# Clustering the Countries for HELP International

**LUM FU YUAN ( B032110251)** **TEH XIAO THONG ( B032110141)** **SIM WENG JIN ( B032110376)** **ANG WEI KANG ( B032110301)**

***Abstract*— IN today's interconnected world, international organizations play a crucial role in addressing global challenges and improving the lives of those in need. HELP International, an esteemed humanitarian NGO, is committed to fighting poverty and providing essential assistance to disadvantaged populations during times of disasters and calamities. With recent fundraising success, raising $10 million, the NGO's CEO faces the critical task of strategically and effectively allocating these funds. To achieve this, it is essential to identify countries that are in dire need of aid and prioritize their support. This project aims to categorize countries based on socio-economic and health factors in order to determine their overall development. This categorization will assist HELP International, an international humanitarian NGO, in strategically and effectively allocating their raised funds of $10 million. The aim is to identify the countries that are in the direst need of aid, allowing the CEO to make informed decisions on where to focus the organization's resources. To achieve this, we employ a data-driven approach using the Fuzzy C-means clustering algorithm. This algorithm allows us to group countries based on similarities in their socio-economic and health indicators. By identifying clusters of low values for indicators, we can pinpoint countries that need immediate attention and support. Through data preprocessing, feature transformation, and cluster analysis, we generate an unambiguous classification of countries. We visualize the clusters and highlight the countries in need of assistance. This information provides valuable insights to the CEO of HELP International, aiding in the decision-making process for the effective utilization of their $10 million funds. By focusing on countries identified as most deserving of assistance, NGOs can maximize their impact and contribute to the development and well-being of those communities. This program serves as a valuable tool for HELP International and other humanitarian organizations in optimizing their resource allocation strategies to ensure that aid is directed to countries in greatest need.**

## I. Introduction

HELP International, an international humanitarian NGO, has successfully raised $10 million to support its mission of alleviating poverty and providing assistance during times of disaster and calamities. With a substantial amount of funds at their disposal, the CEO of HELP International faces the critical task of deciding how to strategically allocate the funds to the countries in need. The objective is to identify countries most in need of assistance, ensure that funds are used effectively, and have a significant impact on targeted communities.

To achieve this, a data-driven approach is employed to leverage socio-economic and health factors that play a vital role in determining the overall development of a country. By analyzing key indicators such as child mortality, income levels, GDP per capita, life expectancy, and other relevant factors, it becomes possible to categorize countries according to their specific needs and challenges.

In this project, we will utilize the Fuzzy C-means clustering algorithm to categorize countries into distinct clusters based on their socio-economic and health indicators. This algorithm allows for the identification of patterns and similarities among countries, enabling a more informed decision-making process. By identifying clusters with lower values for critical indicators such as child mortality, income, and GDP per capita, we can pinpoint countries that require immediate attention and aid.

The outcome of this analysis will provide valuable insights to the CEO of HELP International, aiding in the prioritization of countries and the allocation of funds. By focusing resources on countries identified as having the greatest need, HELP International can maximize the impact of their humanitarian efforts, contributing to the improvement of living conditions and overall development in those communities.

Furthermore, the findings and methodology presented in this project can serve as a valuable tool for other humanitarian organizations facing similar resource allocation challenges. By utilizing data-driven approaches, NGOs can optimize their strategies and ensure that aid reaches those countries and regions that require it the most.

In the following sections, we will discuss the data, perform exploratory analysis, apply the clustering algorithm, and present the results. Through this analysis, we aim to assist HELP International in making informed decisions that will have a tangible and positive impact on the lives of people in need.

## II. GENERAL CONCEPT

Fuzzy logic, proposed by Lotfi Zadeh, is an approach to computing that extends traditional Boolean logic by introducing the concept of "degree of truth" instead of the binary "true" or "false" values. Unlike classical logic, which operates in a crisp, deterministic manner, fuzzy logic allows for the representation and manipulation of vague and imprecise information. Fuzzy C means (FCM) have adapted the concept of fuzzy logic in the clustering process.

### *A. Fuzzy C means (FCM)*

The fuzzy C-means (FCM) clustering algorithm, initially introduced by Dunn and later extended by Bezdek in the 1990s, is a data clustering technique that allows for more flexible and probabilistic cluster assignments. Fuzzy logic is a mathematical framework that deals with uncertainty and allows for the representation and manipulation of vague or ambiguous information. It is particularly useful when dealing with data that may not have clear-cut boundaries or when assigning membership degrees to multiple clusters. FCM aims to group a dataset into N clusters, where each data point is assigned a membership degree to each cluster.

In FCM, data points closer to the center of a cluster have higher membership degrees for that cluster, while data points farther away have lower membership degrees. This approach allows for a more nuanced representation of the data, capturing the uncertainty or ambiguity in cluster assignments.

The FCM algorithm is iterative and aims to find an optimal partition of the dataset by minimizing the weighted within-group sum of the squared error objective function. This objective function measures the total deviation of data points from their assigned cluster centers, taking into account the membership degrees. By adjusting the cluster centers and updating the membership degrees iteratively, the algorithm seeks to minimize this objective function.

The fuzzy logic concept in FCM helps in capturing the uncertainty and fuzziness in the data, allowing for a more nuanced representation of cluster membership. This is particularly beneficial when dealing with complex and overlapping patterns in the dataset, as it allows data points to belong to multiple clusters based on their degree of similarity to each cluster centroid.

## III. THE FRAMEWORK OF CLUSTERING THE COUNTRIES FOR HELP INTERNATIONAL

Clustering the Countries for HELP International aims to identify and assist countries in need based on various socio-economic indicators.

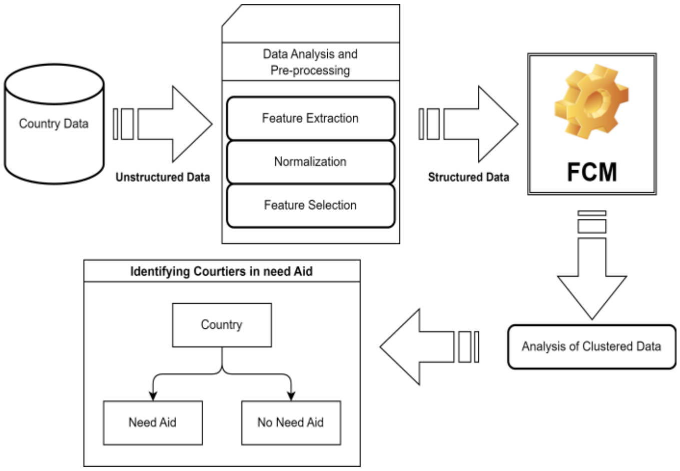


*Figure 1: The Framework for Clustering The Countries*

This framework encompasses several essential steps, starting with obtaining a dataset from a reliable source such as the Kaggle database.

Once the dataset is acquired, data analysis and pre-processing are carried out to ensure the dataset's quality and suitability for clustering. Duplicate and incomplete data items are eliminated to maintain data integrity. Furthermore, unnecessary attributes that do not contribute to the clustering process are dropped from the dataset. Feature extraction and normalization techniques are applied to transform the preprocessed dataset into a suitable format for clustering. This step involves leveraging Python's built-in libraries to extract meaningful features and normalize the data, ensuring that all features are on a similar scale.

The Fuzzy C-Means (FCM) algorithm is selected as the clustering method in this framework. FCM is a popular clustering algorithm that allows for the soft assignment of data points to clusters, considering the degree of membership rather than strict membership. The FCM model is initialized with a predetermined number of clusters (referred to as the C value) and fitted to the preprocessed data using the fit method. Subsequently, the cluster centers are obtained using the centers attribute of the FCM model. Each data point is assigned to a specific cluster based on its predicted membership using the prediction method.

Upon clustering the data, an analysis of the clustered data is performed. This analysis involves calculating the cluster means for each feature or indicator by grouping the data based on the predicted clusters. By comparing these cluster means with specific indicators such as 'child_mort', 'income', 'gdpp', and 'life_expec', clusters with lower values for these indicators are identified. These clusters are likely to represent countries in need of aid due to their unfavorable socio-economic conditions.

To pinpoint the countries in need of aid, the framework filters the dataset based on the identified clusters and extracts the corresponding country names. The list of countries in need is then displayed by printing it to the console. This provides a clear understanding of which countries require assistance based on the clustering analysis.

In conclusion, the framework for clustering the countries for HELP International involves obtaining a dataset, performing data analysis and pre-processing, applying the FCM clustering algorithm, analyzing the clustered data, and identifying the countries in need of aid. By following this framework, HELP International can effectively target its assistance efforts and make informed decisions based on the socio-economic conditions of different countries.

## IV. The Proposed Predictive Modelling

The Clustering model in this paper consists of four major components, i.e. (1) Data Acquisition and Discovery, (2) Data Preparation and Processing, (3) Fuzzy C-Means (FCM) Model, and (4) Result Analysis and Identifying Countries in need Aid.

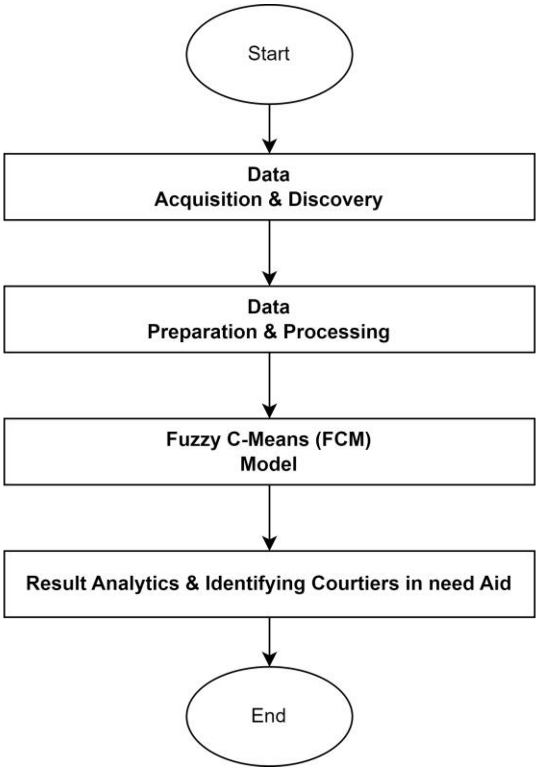


*Figure 2: The Flowchart for the Clustering Model*

### A. Data Acquisition and Discovery

The first step in the proposed predictive modeling framework is obtaining a dataset from a reliable source, such as the Kaggle database. This dataset should contain relevant socio-economic indicators for countries.

The dataset comprises several features that describe different aspects of each country. These features include:

**country:**
Name of the country

**child_mort:**
Death of children under 5 years of age per 1000 live births

**exports:**
Exports of goods and services per capita. Given as a percentage of the GDP per capita

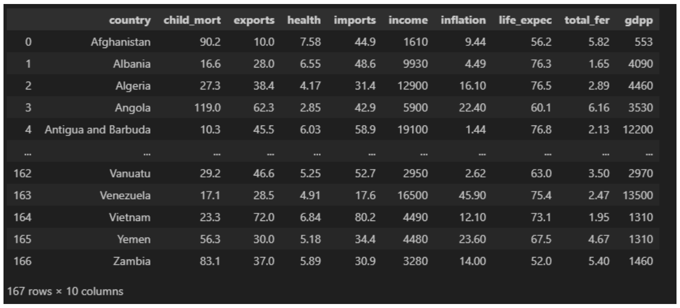
**health:**
Total health spending per capita. Given as a percentage of GDP per capita

**imports:**
Imports of goods and services per capita. Given as a percentage of the GDP per capita

**Income:**
Net income per person

**Inflation:**
The measurement of the annual growth rate of the Total GDP

**life_expec:**
The average number of years a newborn child would live if the current mortality patterns are to remain the same
**total_fer:**
The number of children that would be born to each woman if the current age-fertility rates remain the same.
**gdpp:**
The GDP per capita. Calculated as the Total GDP divided by the total population.

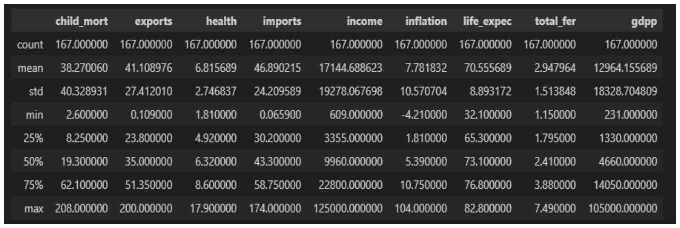


| | country | child_mort | exports | health | imports | income | inflation | life_expec | total_fer | gdpp |
|---|---|---|---|---|---|---|---|---|---|---|
| 0 | Afghanistan | 90.2 | 10.0 | 7.58 | 44.9 | 1610 | 9.44 | 56.2 | 5.82 | 553 |
| 1 | Albania | 16.6 | 28.0 | 6.55 | 48.6 | 9930 | 4.49 | 76.3 | 1.65 | 4090 |
| 2 | Algeria | 27.3 | 38.4 | 4.17 | 31.4 | 12900 | 16.10 | 76.5 | 2.89 | 4460 |
| 3 | Angola | 119.0 | 62.3 | 2.85 | 42.9 | 5900 | 22.40 | 60.1 | 6.16 | 3530 |
| 4 | Antigua and Barbuda | 10.3 | 45.5 | 6.03 | 58.9 | 19100 | 1.44 | 76.8 | 2.13 | 12200 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 162 | Vanuatu | 29.2 | 46.6 | 5.25 | 52.7 | 2950 | 2.62 | 63.0 | 3.50 | 2970 |
| 163 | Venezuela | 17.1 | 28.5 | 4.91 | 17.6 | 16500 | 45.90 | 75.4 | 2.47 | 13500 |
| 164 | Vietnam | 23.3 | 72.0 | 6.84 | 80.2 | 4490 | 12.10 | 73.1 | 1.95 | 1310 |
| 165 | Yemen | 56.3 | 30.0 | 5.18 | 34.4 | 4480 | 23.60 | 67.5 | 4.67 | 1310 |
| 166 | Zambia | 83.1 | 37.0 | 5.89 | 30.9 | 3280 | 14.00 | 52.0 | 5.40 | 1460 |

167 rows × 10 columns

*Figure 3: Data View of columns and values*

After acquiring the dataset, it is crucial to perform data exploration to gain insights into the data's characteristics. This includes checking for standard statistical measures such as mean, median, minimum, maximum, and quartiles for each feature. Understanding the statistical distribution of the data can help identify potential issues like outliers or skewed distributions.

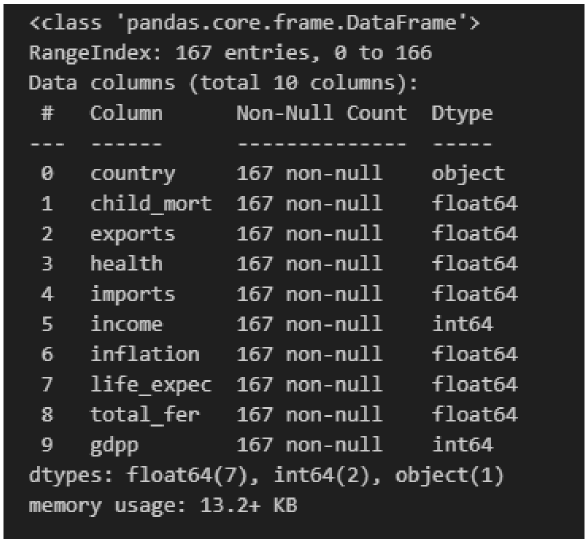


| | child_mort | exports | health | imports | income | inflation | life_expec | total_fer | gdpp |
|---|---|---|---|---|---|---|---|---|---|
| count | 167.000000 | 167.000000 | 167.000000 | 167.000000 | 167.000000 | 167.000000 | 167.000000 | 167.000000 | 167.000000 |
| mean | 38.270060 | 41.108976 | 6.815689 | 46.890215 | 17144.688623 | 7.781832 | 70.555689 | 2.947964 | 12964.155689 |
| std | 40.328931 | 27.412010 | 2.746837 | 24.209589 | 19278.067698 | 10.570704 | 8.893172 | 1.513848 | 18328.704809 |
| min | 2.600000 | 0.109000 | 1.810000 | 0.065900 | 609.000000 | -4.210000 | 32.100000 | 1.150000 | 231.000000 |
| 25% | 8.250000 | 23.800000 | 4.920000 | 30.200000 | 3355.000000 | 1.810000 | 65.300000 | 1.795000 | 1330.000000 |
| 50% | 19.300000 | 35.000000 | 6.320000 | 43.300000 | 9960.000000 | 5.390000 | 73.100000 | 2.410000 | 4660.000000 |
| 75% | 62.100000 | 51.350000 | 8.600000 | 58.750000 | 22800.000000 | 10.750000 | 76.800000 | 3.880000 | 14050.000000 |
| max | 208.000000 | 200.000000 | 17.900000 | 174.000000 | 125000.000000 | 104.000000 | 82.800000 | 7.490000 | 105000.000000 |

*Figure 4: Statistical Distribution of The Data*

Next, it is important to check the data types for each column to ensure they are correctly interpreted. For example, numerical features should be represented as numeric data types, while categorical variables should be appropriately encoded.

Data cleaning involves checking for missing values and duplicate entries. It is essential to ensure that the dataset is complete and free from any data quality issues that could affect the modeling process. Duplicate and incomplete data items are deleted from the dataset.



```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 167 entries, 0 to 166
Data columns (total 10 columns):
 #   Column      Non-Null Count  Dtype
---  ------      --------------  -----
 0   country     167 non-null    object
 1   child_mort  167 non-null    float64
 2   exports     167 non-null    float64
 3   health      167 non-null    float64
 4   imports     167 non-null    float64
 5   income      167 non-null    int64
 6   inflation   167 non-null    float64
 7   life_expec  167 non-null    float64
 8   total_fer   167 non-null    float64
 9   gdpp        167 non-null    int64
dtypes: float64(7), int64(2), object(1)
memory usage: 13.2+ KB
```

*Figure 5: Summary of a DataFrame*

### B. Data Preparation and Processing

After acquiring the dataset, data preparation and processing steps are necessary to ensure the data's quality and suitability for modeling. The previously mentioned data analysis and preprocessing techniques can be applied, including handling missing values, dropping irrelevant columns, and normalizing the data using techniques.

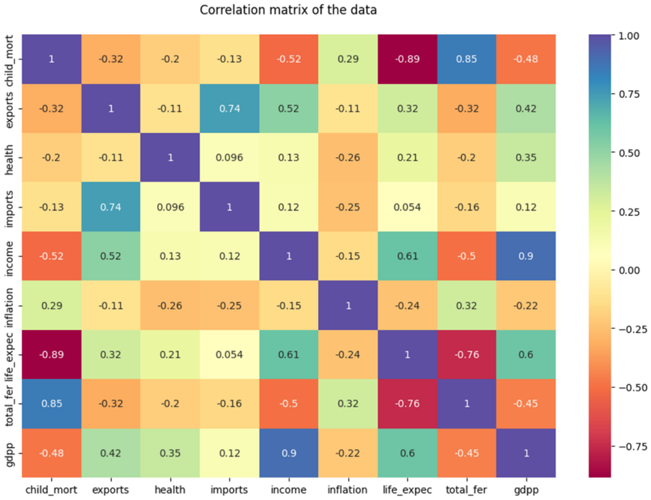


*Figure 6: Correlation Matrix of The Data*

Figure 6 illustrates the correlation matrix of the dataset, providing valuable insights into the relationships between different variables. The correlation matrix helps us understand the extent to which variables are related to each other, which is essential for identifying patterns and dependencies within the data.

These correlation insights provide a deeper understanding of the dataset. For example, countries with higher child mortality rates often exhibit lower life expectancy and higher total fertility rates. Additionally, higher income levels are generally associated with lower child mortality rates, higher life expectancy, and higher GDP per capita.

Understanding these relationships is valuable for further analysis and decision-making. The correlation matrix assists in identifying potential indicators or factors that influence various socio-economic aspects of a country.

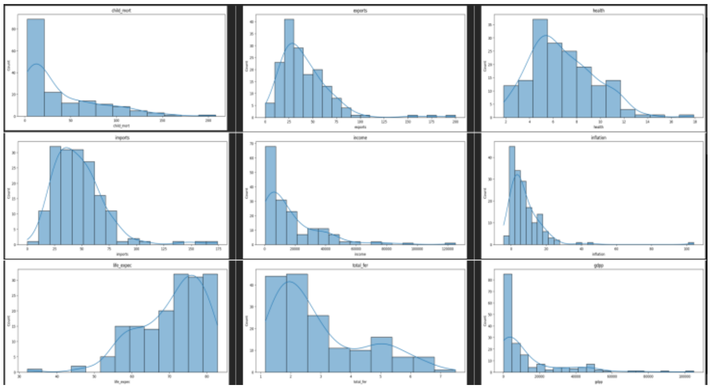


*Figure 7: Data Distribution of each column*

Figure 7 illustrates the Data Distribution of each column. Firstly, the health attribute follows a normal distribution. This means that the majority of countries have similar levels of health spending per capita, with no significant skewness towards higher or lower values.

Secondly, the life expectancy attribute exhibits a left-skewed distribution. This indicates that most countries have higher life expectancies, while a few countries have lower life expectancies. The skewness towards the left suggests that there is a concentration of countries with longer life expectancies.

Lastly, the remaining attributes, including child mortality, exports, imports, income, inflation, total fertility, and GDP per capita, display right-skewed distributions. This means that the majority of countries tend to have lower values for these indicators, while a few countries have higher values. The positive skewness towards the right indicates that there is a tail towards higher values, suggesting the presence of countries with more extreme values in these socio-economic indicators.

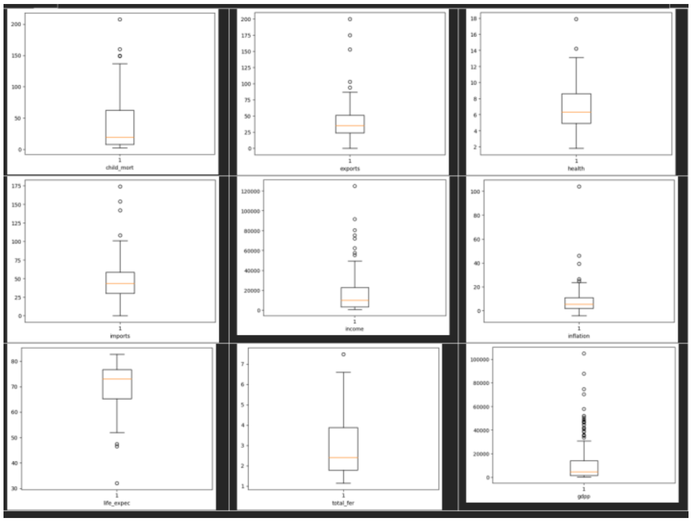


*Figure 8: Detect outliers using box plots*

The presence of outliers in a dataset can have a significant impact on data analysis and modeling. Outliers are data points that deviate significantly from the overall pattern of the data. They can arise due to various reasons such as measurement errors, data entry mistakes, or representing extreme cases within the population being studied.

However, it is important to note that in this analysis, the decision has been made not to remove outliers. This is because outliers can be valuable and informative in understanding the countries that are in critical condition and need of help. Outliers may represent countries facing extreme socio-economic challenges or experiencing unique circumstances that warrant special attention.

### C. FCM Modelling

The Fuzzy C-Means (FCM) algorithm is an unsupervised clustering algorithm that aims to assign data points to clusters based on their proximity to the cluster centers. In this section, the FCM modeling phase is implemented to identify clusters within the dataset. Before applying the FCM algorithm, it is essential to ensure that the columns of the dataset are normalized. Normalization is an essential preprocessing step that ensures all features in the dataset are on a similar scale. The StandardScaler technique is employed to normalize the data because StandardScaler is less sensitive to outliers compared to other methods. After normalization, it is necessary to convert the normalized data back into a structured format for ease of interpretation and further analysis.

Applying the FCM algorithm:

1. Setting the number of clusters:

Before applying the FCM algorithm, sets the desired number of clusters. In this case, the number of clusters is set to 3.

2. Fitting the preprocessed data:

The preprocessed data, which has been normalized using the StandardScaler, is then fitted to the FCM algorithm. The fit() method is called on the FCM model, passing in the preprocessed data as the input.

3. Obtaining the cluster centers:

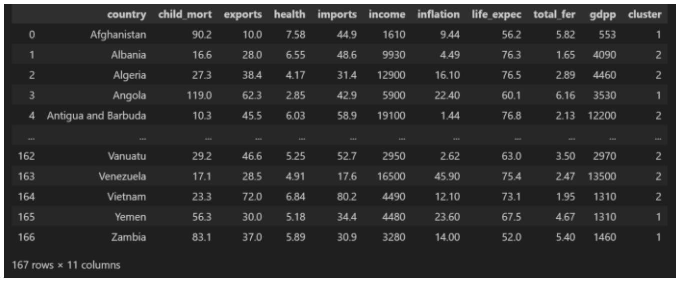
After fitting the data, retrieves the cluster centers using the centers attribute of the FCM model. These cluster centers represent the representative points for each cluster and provide insights into the characteristics of the clusters.

4. Predicting the cluster for each data point:

Using the fitted FCM model predicts the cluster membership for each data point in the preprocessed data. This is done using the predict() method, which assigns each data point to the cluster with the highest membership degree based on their proximity to the cluster centers.

## V. THE RESULTS AND DISCUSSION

The DataFrame below shows the information on countries with clustered values.



| | country | child_mort | exports | health | imports | income | inflation | life_expec | total_fer | gdpp | cluster |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | Afghanistan | 90.2 | 10.0 | 7.58 | 44.9 | 1610 | 9.44 | 56.2 | 5.82 | 553 | 1 |
| 1 | Albania | 16.6 | 28.0 | 6.55 | 48.6 | 9930 | 4.49 | 76.3 | 1.65 | 4090 | 2 |
| 2 | Algeria | 27.3 | 38.4 | 4.17 | 31.4 | 12900 | 16.10 | 76.5 | 2.89 | 4460 | 2 |
| 3 | Angola | 119.0 | 62.3 | 2.85 | 42.9 | 5900 | 22.40 | 60.1 | 6.16 | 3530 | 1 |
| 4 | Antigua and Barbuda | 10.3 | 45.5 | 6.03 | 58.9 | 19100 | 1.44 | 76.8 | 2.13 | 12200 | 2 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 162 | Vanuatu | 29.2 | 46.6 | 5.25 | 52.7 | 2950 | 2.62 | 63.0 | 3.50 | 2970 | 2 |
| 163 | Venezuela | 17.1 | 28.5 | 4.91 | 17.6 | 16500 | 45.90 | 75.4 | 2.47 | 13500 | 2 |
| 164 | Vietnam | 23.3 | 72.0 | 6.84 | 80.2 | 4490 | 12.10 | 73.1 | 1.95 | 1310 | 2 |
| 165 | Yemen | 56.3 | 30.0 | 5.18 | 34.4 | 4480 | 23.60 | 67.5 | 4.67 | 1310 | 1 |
| 166 | Zambia | 83.1 | 37.0 | 5.89 | 30.9 | 3280 | 14.00 | 52.0 | 5.40 | 1460 | 1 |

167 rows × 11 columns

*Figure 9: Data View of columns and values (Clustered)*

### A. Result Analytics and Identifying Countries in Need Aid

Once the Fuzzy C-Means (FCM) clustering algorithm has been applied to the normalized dataset, we can proceed with result analytics and identifying countries in need of aid. This stage involves analyzing the clustered data and determining which countries require assistance based on specific socio-economic indicators.

Visualization plays a crucial role in understanding the clustering results. The clusters are visualized using scatter plots. The dimensionality reduction technique used transforms the original high-dimensional data into a lower-dimensional space while preserving the most important patterns and variations in the data. By plotting the countries' data points in this reduced two-dimensional space, we can observe the clustering behavior and identify patterns among countries based on their socio-economic indicators. Each data point represents a country and its position on the scatter plot. Each data point is plotted with a specific color representing its assigned cluster. Additionally, the cluster centers are plotted as black crosses. This visualization allows us to visually inspect the clusters and their distribution within the feature space.

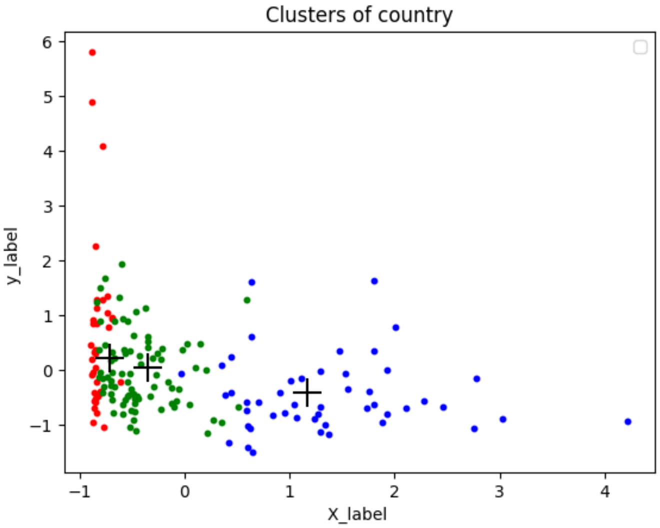
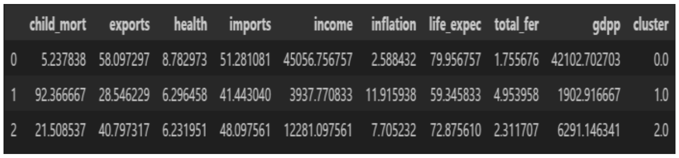


*Figure 10: Data Visualizing the Clusters*

After visualizing the clusters, proceeds to identify the countries in need of aid. To accomplish this, the cluster means for each attribute are calculated by grouping the data based on the predicted clusters. By analyzing the cluster means, we can identify clusters with lower values for indicators such as child mortality, income, GDP per capita, and life expectancy. These indicators are commonly associated with countries in need of aid, as they reflect socio-economic challenges.



| | child_mort | exports | health | imports | income | inflation | life_expec | total_fer | gdpp | cluster |
|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 5.237838 | 58.097297 | 8.782973 | 51.281081 | 45056.756757 | 2.588432 | 79.956757 | 1.755676 | 42102.702703 | 0.0 |
| 1 | 92.366667 | 28.546229 | 6.296458 | 41.443040 | 3937.770833 | 11.915938 | 59.345833 | 4.953958 | 1902.916667 | 1.0 |
| 2 | 21.508537 | 40.797317 | 6.231951 | 48.097561 | 12281.097561 | 7.705232 | 72.875610 | 2.311707 | 6291.146341 | 2.0 |

*Figure 11: Cluster Mean*

Figure 11 shows the Cluster Mean:

Cluster 0: This cluster represents countries with strong and positive socio-economic indicators. These countries exhibit good economic development, high life expectancy, and low child mortality rates. It indicates that these nations have well-established healthcare systems, higher incomes, and overall better living conditions.

Cluster 1: This cluster is characterized by the most negative values across the indicators. Countries in this cluster experience high child mortality rates, the lowest levels of economic development, low GDP per capita, minimal exports and imports, and the lowest life expectancy. These countries face significant socio-economic challenges and require immediate attention and aid.

Cluster 2: This cluster shows average values for all the features when compared to the other clusters. The countries in this cluster demonstrate moderate levels of development across socio-economic indicators. While they may not require immediate aid, targeted interventions, and support can contribute to their further progress.

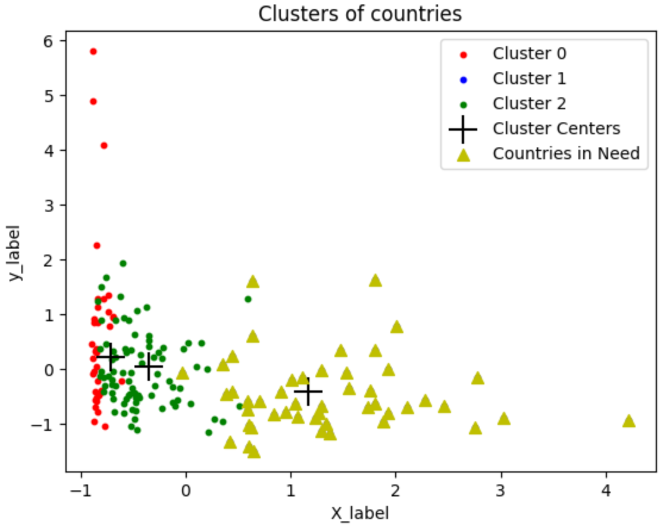
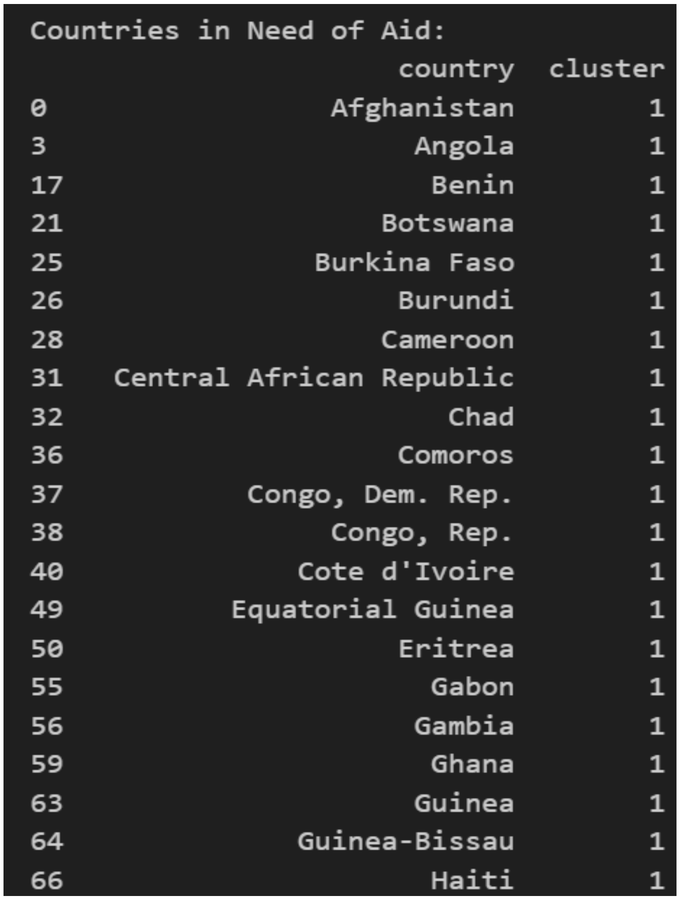


*Figure 12: Data Visualizing the clusters (labeled)*

Visualizes the clusters and the countries in need of aid in a scatter plot. The data points belonging to each cluster are plotted with different colors, while the cluster centers are marked as black crosses. This visualization allows us to visualize the clusters of countries and identify those in need of aid.

Lastly, displays the list of countries in need of aid by printing it to the console. This provides a clear understanding of which countries require assistance based on the clustering analysis. The list can be further utilized for targeted intervention and aid programs, focusing on countries facing socio-economic challenges.



```
Countries in Need of Aid:
                     country  cluster
0                Afghanistan        1
3                     Angola        1
17                     Benin        1
21                  Botswana        1
25              Burkina Faso        1
26                   Burundi        1
28                  Cameroon        1
31  Central African Republic        1
32                      Chad        1
36                   Comoros        1
37          Congo, Dem. Rep.        1
38               Congo, Rep.        1
40             Cote d'Ivoire        1
49         Equatorial Guinea        1
50                   Eritrea        1
55                     Gabon        1
56                    Gambia        1
59                     Ghana        1
63                    Guinea        1
64             Guinea-Bissau        1
66                     Haiti        1
```

*Figure 13: List of countries in need of aid*

In conclusion, the resulting analytics and identification of countries in need of aid involve visualizing the clusters, calculating cluster means, identifying clusters with lower indicator values, and finally, determining the countries in need based on these clusters. The combination of visualization and analysis enables us to gain valuable insights and make informed decisions regarding aid and support to countries experiencing socio-economic difficulties.

## CONCLUSION

As a conclusion, we conducted an analysis to identify countries in need of aid based on socio-economic indicators. We started by acquiring a dataset and performing exploratory data analysis, which helped us understand the data distribution and identify correlations among variables. Outliers were retained as they provided valuable insights into countries facing critical conditions. We then applied the Fuzzy C-Means clustering algorithm to classify countries into distinct clusters. Cluster analysis revealed three clusters: one representing positive values across indicators, another indicating country in dire need of aid, and a third representing countries with average values. Based on the cluster means, we identified 48 countries that require urgent aid. These countries exhibit characteristics such as high child mortality, low economic development, and low life expectancy. The results emphasize the need for targeted support and collaborative efforts to improve the well-being and quality of life in these nations. By prioritizing aid and sustainable development, we can work towards a more equitable and prosperous world for all.

## REFERENCES

[1] Jain, V., & Verma, A. (2019). Cluster Analysis for Social and Economic Development. International Journal of Research in Computer Science, 9(1), 1-5.

[2] Hair Jr, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). Multivariate Data Analysis (8th ed.). Cengage Learning.

[3] Kundu, S., & Mondal, S. (2020). Identifying Developing Countries for Aid Allocation: A Fuzzy Clustering Approach. International Journal of Economics, Commerce, and Management, 8(11), 103-117.

[4] 10. Susana, N. (2005). Fuzzy Clustering via Proportional Membership Model. IOS Press.

[5] Cornelius, T.L. (1998). Fuzzy Logic and Expert Systems Applications. Elsevier Science Publishing Co Inc.

[6] Marie, B. (2001). An Introduction To Many-Valued And Fuzzy Logic. Cambridge University Press.

[7] Fuzzy Logic with Engineering Applications, 3rd Edition Timothy J. Ross

[8] Social determinants of health (who.int) https://www.who.int/health-topics/social-determinants-of-health#tab=tab_1

[9] Unsupervised Learning on Country Data | Kaggle https://www.kaggle.com/datasets/rohan0301/unsupervised-learning-on-country-data?select=Country-data.csv

[10] fuzzy systems optimization identification based on FCM https://asp-eurasipjournals.springeropen.com/articles/10.1186/s13634-020-00706-2