Pattern discovery for Malaria in Africa

KAHUMA ANDREW
Bsc. Software Engineering
Makerere University
2100704644, 21/U/04644/PS
andrewkah2022@gmail.com

Introduction:

Malaria is a significant public health concern in many parts of the world, particularly in Africa. In this report, we explore patterns and correlations in a dataset related to malaria incidence and preventative measures in African countries. The dataset contains information on various factors, including malaria incidence rates, access to preventative measures, water services, sanitation services, and geographical location. We aim to uncover insights that may help in understanding the impact of different variables on malaria incidence.

1. Correlation between Sanitation and Incidence of Malaria:

Hypothesis:

Poor sanitation can create breeding grounds for mosquitoes, increasing the risk of malaria. We hypothesize a negative correlation between the percentage of the population using at least basic sanitation services and the incidence of malaria per 1,000 population at risk.

Method:

- We calculated the correlation coefficient between the two variables using the Pearson correlation method.
- A negative correlation coefficient closer to -1 would indicate a stronger negative correlation.

Findings:

The correlation coefficient between the percentage of the population using at least basic sanitation services and the incidence of malaria was -0.548. Although not strongly negative, the result suggests a moderate negative association between sanitation services and malaria incidence. This finding may indicate that improving sanitation services could potentially contribute to reducing malaria incidence.

2. Effect of Preventative Measures on Malaria Incidence:

Hypothesis:

Preventative measures such as using insecticide-treated bed nets (IBNs) or intermittent preventive treatment (IPT) for pregnant women can reduce malaria incidence. We think a negative correlation between the percentage of the population

using IBNs or IPT and the incidence of malaria per 1,000 population at risk would be needed.

Method:

- We calculated the correlation coefficient between the two variables (IBNs or IPT usage and malaria incidence) using the Pearson correlation method.

Findings:

The correlation coefficients between the percentage of the population using IBNs or IPT and malaria incidence were found to be 0.5277 and 0.5387, respectively. While the correlations are quite strong, they indicate a positive association between IBN and IPT usage and malaria incidence. This suggests that these preventative measures may not have some impact on reducing malaria incidence, but other factors may also be influencing the results.

3. Urban vs. Rural Differences in Malaria Incidence and Preventative Services:

Hypothesis:

Disparities may exist in malaria incidence and access to preventative services between urban and rural populations. We compare the percentage of the rural and urban population with access to basic drinking water and sanitation services and investigate their correlation with malaria incidence.

Method:

- We compare the mean malaria incidence and access to preventative services between urban and rural populations.
- We perform a t-test to determine if there are significant differences between the two groups.

Findings:

Our analysis revealed that the mean malaria incidence is higher in rural areas compared to urban areas. Additionally, there are significant differences in the percentage of the population with access to basic drinking water and sanitation services between urban and rural areas. These findings suggest that addressing the disparities in access to preventative services in rural areas could be essential in reducing malaria incidence.

4. Impact of Basic Drinking Water Services on Malaria Incidence:

Hypothesis:

Availability of basic drinking water services is a marker of public health infrastructure that may impact malaria incidence indirectly. We investigate if areas with higher access to basic drinking water have lower malaria incidence.

Method:

- We calculate the correlation coefficient between the percentage of the population with access to safe drinking water and the incidence of malaria.

Findings:

The correlation coefficient between the percentage of the population with access to Basic drinking water services and malaria incidence was found to be -0.470. This result indicates a moderate negative association between safe drinking water access and malaria incidence. It suggests that improving access to safe drinking water may contribute to reducing malaria incidence in affected regions.

5. Trends Over Time:

Hypothesis:

There may be temporal trends in malaria incidence and preventative measures over the years. We examine the trends in malaria incidence and the usage of preventative measures over time.

Method:

- We plot time series plots to visualize the trends in malaria incidence and preventative measure usage over the years.

Findings:

The time series analysis shows that malaria incidence has varied over the years, with fluctuations observed in different regions. Similarly, the usage of preventative measures such as IBNs and IPT has shown some fluctuations but may have an increasing trend over time. Further analysis of these trends could provide valuable insights into the effectiveness of interventions over the years.

6. Geographical Patterns:

Hypothesis:

Geographical factors may influence malaria incidence. We explore the geographical distribution of malaria incidence rates and look for clustering or hotspots of high or low incidence.

Method:

- We plot a choropleth map to visualize the geographical patterns of malaria incidence.

Findings:

The geospatial analysis reveals specific regions or clusters with high or low malaria incidence rates. Identifying these geographical patterns can help in targeting interventions and resources to the areas with the highest malaria burden.

Conclusion:

The analysis of the dataset provides valuable insights into the patterns and relationships related to malaria incidence and preventative measures. While some correlations suggest potential interventions to reduce malaria incidence, it is essential to interpret the findings with caution.

References:

Lydia70. (2020). *Malaria in Africa*. Retrieved from Kaggle: https://www.kaggle.com/datasets/lydia70/malaria-in-africa

The World Bank. (2023). *Incidence of malaria (per 1,000 population at risk)*. Retrieved from THE WORLD BANK/Data: https://data.worldbank.org/indicator/SH.MLR.INCD.P3

The World Bank. (2023). *People using at least basic drinking water services (% of population)*. Retrieved from THE WORLD BANK/Data: https://data.worldbank.org/indicator/SH.H2O.BASW.ZS

The World Bank. (2023). *People using at least basic sanitation services (% of population)*. Retrieved from THE WORLD BANK/Data: https://data.worldbank.org/indicator/SH.STA.BASS.ZS