Strategic Aid Allocation: A Data-Driven Approach to Maximize Impact

P Parth Patel
4 min read · Just now

Introduction

Humanitarian organizations often face the challenge of strategically allocating limited resources to achieve maximum impact. With $10 million raised, our goal was to identify the most vulnerable countries requiring aid and optimize fund allocation for poverty alleviation and community development. Using advanced data analytics, this project leverages clustering algorithms, exploratory data analysis (EDA), and predictive modeling to provide actionable insights.

This blog details the journey of solving the problem, from data preprocessing and EDA to hypothesis testing, modeling, and deployment,

alongside key insights and recommendations.

Problem Statement

A humanitarian NGO must strategically allocate $10 million to maximize the impact of aid. The challenge is to:

1. Identify countries in dire need based on socio-economic and health indicators.
2. Group countries into clusters with similar needs to streamline resource allocation.
3. Provide actionable recommendations to improve outcomes in critical areas like health, education, and infrastructure.

Dataset Overview

The dataset used includes the following key features:

- **Country**: Name of the country.
- **Child Mortality**: Deaths of children under 5 years per 1000 live births.
- **Exports**: Exports as a percentage of GDP.
- **Health**: Health spending as a percentage of GDP.
- **Imports**: Imports as a percentage of GDP.
- **Income**: Net income per person.

- **Inflation**: Annual inflation rate.
- **Life Expectancy**: Average number of years a newborn is expected to live.
- **Fertility Rate**: Average number of children born per woman.
- **GDP per Capita**: GDP per person.

Step 1: Exploratory Data Analysis (EDA)

Univariate Analysis

Key Observations:

- Countries with high child mortality rates often have low income and life expectancy.
- A large disparity exists in GDP per capita among countries.

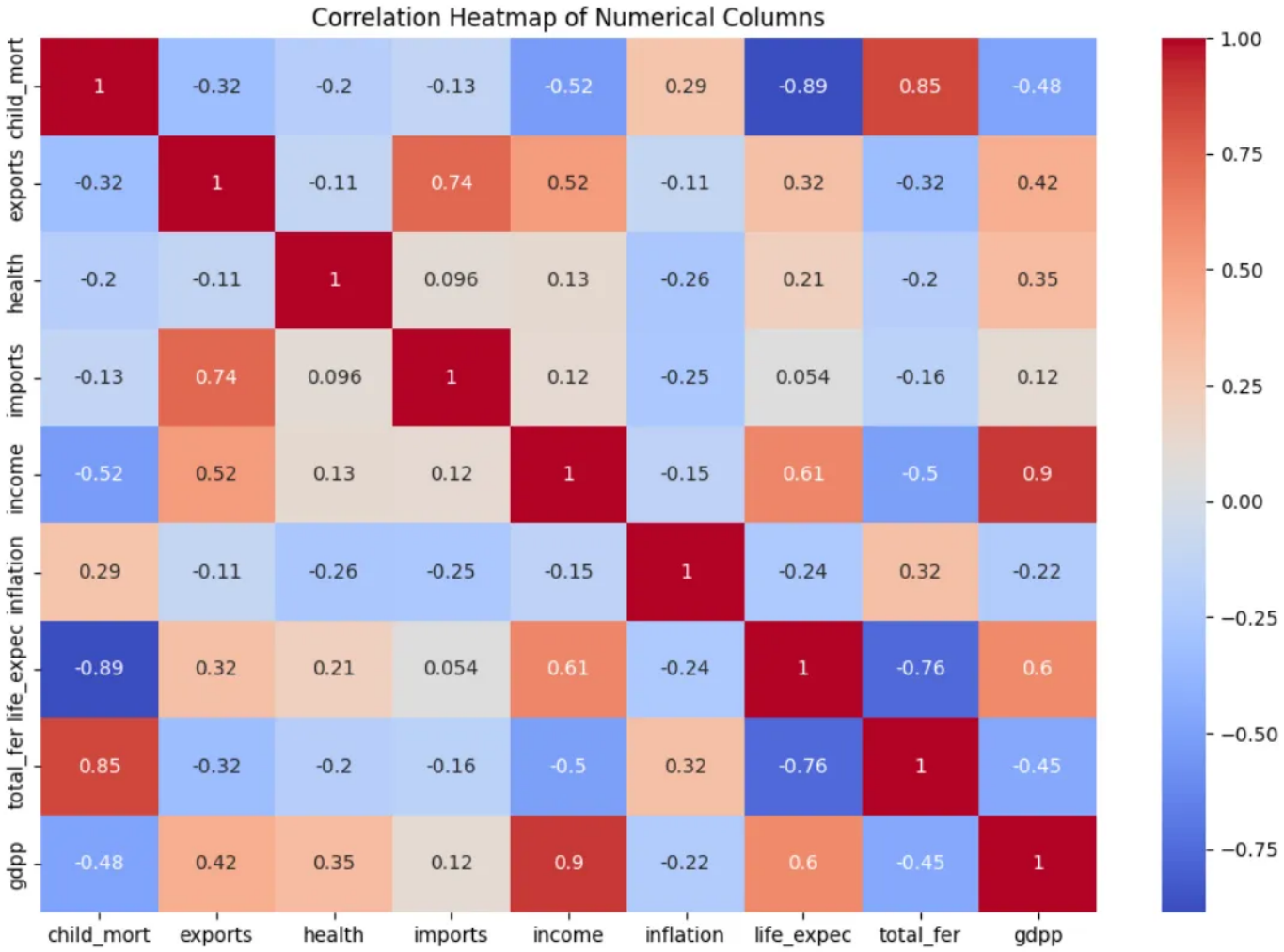
Insights:

- Child mortality rates ranged from 2.6 to 208 per 1000 live births.
- Health spending was low in countries with high child mortality.

Visualization:

- Distribution plots for GDP per capita, child mortality, and health spending revealed skewness. Log transformations were applied to

normalize the data.



Heatmap illustrating key performance indicators across various countries, highlighting areas requiring focused aid allocation

Bivariate Analysis

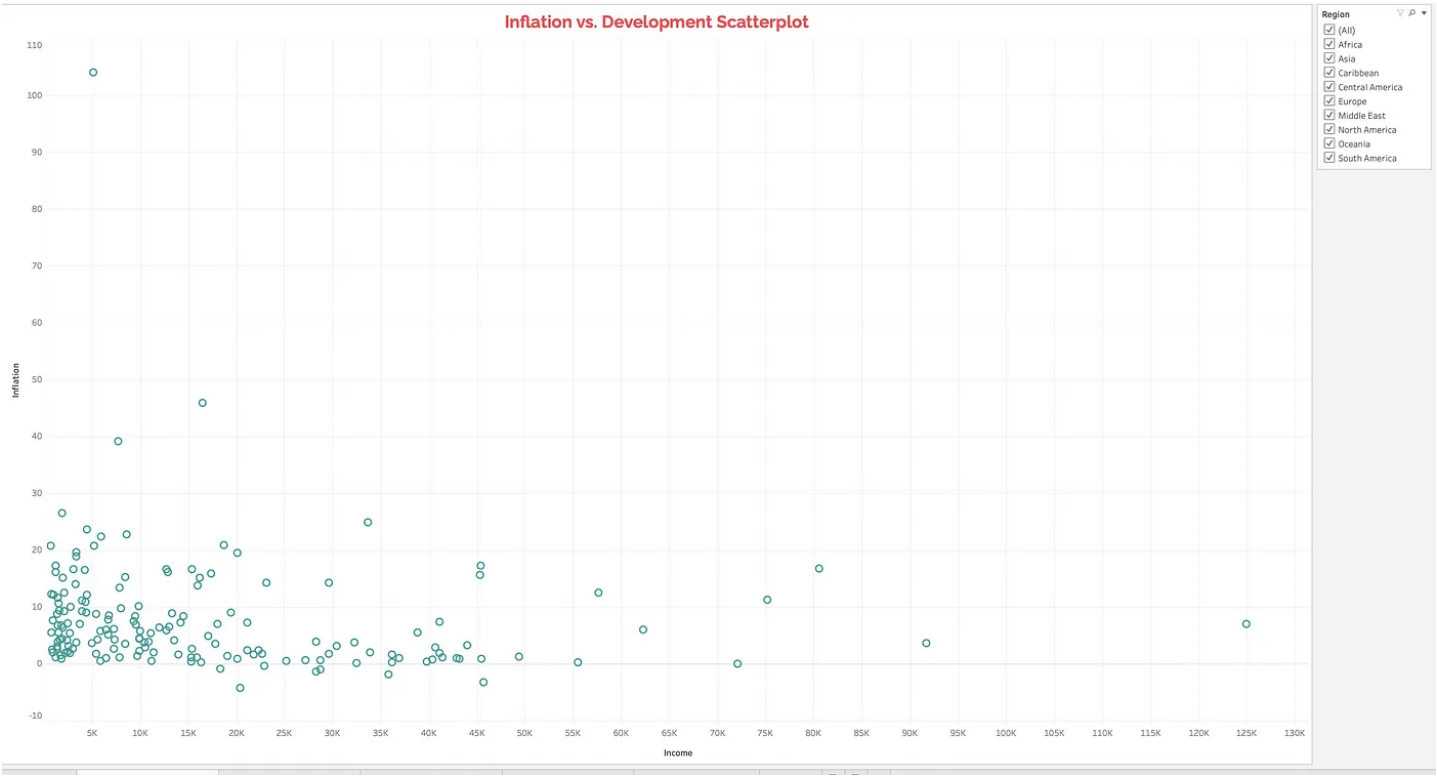
Key Relationships:

- A strong negative correlation (ρ = -0.63) exists between child mortality and income.

- Positive correlation ($\rho = 0.25$) between health spending and life expectancy.

Visualization:

- Correlation heatmaps highlighted critical relationships between features.



Scatterplot illustrating the relationship between inflation rates and development indicators across various countries.

Outlier Detection and Treatment

Techniques:

- Winsorization was applied to cap outliers in child mortality, exports, and imports.

Step 2: Hypothesis Testing

Hypothesis 1: Increased health spending improves life expectancy.

- **T-test Results:**
- **t-statistic:** -97.99
- **p-value:** 4.22 × 10^-247
- **Conclusion:** Reject the null hypothesis. Increased health spending significantly correlates with higher life expectancy.

Hypothesis 2: Higher income reduces child mortality.

- **T-test Results:**
- **t-statistic:** -13.68
- **p-value:** 4.27 × 10^-34
- **Conclusion:** Higher income is strongly associated with lower child mortality rates.

Step 3: Feature Engineering

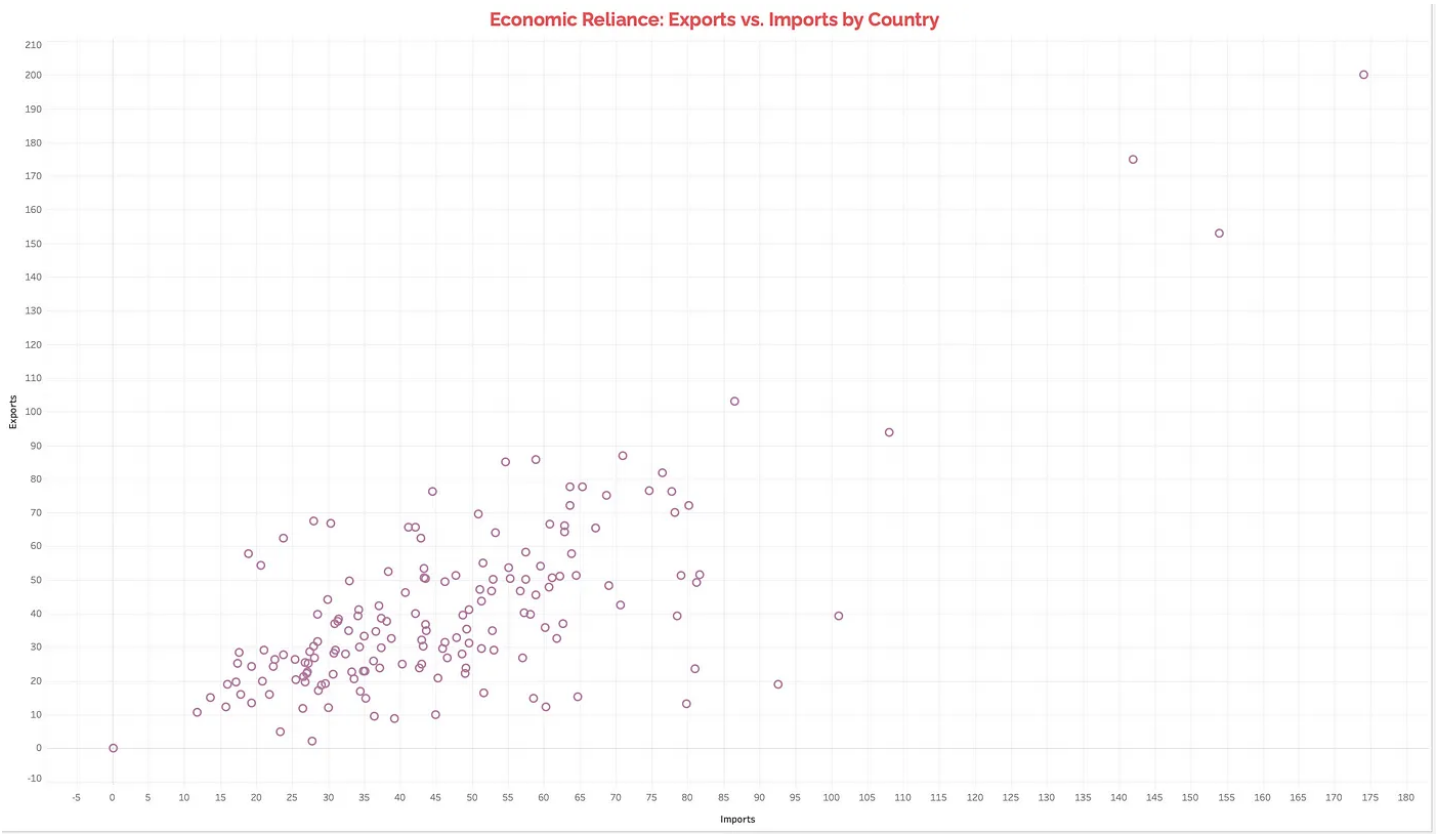
New Features Created:

- **Import-Export Ratio**: Exports/Imports to understand economic reliance.
- **Log GDP**: Applied log transformation to normalize GDP per capita.
- **Regional Encoding**: Grouped countries into regions (e.g., South Asia, Africa) for better clustering.

Scaling:

- StandardScaler was applied to normalize numerical features.

Step 4: Clustering and Modeling



Clustering Techniques

K-Means Clustering:

- Optimal number of clusters: 3 (determined using the Elbow Method).
- Clusters represented distinct socio-economic profiles.

Hierarchical Clustering:

- Visualized through dendrograms.
- Confirmed 3-cluster structure.

DBSCAN:

- Identified noise points and outliers but showed less-defined clusters compared to K-Means.

PCA for Dimensionality Reduction:

- Reduced features to two principal components for visualization.

Cluster Insights

- **Cluster 0**: High child mortality, low income, and low GDP.
- **Cluster 1**: Moderate child mortality and income levels.

- **Cluster 2:** Low child mortality, high income, and high GDP.

Step 5: Deployment

Model Deployment

K-Means Model Pickle:

- Exported using Pickle for integration into dashboards.

Interactive Dashboard:

- Built with Python and hosted on Google Colab.
- Visualized country clusters, correlations, and key metrics.

Deployment Link: *Colab Notebook*

GitHub Repository: *GitHub Repo*

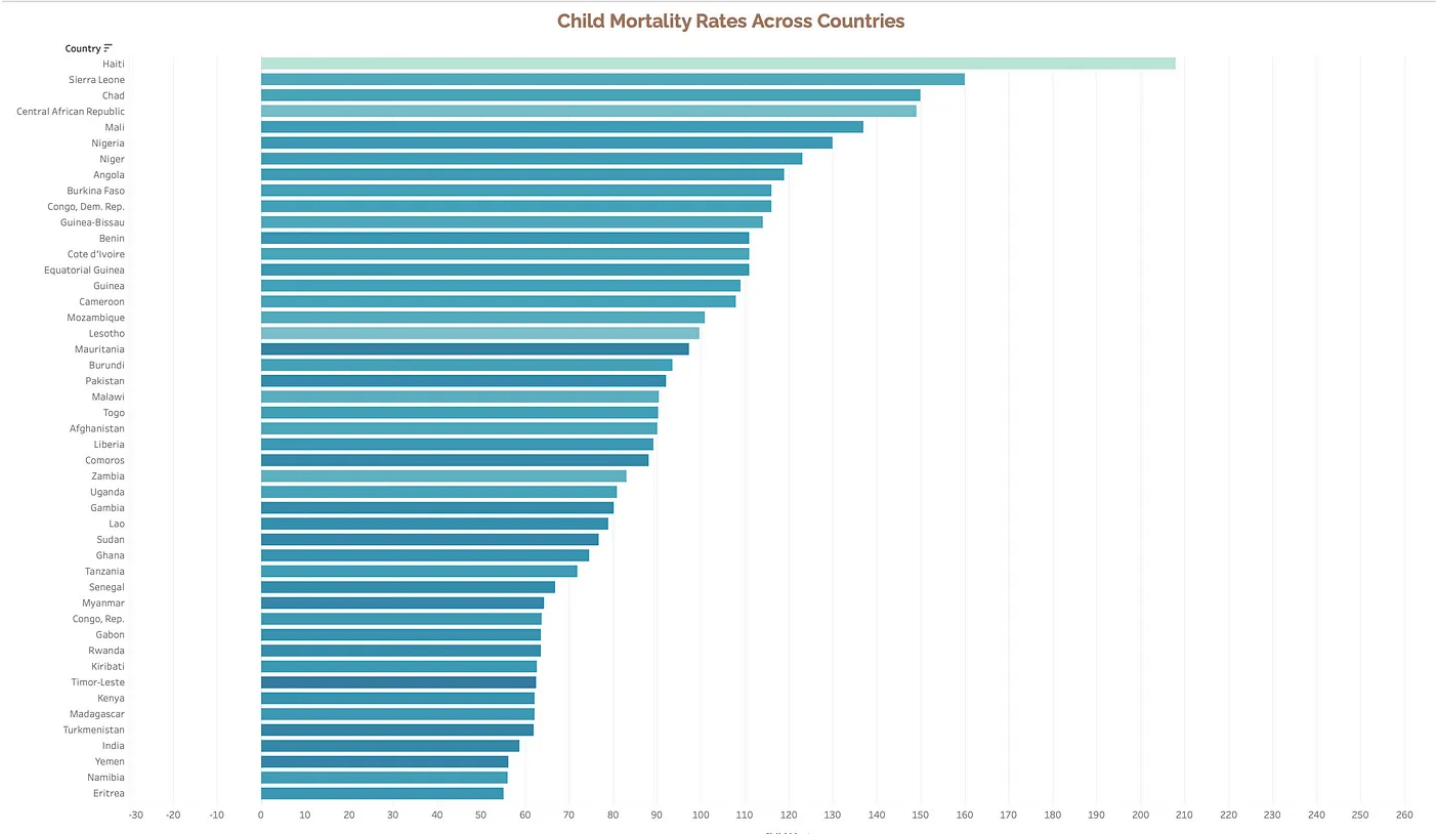
Tableau Dashboard: *Tableau Dashboard*

Key Insights and Recommendations

Invest in Healthcare:

- Fund immunization programs and enhance maternal healthcare to

reduce child mortality.



Strengthen Local Health Systems:

- Collaborate with local organizations to supply medical resources and train healthcare workers.

Address Economic Disparities:

Invest in education and vocational training to improve income levels and reduce poverty

Promote Food Security:

- Support sustainable agriculture and nutritional education to address malnutrition.

Long-Term Development Projects:

- Focus on clean water access, renewable energy, and education to address root causes of poverty.

Conclusion

This project demonstrates how data analytics can guide impactful decision-making in resource allocation. By clustering countries based on socio-economic and health metrics, we provided actionable recommendations to maximize aid impact. Future work can incorporate real-time data to refine insights and adapt strategies dynamically.

P

Written by Parth Patel

0 Followers · 4 Following

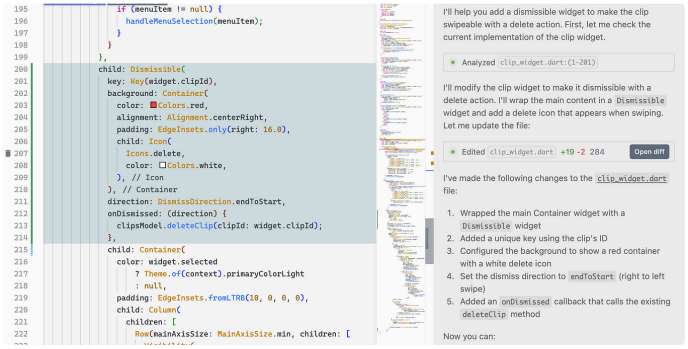
Edit profile

No responses yet

What are your thoughts?

Respond

Recommended from Medium



Austin Starks

You are an absolute moron for believing in the hype of “AI...

All of my articles are 100% free to read. Non-members can read for free by clicking my...

Jan 11 5.4K 223

In Coding Beauty by Tari Ibaba

This new IDE just destroyed VS Code and Copilot without even...

Wow I never thought the day I stop using VS Code would come so soon...

Jan 17 1.6K 64

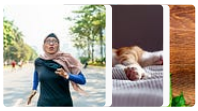
Lists

Staff picks

804 stories · 1588 saves

Stories to Help You Level-Up at Work

19 stories · 927 saves

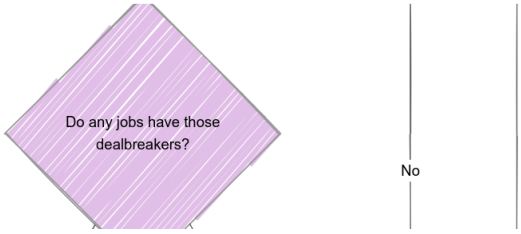


Self-Improvement 101

20 stories · 3244 saves

Productivity 101

20 stories · 2742 saves



In Python in Plain English by Kiran Maan

Just Stop Writing Python Functions Like This!!!

I just reviewed someone else's code and I was just shocked.

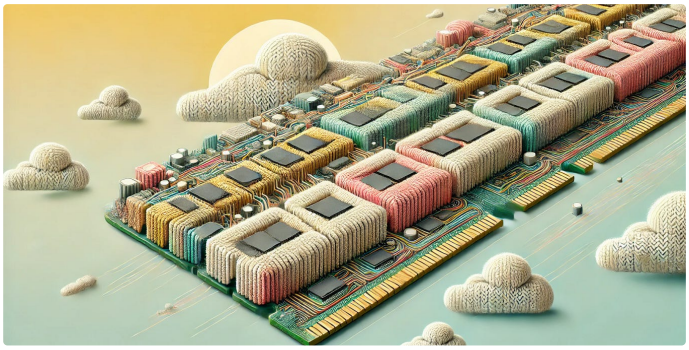
Jan 19 1.4K 36

Mark Shrime, MD, PhD

The Anatomy of a Good Decision

Make the hardest decisions of your life like a surgeon

Dec 16, 2024 3K 36



In Level Up Coding by Dr. Ashish Bamania

Memory Layers Are Supercharging LLMs Like Never...

A deep dive into how Memory layers work and how they supercharge the performanc...

5d ago 849 21

Alberto Romero

DeepSeek Is Chinese But Its AI Models Are From Another Planet

OpenAI and the US are in deep trouble

4d ago 1.7K 31

See more recommendations