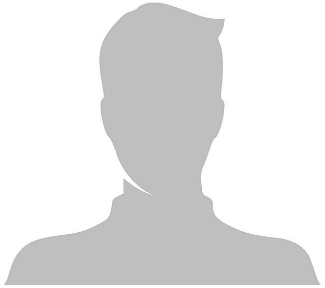
The Impact of Maternal Education on Child Malaria Rates

25/05/2024

Our Team



Usman Hassan
Role: Presenter 1

Abdur-Rahman Olaniyan
Role: Presenter 2

Project Lead
Name: Bamidele Faith

Assistant Project Lead
Name: Wellspring Praise

Query Analyst
Name: Blessing Domola

Other Active Members

Name: Ayodeji Saberedowo
Name: Adeleye Kayode
Name: Christian Utsu
Name: Ibifaka Campbell
Name: Fajemisin Adeniyi
Name: Olubola Titilope Adegbosin

Problem Statement

- Despite efforts, malaria remains a major health issue in many regions. This project explores the impact of maternal education on reducing child malaria rates, lacking comprehensive analysis.

- Malaria persists as a significant health challenge. This study investigates whether higher maternal education correlates with lower child malaria rates, filling the gap in detailed analysis.

- Malaria continues to affect many regions severely. This project aims to provide detailed evidence on how maternal education can reduce child malaria rates, informing targeted interventions..

Existing Solutions

- Previous studies, such as Joseph D. Njau (2014), have explored the statistical relationship between maternal education and childhood malaria infection. These studies aimed to highlight these relationships and inform strategies to reduce the disease burden among children.

- The previous solution did not specifically address the levels of maternal education and instead generalized across all childhood ages and regions. However, children under five are particularly vulnerable, constituting 70% of malaria deaths, necessitating a focused approach on this age group and regional variations.

Our approach

- Data Gathering: We obtained the dataset from UNICEF's official website https://data.unicef.org/topic/child-health/malaria/, which provided comprehensive data on child health and malaria.

- Data Description/Used: The dataset consisted of 17 sheets, and we carefully selected the three most relevant sheets that aligned with our project's objectives. This ensured that our analysis was focused and effective.

- Data Visualization: We utilized the Seaborn library's histogram function to visualize the correlation between mothers' educational levels and their children's malaria rates. This visualization provided valuable insights into the relationship between these variables.

- Model Training: To further explore and validate our findings, we used Linear regression model for get pattern from our dataset and we divided the dataset into training and testing sets. We used the mother's education level as the dependent variable and the National column as the independent variable.

- Model Evaluation: We employed the coefficient of determination (R-squared) to assess the strength of the relationship between the trained and test variables.

Dataset Description

- Data Collection and Preparation
  Source: UNICEF Data
  Sheets Used: MLRCARE, MLRDIAG, MLRACT
- Data Wrangling/EDA
  Handled null values, duplicates, and filtered necessary indicators. Each sheet was cleaned and analyzed separately
- Data Description
  Key Features: Country code, country names, UNICEF reporting and program regions, World Bank income groups, Year, data sources, Malaria infection rates, Gender distribution, Rural and urban populations, Income groups, Maternal education levels
- Data Visualization
  Tools Used : MatplotLib And Seaborn
- Key Observations :
  MLRCARE: Percentage of children (under age 5) with fever for whom advice or treatment was sought.
  MLRDIAG: Malaria Diagnostic Usage - Percentage of febrile children (under age 5) who had malaria diagnosis.
  MLRACT: First-line treatment (ACT) for children under age 5 with fever-
- Data Modelling
  Algorithm Used: Linear regression
  Evaluation Metrics: Mean Absolute Error (MAE), r2_score and Root Mean Square Error (RMSE) to assess model performance.

Model

- Linear Regression was chosen as the modeling technique, proving to be the most suitable approach for our dataset. The coefficients generated by the model revealed a significant correlation between maternal education level and children's malaria rate. This analysis enabled us to draw meaningful conclusions about the impact of maternal education on children's health outcomes, specifically in regards to malaria rates. The model's results provided a deeper understanding of the relationship between these variables, supporting our project's objectives.

- **Step 1:** We initiated our analysis by importing the necessary libraries, including Matplotlib, Seaborn, and Pandas, to facilitate data manipulation and visualization..

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_absolute_error, mean_squared_error, r2_score
from sklearn.linear_model import Ridge
from sklearn.linear_model import Lasso
from sklearn.impute import KNNImputer
from sklearn.preprocessing import MinMaxScaler
from sklearn.model_selection import train_test_split, cross_val_score
```

Model

- **Step 2:** We began by examining the provided dataset, focusing on the three relevant sheets that contained the data for our correlation analysis. This step enabled us to understand the structure and content of the data.

```
pd.set_option("display.max_columns", None)
mlrcare = pd.read_excel("Final_malaria.xlsx", sheet_name = 0)
mlrcare.head(2)
```

	ISO	Countries	UNICEF Reporting region	UNICEF Programme Region	World Bank Income Group (2022)	Year	Short Source	Long Source	National	Male	Female	Rural	Urban	Poorest	Second	Middle	Fourth	Richest	None	Primary	Sec High
0	AFG	Afghanistan	South Asia	ROSA	Low income	2015	DHS 2015	Demographic and Health Survey 2015	63.2	NaN	NaN	62.3	65.9	61.1	62.0	61.1	67.7	64.1	NaN	NaN	N
1	AFG	Afghanistan	South Asia	ROSA	Low income	2018	Afghanistan Health Survey 2018	Afghanistan Health Survey 2018. Amsterdam: KIT...	62.1	60.9	63.2	63.1	59.5	56.6	67.2	59.9	63.5	63.6	60.9	66.1	N

Model

- **Step 3:** We cleaned the dataset by removing columns that were deemed irrelevant to our analysis, ensuring a more focused and efficient exploration of the data.

```
mlrcare = mlrcare.drop(columns= ["ISO","UNICEF Reporting region","UNICEF Programme Region",
                                 "World Bank Income Group (2022)","Long Source","Short Source"], axis=1)
```

- **Step 4:** Using KNN impute to Handle missing values because each entity has different pattern of values.

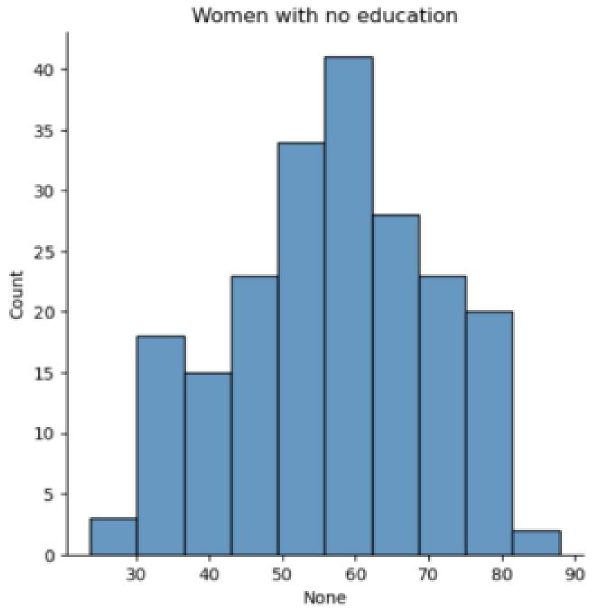
```
# Columns to impute
columns_to_impute = ['National', 'Male', 'Female', 'Rural', 'Urban', 'Poorest', 'Second', 'Middle', 'Fourth', 'Richest', 'None', 'Primary', 'Sec & Higher']

#Create a KNN Imputer object
imputer = KNNImputer(n_neighbors=5, weights='uniform')

#Fit and transform the data
imputed_data = imputer.fit_transform(mlrcare[columns_to_impute])

# Convert the imputed data back to a pandas DataFrame
imputed_data = pd.DataFrame(imputed_data, columns=columns_to_impute)

# Replace the original columns with the imputed data
mlrcare[columns_to_impute] = imputed_data
```

```
mlrcare.isna().sum()
```

Model

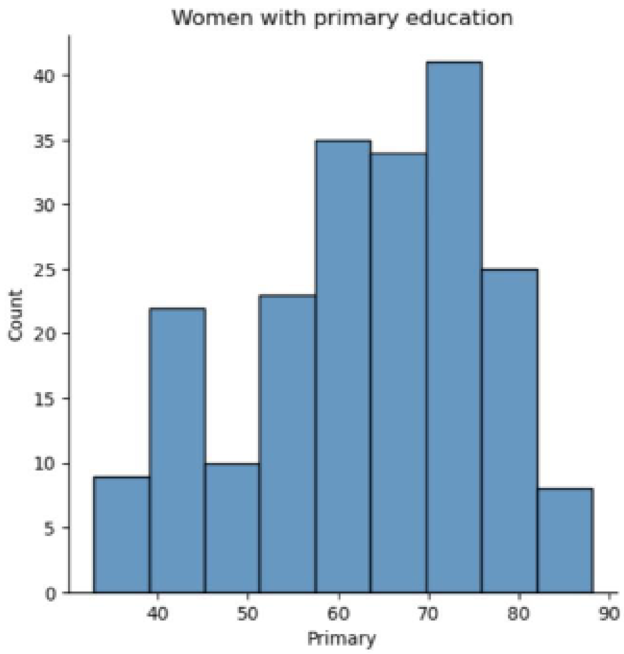
- **Step 5:** We created visualizations for each of the three sheets to gain a deeper understanding of the data distribution and relationships.

❑ **MLRCARE** - Percentage of children (under age 5) with fever for whom advice or treatment was sought.

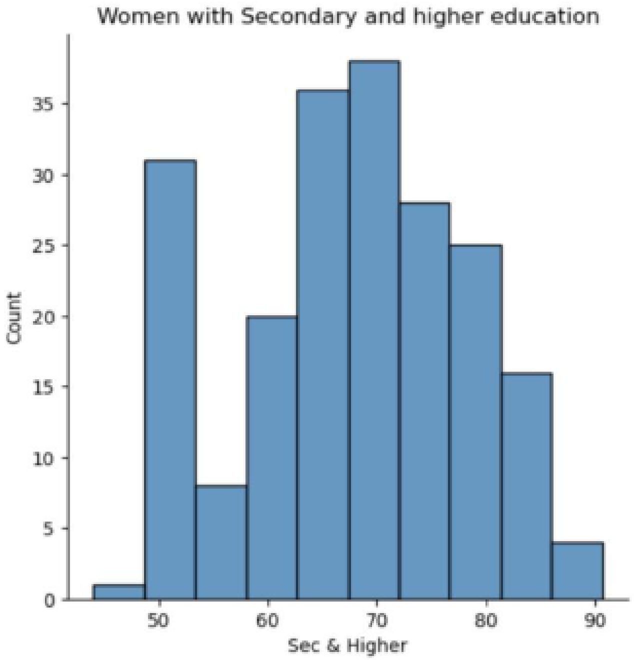


The highest count is seen in the age group of 55-60%, with around 40 Countries. This implies that about 55-60% of children with fever in most countries had their mothers with no education sought advice for their treatment.

Model



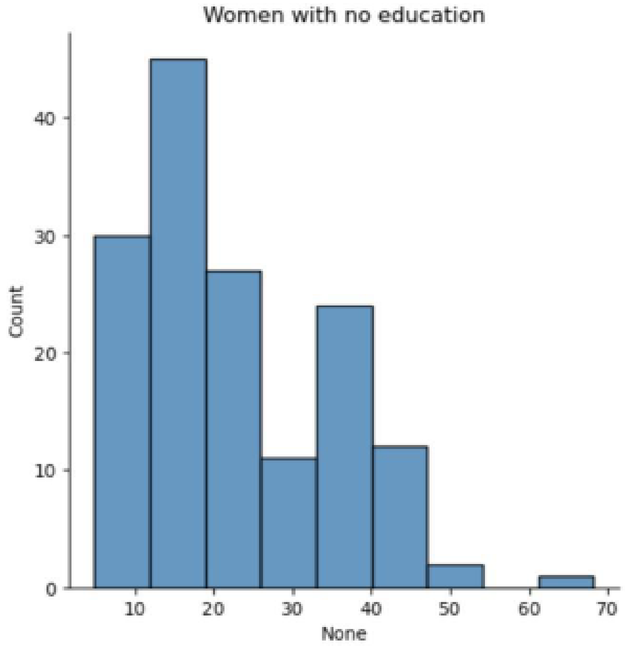
The visual above implies that about 65% of children with fever in most countries had their mothers with primary education sought advice for their treatment.



The visual above implies that about 70% of children with fever in most countries had their mothers with Secondary or Higher education sought advice for their treatment.

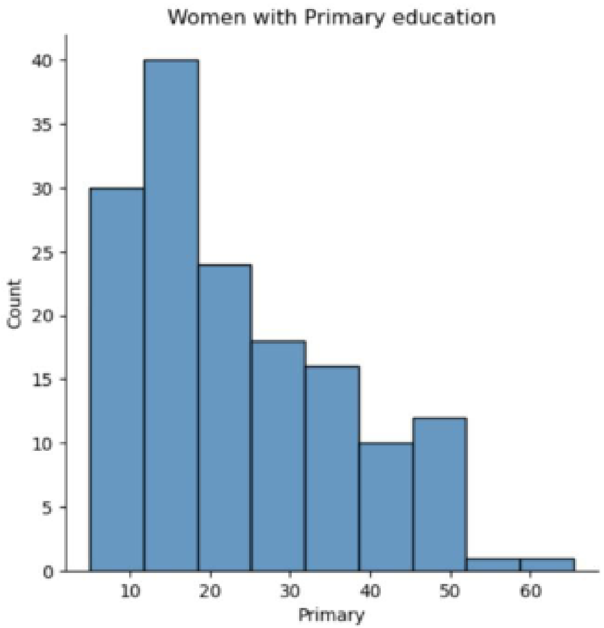
Model

- **MLRDIAG -** Malaria Diagnostic Usage - Percentage of febrile children (under age 5) who had malaria diagno
  The visualization below analyses our second sheet.

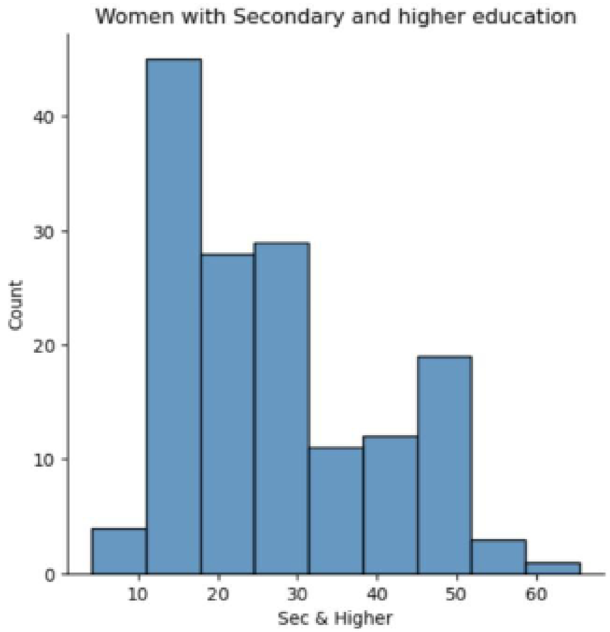


This plot implies that about 12-18% of children who were tested for malaria in most countries had their mothers having no education.

Model



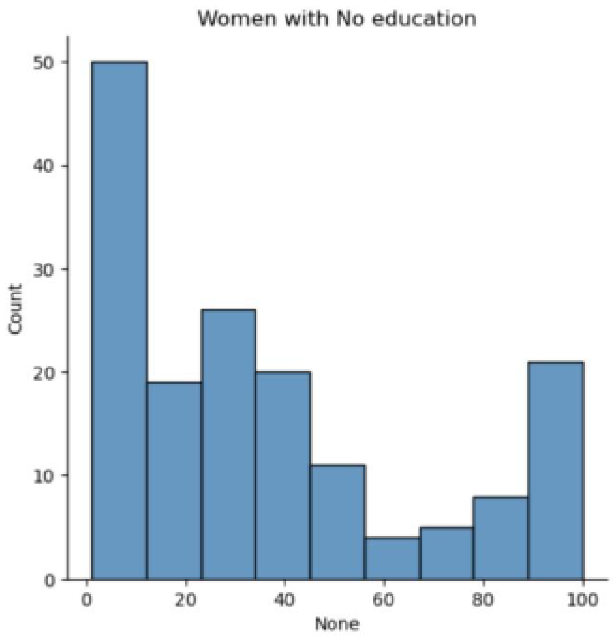
The above plot infers that about 12-18% of children who were tested for malaria in most countries had their mothers having **primary** education.



The above plot infers that about 12-17% of children who were tested for malaria in most countries had their mothers having **Secondary and Higher** education.

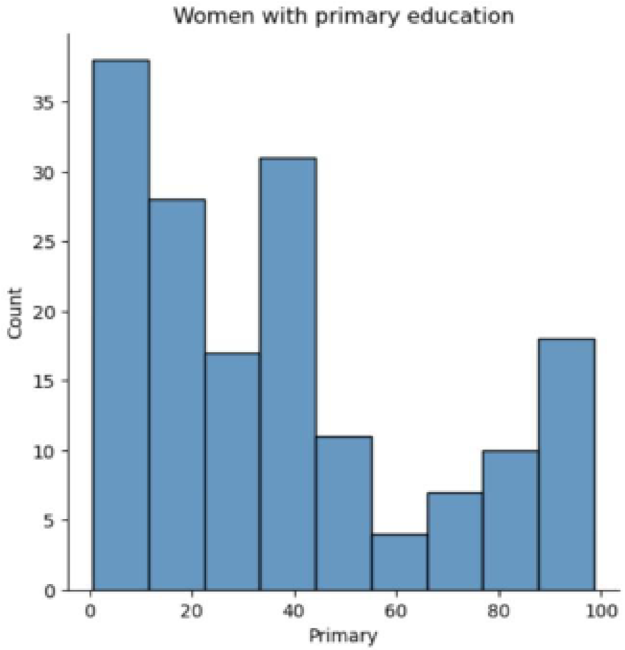
Model

- **MLRACT -** First-line treatment (ACT) for children under age 5 with fever- Percentage of febrile children (under age 5) receiving ACT (first-line anti- malarial drug), among those receiving any antimalarial drugs. This is our third sheet. The visualization below analyses our third sheet.

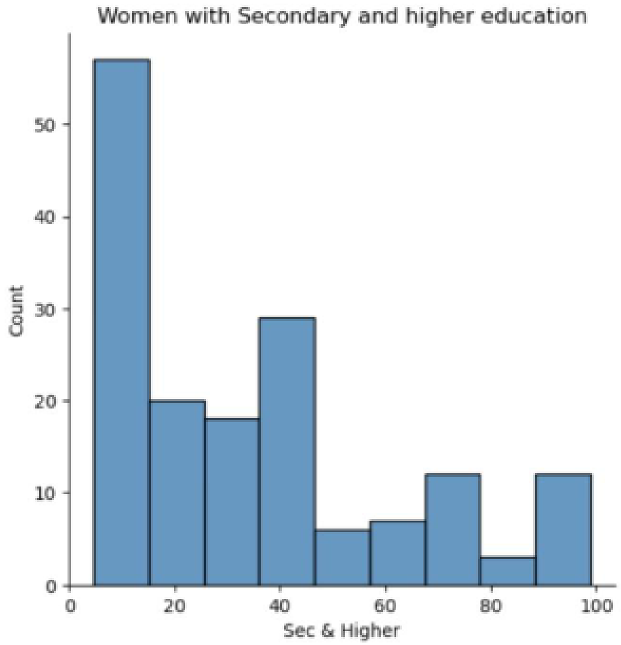


The chart above indicates that about 1-10% of children who recieves antimalaria treatment in most countries had their mothers having no education.

Model



The chart above indicates that about 1-15% of children who recieves antimalaria treatment in most countries had their mothers having primary education.



The chart above indicates that about 5-15% of children who recieves antimalaria treatment in most countries had their mothers having Secondary and higher education.

Model

- **Step 6:** We divided the dataset into training and testing sets to ensure model robustness and generalizability thereby reducing the risk of overfitting and improving reliability.

```
# building model for the mlrcare dataset
X = mlrcare_selected.drop(columns = ['National'])
y = mlrcare_selected[['National']]
```

```
x_train, x_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
```

- **Step 7:** MinMaxScaler - a specific technique for normalization, was used to scale the values to a common range to prevent feature dominance.

```
numerical = ['None', 'Primary', 'Sec & Higher']
```

```
scaler = MinMaxScaler()
scaler.fit(x_train[numerical])
```

```
▾ MinMaxScaler
MinMaxScaler()
```

```
x_train[numerical] = scaler.transform(x_train[numerical])
x_test[numerical] = scaler.transform(x_test[numerical])
```

Model

- **Step 8:** We employed linear regression, calculating the coefficient of determination (R-squared) to assess model performance and the strength of the correlation between maternal education level and children's malaria rate.

```
model = LinearRegression()
model.fit(x_train, y_train)
```

```
▾ LinearRegression
LinearRegression()
```

```
coefficients = model.coef_
print(coefficients)
```

```
[[36.32372328 14.64173702 10.60432609]]
```

From the model, we got 3 different coefficients, which indicated the coefficients of (None, Primary, Sec & Higher) variables respectively. The whole procedures were performed to the three sheets we used and it provided a deeper relationship into the effects of maternal education on the rate of malaria on children.

Summary

- Our project was aimed to address a critical public health issue by investigating the impact of maternal education on child malaria rates. Through this research, we were able to uncover valuable insights into how a mother's educational attainment can significantly influence her children's health, particularly in malaria-prone regions.

- This datasets had seventeen excel sheets. It was very difficult for us to decide which of the sheets we were going to use. At first we all agreed to use the MLRDIAG sheet but we could not derive insight from only that sheet. At the end, we chose MLRCARE, MLRDIAG, and MLRACT sheet. They were the only sheets that contain the informations we needed for our project. We had to process the data in order to use it with python.

- The results of this indicated that maternal education indeed has influence on their children's malaria rate especially in vulnerable children. The critical analysis showed that more children from rural areas have uneducated mothers and these children are less likely to get treatment. It also indicate that more children from urban areas have better access to health care because they have more educated mothers. The urban area being closer to healthcare facilities makes it easier for the educated woman to access healthcare.

Summary

- The findings indicated that the higher a mother's education, the lesser chance of the child being infected with malaria.

Recommendations:

- Build more schools in the rural areas and enlighten parents in those areas about the importance and the necessity of sending their children to school.
- A multifaceted approach that includes investment in maternal education should be taken by the government.
- Organize campaigns to teach women mostly in the rural areas the importance of using an insecticide-treated mosquito net and also the importance of seeking care when her child has fever.
- Make healthcare available in the rural areas
- During pregnancy, hospitals should organize short lectures to educated women on malaria symptoms in children and the importance of treating the disease.