Tackling Malnutrition in Africa: A Comprehensive Data Analysis

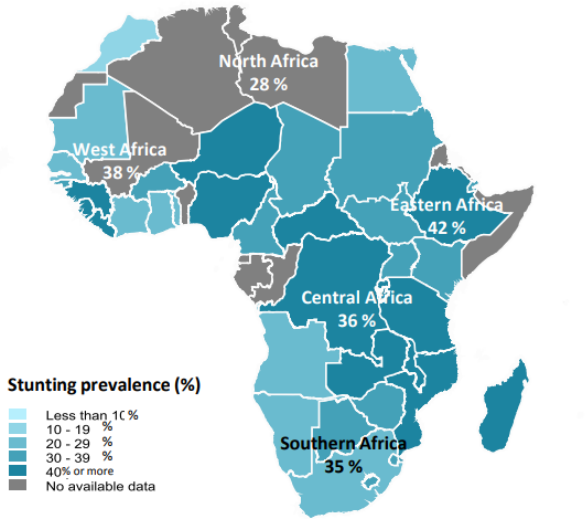
Nithin V.K. Rachakonda, Nithisha S. Marripelly, Koushik S. Kadiyala

Abstract

Malnutrition has been an issue in Africa for centuries. In this paper using the data from the world bank databases, we test a few factors that affect the prevalence of Malnutrition in African countries. Pandas and Matplotlib libraries from pandas have been used to clean and visualize the extracted data. Pearson correlation coefficient, Spearman correlation coefficient and p-value are used to confirm the relation statistically. This paper can be used to draft policies by respective governments or policy makers.

Introduction

Malnutrition is a significant issue in Africa, affecting the lives of millions of individuals and hindering the continent's overall growth. According to UN data, at least 1 in 3 children under 5 years of age in Africa were stunted in 2011. It is indeed ironic that the continent abundant in natural resources is facing this issue.



Percentage of children under 5 years whose height-for-age falls below -2 standard deviations.

This problem not only affects the population's immediate health, but it also gets in the way of broader economic development and social stability. Understanding the complexity and in-depth consequences of malnutrition in Africa is critical to developing effective strategies to address this persistent problem and revitalize communities across the continent. This paper aims to dive into depths of this issue, analyze the data available from world bank databases, and evaluate four proposed hypotheses to alleviate the problem.

Hypotheses

- The proportion of the population who cannot afford a healthy diet is negatively correlated with GDP per capita of their respective country.
- Higher levels of Malnutrition are associated with lower maternal education levels.
- The prevalence of malnutrition is positively correlated with food prices.
- Higher immunization rates are associated with lower malnutrition rates.

Methodology

Data Extraction: The data has been extracted from the world data indicators and Africa data indicators, both of which are databases of the world bank that can be accessed from their website. The data has been downloaded as .csv files.

Data Cleaning: Data has been cleaned using Pandas library from python. The downloaded .csv files have been cleaned and stored in new .csv files for further steps. In the process of cleaning following assumptions were made:

- Malnutrition Index is calculated by taking the mean of stunting, wasting and underweight data.
- The data is of children under 5 years of age and we assume that this data reflects the malnutrition rates of the entire population.

Statistical Analysis: The cleaned data has been put into statistical analysis to check if there exists a correlation between variables that have been listed out in hypotheses. Pearson correlation co-efficient and Spearman correlation co-efficient and p-value were found to draw a conclusion on the relation between the variables.

The correlation coefficients possess values from -1 to 1, -1 being a perfect negative relation and 1 being the perfect positive relation and 0 concluding no strong relation. Pearson correlation coefficient tells us about a linear relation whereas a spearman correlation coefficient tells us about a monotonic relation.

The p-value is a statistical measure that determines the strength of the evidence against a null hypothesis (hypothesis that suggests that there is no effect or no relationship between variables). A small p-value, less than 0.05 suggests convincing evidence against null hypothesis i.e., strong support towards the existence of a relationship between the variables.

Data Visualization: After undergoing through statistical analysis, the variables which will be proven statistically to have a relation between them will be visualized using Matplotlib library of python. Scatter plots will be plotted of the African countries over the data variables to visualize the relation between them.

Note – From data cleaning to data visualization, all the study has been done on python.

Results and Discussion

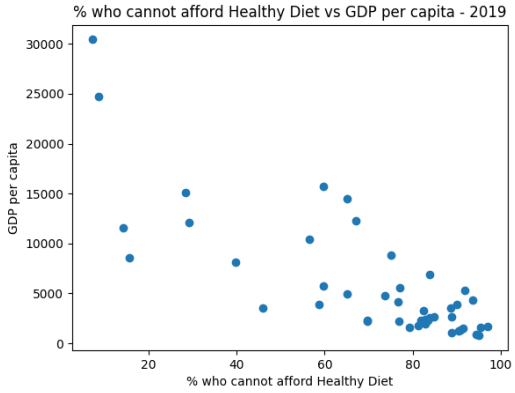
(i) Food affordability vs GDP percapita

The proportion of the population that cannot afford a healthy diet has been considered as the Affordability of food and GDP of their respective country is considered as GDP. Obtained results are as below:

Spearman correlation coefficient: -0.74
Pearson correlation coefficient: -0.79
P-value: 0.0000000045

The consistency between the Spearman and Pearson correlations reinforces the conclusion that as % who cannot afford healthy diet increases, GDP per capita tends to decrease, and vice versa. The extremely low p-value suggests that the observed correlations are highly statistically significant. In summary, there is a highly significant and strong negative relationship between the two variables.

The plot agrees with the statistical views on the data as we can clearly see that there is a negative relation between % who cannot afford healthy diet and GDP per capita.



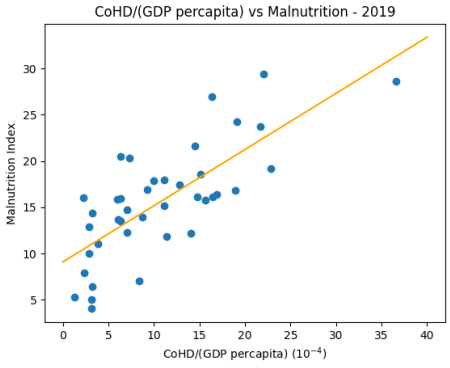
Implementing or enhancing income support programs to assist individuals or households with lower income levels, ensures they can also afford a healthy diet despite potential economic disparities.

(ii) Food prices vs Malnutrition

Cost of a healthy diet per person per GDP (CoHD_GDP per capita) of a country is considered as affordability of food by an average person of a country. On the other hand malnutrition index is considered. Obtained results are as below:

Spearman correlation coefficient: 0.73
Pearson correlation coefficient: 0.75
P-value: 0.00000012

The consistency between the Spearman and Pearson correlations reinforces the conclusion that as CoHD_GDP per capita increases, malnutrition also tends to increase. The extremely low p-value suggests that the observed correlations are highly statistically significant.



The plotted data corroborates the statistical perspective, displaying a distinct positive correlation between CoHD_GDP per capita and the malnutrition index.

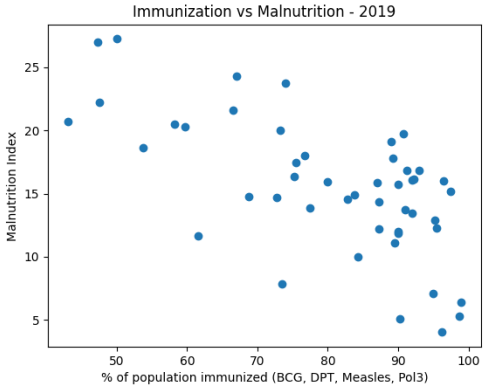
We can improve market access for farmers by enhancing transportation infrastructure and establishing fair trade practices. This can contribute to more competitive and stable food prices. We can try to implement policies aimed at stabilizing food prices, such as strategic food reserves, buffer stocks, and price ceilings on essential food items.

(iii) Immunization rates vs Malnutrition

Mean of proportion of Immunized population for BCG, DPT, Pol and Measles is considered for Immunization rate for a country. The other variable is Malnutrition index we have derived. Obtained results are as below:

Spearman correlation coefficient: -0.59
Pearson correlation coefficient: -0.65
P-value: 0.00001

The consistency between the Spearman and Pearson correlations further supports the finding that higher immunization rates correspond to lower instances of malnutrition. The remarkably low p-value strongly indicates the high statistical significance of these observed correlations.



The visual representation aligns with the statistical analysis, illustrating a clear negative relationship between immunization rates and the malnutrition index in the data.

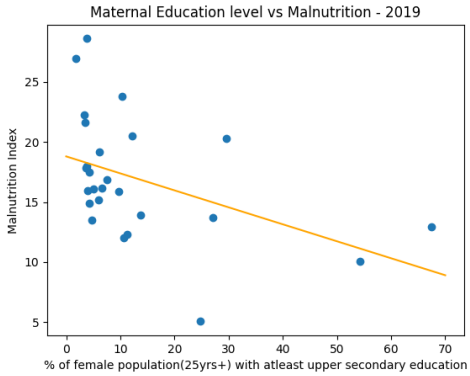
Identifying and implementing targeted interventions in areas with high malnutrition rates. These interventions may include mobile clinics, door-to-door campaigns, and community-based programs that can result in better immunization rates.

(iv) Maternal Education vs Malnutrition

In the world bank database, there were many education indicators available that can be proxied for maternal education rate. Among those the best proxy for maternal education levels has been decided to be educational attainment up to at upper secondary, in female population (25yrs+). This is because it is a more advanced level of education than primary education and lower secondary education where nutritional science would not be taught, but not as advanced as post-secondary education or bachelor's degree. This data has been related to the malnutrition index and the obtained results are as below:

Spearman correlation coefficient: -0.59
Pearson correlation coefficient: -0.44
P-value: 0.0014

There is evidence to support a negative correlation between the two variables. The Spearman correlation, being -0.59, suggests a stronger monotonic relationship compared to the Pearson correlation, which is -0.44. Low p-value indicates that the observed correlations are statistically significant.



The graph agrees with the statistical analysis, clearly illustrating an evident negative association between proportion of female population (25yrs+) with at least secondary education and the malnutrition index in the dataset.

Strengthening antenatal and postnatal care services to ensure that mothers receive proper nutritional guidance during pregnancy and after childbirth, including support for breastfeeding and complementary feeding practices.

Conclusion and Future Work

We stated four hypotheses that are related to malnutrition in Africa. We have extracted data from the world bank databases and cleaned it according to our needs. We subsequently visualized the proposed hypotheses, supported by statistical analysis, indicating that the selected variables either directly or indirectly contribute to or increase the issue

of malnutrition in Africa. We later discussed possible solutions the policy makers can look into in order to eliminate or improve the situation of malnutrition in the continent.

We make sure the source code used to clean and visualize the data is openly available for the research community who would want to make their own research continuing from our conclusion.

References

1) Data Bank – The World Bank

2) Akombi, B. J., Agho, K. E., Merom, D., Renzaho, A. M., & Hall, J. J. Child malnutrition in sub-Saharan Africa: A meta-analysis of demographic and health surveys (2006-2016). PLoS ONE, 2017, 12(5).

3) Schober, Patrick MD, PhD, MMedStat; Boer, Christa PhD, MSc; Schwarte, Lothar A. MD, PhD, MBA (2018). Correlation Coefficients: Appropriate Use and Interpretation. Anesthesia & Analgesia, May 2018, 126(5), p 1763-1768.