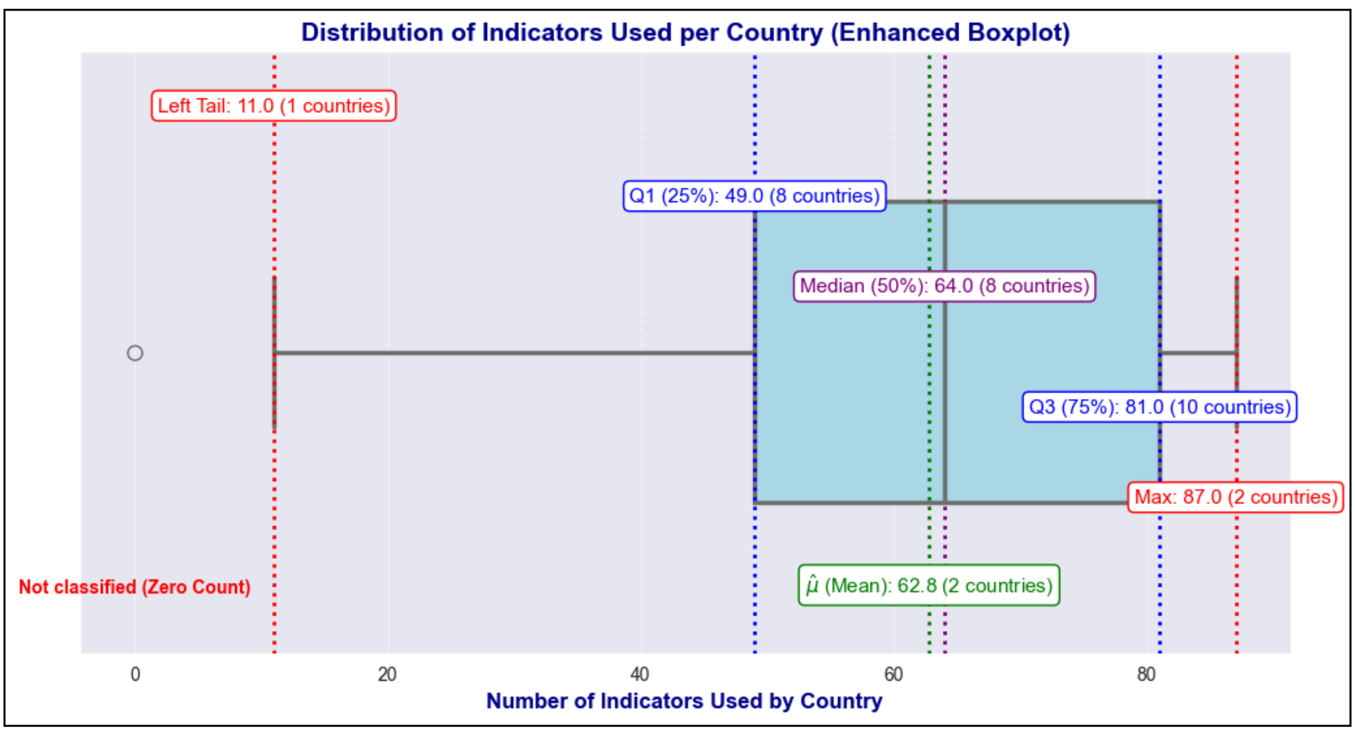

Figure 1. Distribution of Indicators Used per Country

2.2 Initial Multidimensional Index Structure

The primary objective of our ranking method is to systematically assess and rank countries using a set of indicators relevant to multidimensional poverty. This ranking serves as a foundational step for model selection and further analysis. By establishing a structured ranking system, we create a framework for identifying and classifying countries based on poverty levels, enabling us to highlight key areas for improvement that align with SAP's sustainability efforts.

2.2.1 "Less Poverty" vs "More Poverty"

At this stage, our team conducted a thorough analysis of all 88 indicators that could impact multidimensional poverty. Each indicator was carefully examined to determine whether a higher value corresponded to higher or lower levels of poverty. This assessment was based not only on our intuition but also on external research, including scholarly articles, reports from international organizations, and statistical studies. These sources helped clarify the direction and impact of indicators that were not immediately apparent. When going through each indicator one by one, we noticed that they fell under one of the three cases: Intuitive, Intuitive but Needs Confirmation, and Requires Further Research. Below shows examples of indicators that we believed fell into each category:

- **Intuitive:** Access to Electricity (% of population)
- **Intuitive but Needs Confirmation:** Children in Employment (% of children ages 7-14) [4]
- **Requires Further Research:** Urban Population (% of total population) [5]

For the indicators that required further research due to their unclear impact on poverty, we excluded them from further analysis, resulting in 78 total indicators that will be used for future steps.

2.2.2 Multidimensional Index Structure

Following a thorough evaluation of the indicators, work on the MPI began. The process involved several important steps:

1. **Yearly Ranking:** Each year, countries were ranked based on the finalized dataset from Section 2.1. The ranking methodology was adjusted based on the type of indicator. For some indicators, higher values represent better performance in poverty reduction (e.g., Access to Electricity (% of population)), while for others, lower values indicate better outcomes (e.g., Children out of Schools). To ensure consistency, countries with smaller ranks for a given indicator and year were considered to be performing better relative to others, meaning lower ranks correspond to lower poverty levels.

2. **Indicator-Level Averaging:** To create a comprehensive measure of each country's performance across time, we averaged the yearly ranks for each indicator. This ensured that each indicator was given equal weight in the final score. As a result, we obtained a mean rank per country for each indicator. A lower mean rank for an indicator suggests that a country performs better in that dimension of poverty relative to others.

3. **Country-Level Averaging:** Finally, to determine an overall MPI for each country, we averaged the ranks across all indicators. The *mean_rank* represents a country's overall position in terms of multidimensional

poverty. A higher mean rank indicates greater poverty, while a lower mean rank suggests better overall conditions.

As a result, we created a dataframe with the rankings for each country for all indicators. Ranking and scoring based on these indicators culminate into a robust MPI that reflects the multifaceted nature of poverty to provide a robust basis for policy recommendations in subsequent analyses.

2.3 Final Multidimensional Index Structure

2.3.1 Rationale for Differential Weighting

Upon viewing the initial rankings, it became evident that all indicators did not have uniform availability of data for the 265 countries under observation. Some of the indicators had data for all every country, but others were less covered. The differential coverage of data can bring about a bias towards more covered indicators if managed poorly since those with limited data can be skewed or less representative. To offset this and give an equal weight to all indicators in our last rankings, we decided to assign weights to each indicator differently based on the extent of their data coverage. The differential weighting is to reinforce the MPI's robustness by reducing the influence of indicators with sparse data and giving precedence to those with more extensive data coverage.

2.3.2 Weighting Methodology

The weight of each indicator was determined by the proportion of countries with available data out of the total countries covered in the study. Normalization is the relative weight of each indicator, with those having more data coverage having proportionately more influence in the poverty index. We then used the pre-calculated index in Section 2.2.2 and scaled it for each country by multiplying each indicator by its respective weights. We used this scaled ranking to compute the final mean rank for each country, which is the foundation of the final rankings used in model selection.

2.3.3 Building the Final Rankings DataFrame

The final dataframe *df_transformed.csv* generated as part of the analysis and rankings contains a full panel of indicators for several dimensions of poverty and development for all countries. The rows of the dataframe represent each country, and the column of the dataframe is a specific indicator for both 2022 and 2023 along with the MPI. The MPIs were ranked in **ascending order of the mean ranking**, and the country that had the lowest mean ranking was ranked as least affected by the multidimensional measures. This ranking provides a crucial foundation for model selection because it helps create countries that require urgent intervention and facilitates targeted and effective policy recommendations.

2.4 Machine Learning Modelling

2.4.1 Unsupervised Learning Method

For the unsupervised learning method, we decided to leverage K-Means clustering to cluster countries into meaningful groups. The idea behind this was to group countries with similar MPI index so that each group would have similar variances within each group. We stress tested multiple values of n (1 to 10) trying to find multiple groups and it suggested an optimal cluster size of three; the three clusters could be viewed as Lower, Middle, and Upper class. However, we noticed that the clusters were extremely imbalanced and dividing into two groups gave more balanced groups in terms of counts. By doing so, we were able to simplify the groups into

Upper and Lower Class. This decision allowed for a more robust classification while still capturing meaningful differences between countries.

After clustering, we classified 152 countries into the Upper class and 113 countries into the Lower class, adding this classification as a new column to the dataset for further analysis. When analyzing the clusters, Countries such as Singapore, Switzerland, Cayman Islands, South Korea, and Netherlands were assigned to the Upper class. On the other hand, countries like Afghanistan, Africa Eastern and Southern, Africa Western and Central, American Samoa, and Angola were assigned to the Lower class. This confirms that our clusters have some meaning behind them. To determine which indicators best differentiate between Upper and Lower class countries, we performed the following:

1. **Calculated the average mean_rank per indicator** for each cluster.
2. **Computed the difference (diff)** between the two groups to measure the **magnitude of separation**.
3. **Performed a t-test** for each indicator to determine **statistical significance** in distinguishing Upper and Lower class countries.

Each country was treated as an **independent sample**, and we tested each indicator under the following **null hypothesis**:

"The mean value of this indicator is equal in both Upper and Lower class countries."

A **p-value threshold of 0.05** was used to determine statistical significance.

After running the t-tests, we only retained the indicators with a statistically significant difference (p-value < 0.05). Then an inner join was performed between the **ranked list of indicators based on magnitude of difference** and the **set of indicators with statistically significant differences.** After conducting the inner join, we were able to come up with the final list which was sorted in **descending order** to ensure that the **most influential indicators** appear at the top.

2.4.2 Supervised Learning Method

After obtaining the clusters from the previous step, we explored a different approach by transforming the problem into a supervised machine learning task to extract feature importance. Using the selected indicators from 2022 and 2023 as features and the clusters as the target variable, we formulated a binary classification problem to predict whether a country belongs to the “Upper” or “Lower” class.

In our dataset, each row represents a country, with the indicators from 2022 and 2023 as features and the class label (0 or 1) as the output. During the earlier stages of data exploration and cleaning, we decided not to impute values for indicators that were entirely missing across all years. As a result, some countries have missing values for certain features. To handle these missing values, we use a K-Nearest Neighbors (KNN) imputer, which estimates missing values based on the nearest neighbors in the dataset.

To build our classification model, we experiment with two machine learning models that are suitable for classification and provide insights into feature importance: logistic regression, and random forest. For tabular data, tree-based models, such as random forest, often outperform linear models due to their ability to capture complex interactions between features. Unlike logistic regression, which assumes a linear decision boundary, tree-based models can learn non-linear relationships and automatically handle feature interactions without the need for manual feature engineering. Additionally, they are robust to outliers and missing values, making them particularly effective for real-world datasets with varying data quality.

Given these advantages, we expect the tree-based model to outperform logistic regression. However, model performance is always dataset-dependent, so we include logistic regression. By evaluating these models, we aim to determine the most important indicators influencing poverty classification and use these insights to guide further analysis.

For model training, we used an 80/20 split, where 80% of the data was allocated for training and 20% for testing. Our primary objective was to maximize accuracy, which serves as a general measure of how well the model classifies countries into the “Upper” and “Lower” categories.

Feature importance can be determined in different ways depending on the model. In logistic regression, it can be found using the absolute values of the model’s coefficients, where larger values indicate stronger influence on classification. In random forest, feature importance is obtained through the .feature_importances_ attribute, which measures how much each feature contributes to reducing impurity across decision trees.

In the next section, we will compare features across both the unsupervised and supervised learning method, to identify key indicators that consistently influence poverty classification.

3.0 Results and Discussion

3.1 Top 5 Affected Countries

Identifying the countries most affected by multidimensional poverty is a crucial outcome of our analysis. Based on the analysis and final rankings, Table 2 shows the countries are the most affected by multidimensional poverty in descending order. In addition, Figure 2 shows a heatmap of all countries, colored by the intensity of their respective MPI. One key aspect we notice is that countries associated with poverty lie in the continent of Africa, and that countries that are not associated with it lie in Europe.

Table 2: Top 5 Most Affected Countries

Country Name	MPI (mean_rank)
Eritrea	109.8964
Somalia	101.3538
Solomon Islands	101.2490
South Sudan	100.7740
Equatorial Guinea	100.2582

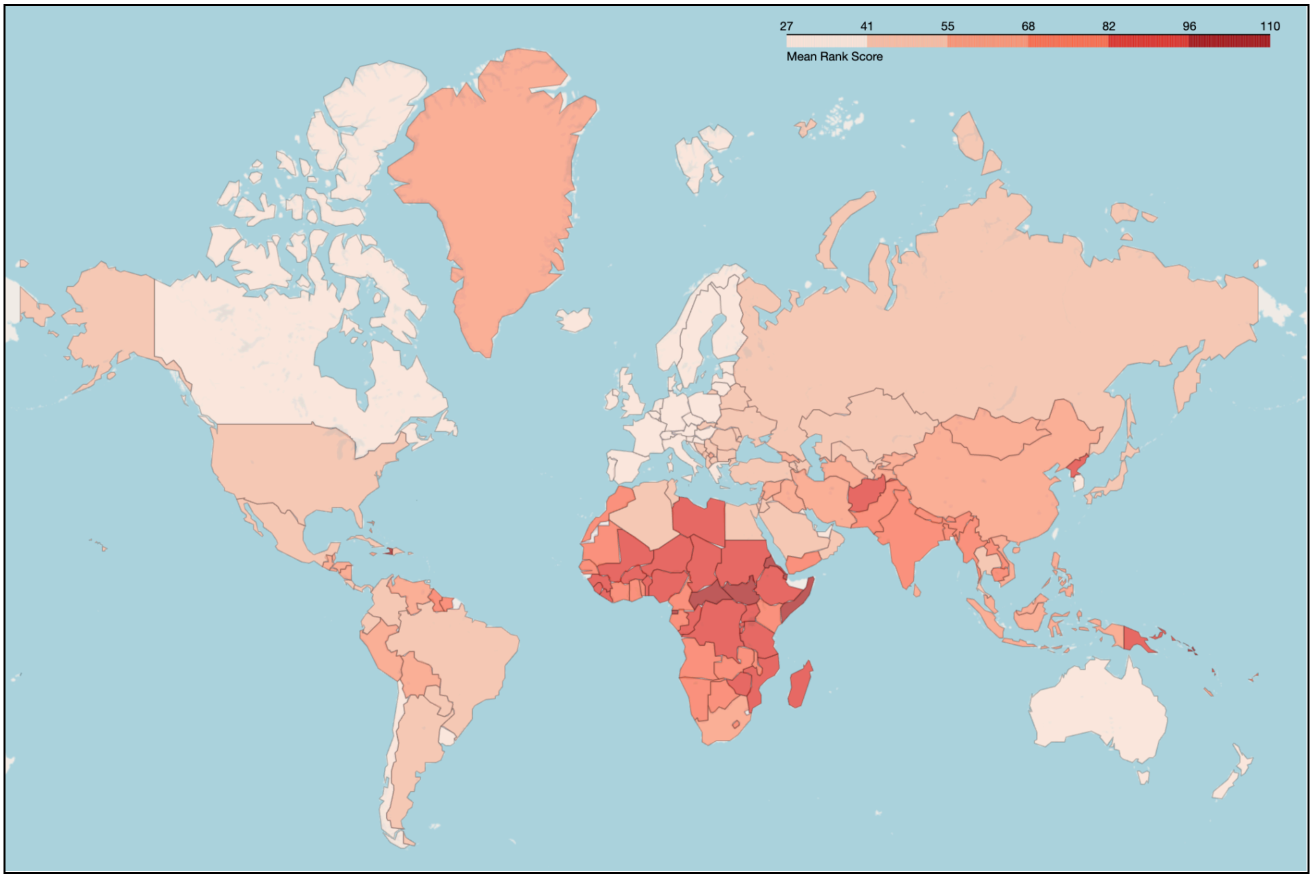

Figure 2: Heatmap of Countries Coloured by MPI

3.2 Unsupervised and Supervised Learning Results

Before we obtained feature importance using the supervised learning method, we first selected the best-performing model. Test accuracies for Random Forest and Logistic Regression were 96.2% and 92.5% respectively. As expected, the tree model outperformed Logistic Regression and thus we used Random Forest to obtain feature importance. Table 3 presents the top 10 most important features for both unsupervised and supervised methods in descending order. (See the complete list in FinfalFeatureImportance.xlsx in our github repository.) Surprisingly, the majority of these top features are environmental variables—a finding we explore in further detail in the next section.

Table 3: Top 10 Most Important Features for Both Methods

Unsupervised Method	Supervised Method
2022_Access to clean fuels and technologies for cooking (% of population)	2023_Access to electricity (% of population)
2022_Access to clean fuels and technologies for cooking, rural (% of rural population)	2022_Access to electricity (% of population)
2022_Access to clean fuels and technologies for cooking, urban (% of urban population)	2023_Access to electricity, rural (% of rural population)

2022_Access to electricity (% of population)	2022_Access to electricity, urban (% of urban population)
2022_Access to electricity, rural (% of rural population)	2022_Births attended by skilled health staff (% of total)
2022_Access to electricity, urban (% of urban population)	2022_Access to electricity, rural (% of rural population)
2022_Adequacy of social insurance programs (% of total welfare of beneficiary households)	2023_Births attended by skilled health staff (% of total)
2022_Adequacy of social protection and labor programs (% of total welfare of beneficiary households)	2023_Literacy rate, adult female (% of females ages 15 and above)
2022_Adjusted net national income per capita (current US$)	2022_Access to clean fuels and technologies for cooking, rural (% of rural population)
2022_Adjusted savings: education expenditure (current US$)	2023_People using at least basic drinking water services (% of population)

3.3 Combining Unsupervised and Supervised Learning Results

Both techniques produce a list of important features and their relative rank. In order to combine these results, we first make a ranking column for both techniques in such a manner that the rank is 1 for the most important feature. In order to combine the output of both the unsupervised and supervised learning techniques in an organized manner, we perform an inner join on year and feature to yield a combined dataset with the feature and its rank of importance from both techniques.

For comparability, we normalize the ranks using Min-Max scaling, scaling them to the range 0 to 1. The ranked values are added to get a global ranking of feature importance by both approaches. A value of nearer to 0 indicates higher overall importance, while a value of nearer to 1 indicates lower importance. Once the last set of most important features is identified, we examine the top 20 features and identify the topics they belong to. We examine the frequency of each topic in the top features to identify which topics are most common in identifying poverty classification.

3.4 Analyzing Topics and Features

Table 4 presents the 20 most important features along with their associated topics. For our analysis, we focused on environment, health, and education, which were among the most prevalent topics, comprising 7, 6, and 3 indicators, respectively. Our models captured more features in 2022 owing to the pandemic; the lower class suffered the most compared to the upper class countries which is the reason why there is a wide gap between the two groups. The lower class countries' weaker health facilities, lower financial buffers, and diminished safety nets created more severe economic repercussions and sluggish recoveries compared to wealthier countries. As a result, the gap between lower- and upper-income countries widened. The indicators measuring inequality showed significant differences in mean after the pandemic. In this section, we explore these key topics and examine their relationship with poverty through further research and analysis.

Table 4: Summary of 20 Most Important Indicators

Indicators
Economic Policy & Debt
National accounts: Adjusted savings & income
2022_Adjusted net national income per capita (current US$)
Education
Outcomes
2022_Literacy rate, adult female (% of females ages 15 and above)
2022_Literacy rate, adult male (% of males ages 15 and above)
2022_Literacy rate, youth male (% of males ages 15-24)
Environment
Energy production & use
2022_Access to clean fuels and technologies for cooking (% of population)
2022_Access to clean fuels and technologies for cooking, rural (% of rural population)
2022_Access to clean fuels and technologies for cooking, urban (% of urban population)
2022_Access to electricity (% of population)
2022_Access to electricity, rural (% of rural population)
2022_Access to electricity, urban (% of urban population)
2023_Access to electricity (% of population)
2023_Access to electricity, rural (% of rural population)
Health
Disease prevention
2022_People using at least basic drinking water services (% of population)
Health systems
2022_Current health expenditure per capita (current US$)
Population: Dynamics
2022_Age dependency ratio, old (% of working-age population)
Reproductive health
2022_Adolescent fertility rate (births per 1,000 women ages 15-19)
2022_Births attended by skilled health staff (% of total)
Social Protection & Labor
Economic activity
2022_Children in employment, wage workers (% of children in employment, ages 7-14)
2022_Children in employment, wage workers, male (% of male children in employment, ages 7-14)
2022_Children in employment, work only (% of children in employment, ages 7-14)

3.4.1 Environment

Access to reliable and modern energy services is a crucial factor influencing poverty levels. The "Environment: Energy Production & Use" category, particularly access to electricity and clean cooking fuels, plays a vital role in socioeconomic development.

Electricity access is fundamental to improving quality of life and economic opportunities. It enables better education through lighting, enhances healthcare services, and facilitates income-generating activities. However, as of recent data, approximately 1.18 billion people worldwide struggle to utilize electricity effectively and a total of 733 million people who lack any electricity connection at all [6]. This form of poverty is most prevalent in low-income regions, highlighting a strong correlation between electricity access and poverty levels [7].

In many developing countries, traditional biomass fuels such as wood and coal are still widely used for cooking. This practice leads to indoor air pollution, which causes approximately 3.7 million premature deaths annually [8]. Additionally, reliance on inefficient cooking fuels perpetuates poverty by negatively impacting health and limiting economic productivity. Transitioning to clean cooking fuels and technologies not only reduces health risks but also has the potential to lower income inequality and promote economic development [9].

Enhancing access to electricity and clean cooking solutions is essential for poverty alleviation. These factors are closely linked to improvements in health, education, and economic growth, which demonstrate their importance in sustainable development strategies.

3.4.2 Health

Our analysis indicated that health was the second most significant topic that influenced the MPI. This topic includes several key aspects within health: Reproductive Health, Health Systems, Disease Prevention, and Population Dynamics.

Reproductive health has remained a priority issue in the international health agenda, with a more specific emphasis on access and quality of maternal health care services. The World Bank reports far-reaching efforts at increasing services such as **skilled birth attendance by staff** and **reducing adolescent fertility rates**. For instance, significant progress has been made in increasing births attended by skilled health personnel across the globe, a key determinant of reducing maternal and neonatal mortality. Yet, challenges persist, particularly in Sub-Saharan Africa and South Asia, where adolescent fertility rates are still significant, impacting the health and economic opportunities for young women [10].

Global health systems have faced increased scrutiny, especially after COVID-19 revealed and worsened existing weaknesses. The World Bank's $27 billion global portfolio includes 160 projects on creating a more resilient health system to achieve universal health coverage (UHC) [11]. This investment is not only in health infrastructure but also looking to **make healthcare more affordable utilizing digital solutions.** Despite these investments to achieve universal health coverage, the majority of low and middle-income countries face fiscal pressures that can compromise sustained investments.

Disease prevention is instrumental in reducing the burden of both communicable and non-communicable diseases. The World Bank interventions are centred on encouraging universal access to **water and sanitation across 34 countries** in strengthening the capabilities of their water-related institutions, which are at the centre of prevention of countless diseases [12]. While progress has been achieved, the COVID-19 pandemic has shown the importance of more robust public health surveillance and response systems by emphasizing the application of preventive measures in controlling disease spread.

Population dynamics, such as age dependency ratios, significantly affect the sustainability and economic viability of healthcare systems. A constant increase in **old-age dependency ratio** poses a challenge to healthcare systems since aging populations require more medical and social care [13]. Such dynamics influence future healthcare demand planning and economic policy to adapt to a changing demographic structure.

The insights gained from our analysis on the health sector indicate the complex interaction between reproductive health care, health system strength, disease control activities and population trends which are all fundamental to transitioning individuals out of poverty and improving overall health in society.

3.4.3 Education

Another key topic that we found to be important was education, specifically the indicators related to literacy. Literacy is a key driver of economic and social well-being, making it one of the most critical factors in poverty reduction. Individuals with higher literacy skills have better access to employment opportunities, higher wages, and economic mobility, reducing their risk of poverty. Beyond income, literacy enables informed decision-making in healthcare, financial management, and civic participation, improving overall quality of life [14]. Conversely, low literacy rates can perpetuate cycles of poverty by restricting access to education, job opportunities, and essential services. Investing in literacy programs can help break these cycles, fostering sustainable development and economic resilience.

3.5 Policy Recommendations

In our comprehensive analysis of the three prevalent sectors of health, environment, and education, we aim to give comprehensive policy recommendations to address the needs of the top five most affected countries that we have obtained in our analysis. Additionally, we ensure that these recommendations align with SAP's sustainability goals, enhancing their feasibility and relevance.

In the realm of environmental sustainability, SAP's big data handling and analysis features can play a crucial role in monitoring renewable energy schemes. For example, SAP's analytics and IoT platforms are able to monitor the efficiency and output of renewable energy installations and monitor smart grids to ensure that environmental projects remain effective and sustainable.

For health care programs, SAP can provide robust healthcare analytics solutions that can facilitate health trend forecasting, management of health care supplies, and efficient allocation of resources. SAP's solutions can integrate data from telemedicine and mobile health clinics to enable health professionals access to patient records and availability of resources in real time, enhancing the quality of care provided even in remote areas.

In the field of education, SAP software solutions can be utilized to make educational institutions' management efficient and enhance the learning process in countries mentioned above. SAP SuccessFactors can effectively handle teacher recruitment and training procedures, while SAP's digital learning environments can provide scalable resources to enhance student access to quality education.

The combination of SAP's capability and technology with the comprehensive, data-driven, and tailored policy recommendations is to the effect that interventions not only become customized for the distinctive needs of the most affected countries but also scalable and sustainable. SAP's ability to process and analyze high levels of data can provide policymakers with the intelligence they require to make informed decisions, monitor the efficacy of interventions, and also make changes in real time, guaranteeing that they maximize benefit to those who need it most.

These recommendations will pave the way for transformative impacts in health, environment, and education, achieving long-term and sustainable development in the world's poorest regions. The strategic alignment ensures that SAP's technological innovation directly contributes to the global efforts aimed at eradicating poverty and overall sustainability.

4.0 Conclusion & Next Steps

The MPI analysis across the globe from 2022 and 2023 reveals that it constitutes a significant source of information into devastating and far-reaching effects of poverty, especially magnified by the COVID-19 pandemic, that have reversed years of gains and pushed millions into poverty. Through in-depth examination across health, environment, and education, we have pinpointed priority areas where immediate intervention is essential and where the capabilities of SAP can be effectively utilized.

Our findings emphasize the need for evidence-based, targeted policy interventions aligned with global sustainability goals. SAP's enterprise applications and analytics enable efficient deployment, monitoring, and optimization of these efforts. By leveraging these strengths, SAP and its collaborators can address critical challenges in the most affected regions.

In summary, tackling poverty requires urgent, sustained action. While the path is challenging, the right tools and partnerships can drive meaningful, long-term change. With its resources and expertise, SAP has the opportunity to lead efforts in creating a more sustainable and equitable world.

One limitation of our analysis is that we focused solely on data from 2022 and 2023. As a future step, we could do more research to incorporate historical data to uncover deeper patterns and trends. Another limitation lies in our initial approach to handling NULL values, since we primarily relied on forward and backward filling. For future next steps, more sophisticated imputation methods could be used to improve data accuracy and consistency across years. Additionally, while our policy recommendations were based on the most important indicators, moving forward, a more in-depth examination of these countries could help determine whether our selected indicators align with their specific needs, leading to more targeted and effective recommendations.

References

[1] United Nations, “Ending poverty,” United Nations, 2023.
https://www.un.org/en/global-issues/ending-poverty

[2] United Nations, “UN Report Finds COVID-19 Is Reversing Decades of Progress on poverty, Healthcare and Education,” United Nations, 2020.
https://www.un.org/nl/desa/un-report-finds-covid-19-reversing-decades-progress-poverty-healthcare-and

[3] World Bank, “COVID-19 leaves a legacy of rising poverty and widening inequality,” *World Bank Blogs*, Oct. 07, 2021.
https://blogs.worldbank.org/en/developmenttalk/covid-19-leaves-legacy-rising-poverty-and-widening-inequality

[4] Z. Naeem, F. Shaukat, and Z. Ahmed, “Child labor in relation to poverty,” International Journal of Health Sciences, vol. 5, no. 2 Suppl 1, p. 48, Jul. 2011, Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC3533357/

[5] “World Bank’s Assessment Warns of Rising Urban Poverty in Madagascar and Recommends Broad-Based and Sustained Growth,” World Bank.
https://www.worldbank.org/en/news/press-release/2024/02/21/world-banks-assessment-warns-of-rising-urban-poverty-in-afe-madagascar-and-recommends-broad-based-and-sustained-growth

[6] B. Min, “1.18 billion around the world are unable to use electricity,” World Bank Blogs, 2024.
https://blogs.worldbank.org/en/opendata/1-18-billion-around-the-world-are-unable-to-use-electricity

[7] H. Ritchie, M. Roser, and P. Rosado, “Energy Access,” Our World in Data, 2019.
https://ourworldindata.org/energy-access

[8] IEA, “Access to clean cooking – SDG7: Data and Projections – Analysis,” IEA, 2022.
https://www.iea.org/reports/sdg7-data-and-projections/access-to-clean-cooking

[9] S. Lee, “Clean cooking fuels have great potential to reduce income inequality, researchers find,” Phys.org, Feb. 10, 2025. https://phys.org/news/2025-02-cooking-fuels-great-potential-income.html? (accessed Feb. 24, 2025).

[10] Sexual and reproductive health and rights,
https://www.unwomen.org/sites/default/files/2023-11/unfpa_observer_paper_unfpa.pdf (accessed Feb. 24, 2025).

[11] World Bank, “Overview,” World Bank, 2017. https://www.worldbank.org/en/topic/health/overview

[12] W. B. Group, “The Global Water Security and Sanitation Partnership (GWSP) 2024 Annual Report,” World Bank, Nov. 15, 2024. https://www.worldbank.org/en/topic/water/publication/the-gwsp-2024-annual-report

[13] “Age dependency ratio, old (% of working-age population) | Data,” data.worldbank.org.
https://data.worldbank.org/indicator/SP.POP.DPND.OL

[14] Concern Worldwide US, “6 Benefits of literacy in the fight against poverty,” concernusa.org, Aug. 27, 2020.
https://concernusa.org/news/benefits-of-literacy-against-poverty/