

Global Malaria Trends Analysis (2015–2023)

Project Overview

Malaria remains one of the most critical global public health challenges, disproportionately affecting developing regions.

This project analyzes **global malaria trends from 2015–2023** using data from the **World Health Organization (WHO)** to uncover patterns in:

- Confirmed malaria cases
- Deaths due to malaria
- Treatment coverage

The goal is to **support data-driven decision-making** for governments, NGOs, and health organizations by identifying high-risk regions and gaps in treatment effectiveness.

Business & Analytical Objectives

1. Analyze the **global trend** of malaria cases over time
2. Identify **continents and countries** with the highest confirmed cases and death rates
3. Compare **confirmed cases vs. treated cases** across top-affected countries
4. Understand **regional risk patterns** based on geography and climate
5. Visualize **country-level contribution** to global malaria burden

Data Sources

We used 3 **WHO Global Health Observatory (GHO)** Datasets.

1. GHO's Number of confirmed malaria cases ([Link](#)).
2. GHO's Malaria deaths, number, estimate ([Link](#)).
3. GHO's Number of malaria cases treated with any first line TX courses (including artemisinin-based combination therapies (ACTs)) ([Link](#)).

Data Preprocessing (Python)

Raw datasets were cleaned and optimized before analysis.

Key Steps:

- Removed irrelevant columns
- Standardized column names
- Checked for **null values and duplicates**
- Merged three datasets using a **left join**
- Handled missing values:
 - Dropped rows with missing death data
 - Replaced missing treated-case values with 0
- Exported a clean, analysis-ready dataset

Why this matters: Clean data ensures reliable visualizations and accurate insights.

Code 1: Python Code for Merging Datasets

```
✓ 0s  ⏎ import pandas as pd

data1 = pd.read_excel('/content/Number of confirmed malaria cases.xlsx')
data2 = pd.read_excel('/content/Malaria deaths, number, estimate.xlsx')
data3 = pd.read_excel('/content/Number of malaria cases treated with any first line tx courses

merged_data = pd.merge(data1, data2, on=['Continent', 'Country', 'Year'], how='left')

merged_data = pd.merge(merged_data, data3, on=['Continent', 'Country', 'Year'], how='left')

print(merged_data.head())
```

Figure 1: Head of Merged Dataset

```
→   Continent Country Year Confirmed_cases Deaths Treated_cases
  0    Africa  Angola  2023      9 098 006  16169.0     1 316 833
  1    Africa  Angola  2022      7 858 860  20286.0     7 858 802
  2    Africa  Angola  2021      8 325 921  20493.0     7 793 251
  3    Africa  Angola  2020      7 343 696  17198.0           NaN
  4    Africa  Angola  2019      7 054 978  15253.0           NaN
```

Code 2: Python Code to check for null and duplicate values

```
✓ 0s  ⏎ print(f"Null values per column: {merged_data.isnull().sum()}")
print(f"Total duplicated rows: {merged_data.duplicated().sum()}")

→ Null values per column: Continent          0
   Country          0
   Year            0
   Confirmed_cases  0
   Deaths          15
   Treated_cases   670
   dtype: int64
   Total duplicated rows: 0
```

Code 3: Python Code for Dropping and Replacing Null Values

```
▶ cleaned_data = merged_data.dropna(subset=['Deaths'])
  cleaned_data['Treated_cases'] = cleaned_data['Treated_cases'].fillna(0)

  print(cleaned_data.head())
→ Continent Country Year Confirmed_cases Deaths Treated_cases
  0 Africa Angola 2023 9 098 006 16169.0 1 316 833
  1 Africa Angola 2022 7 858 860 20286.0 7 858 802
  2 Africa Angola 2021 8 325 921 20493.0 7 793 251
  3 Africa Angola 2020 7 343 696 17198.0 0
  4 Africa Angola 2019 7 054 978 15253.0 0
```

Code 4: Recheck for Null values and Duplicates

```
✓ 0s [9] print(f"Null values per column: {cleaned_data.isnull().sum()}")
  print(f"Total duplicated rows: {cleaned_data.duplicated().sum()}")

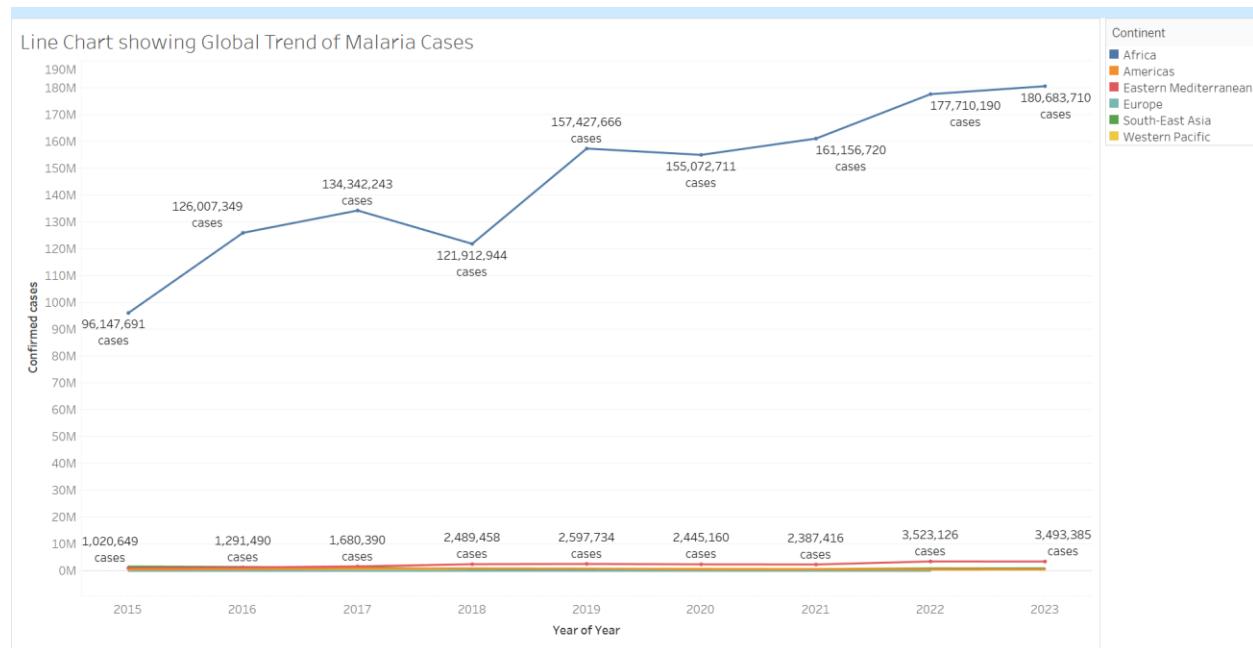
→ Null values per column: Continent 0
  Country 0
  Year 0
  Confirmed_cases 0
  Deaths 0
  Treated_cases 0
  dtype: int64
  Total duplicated rows: 0
```

Exploratory Data Analysis (Tableau)

1. Global Trend of Malaria Cases

A line chart was used to visualize confirmed malaria cases from 2015–2023.

Figure 2: Line Chart Representing Global Malaria Trend



Insight:

Global malaria cases show an **overall upward trend**, indicating either:

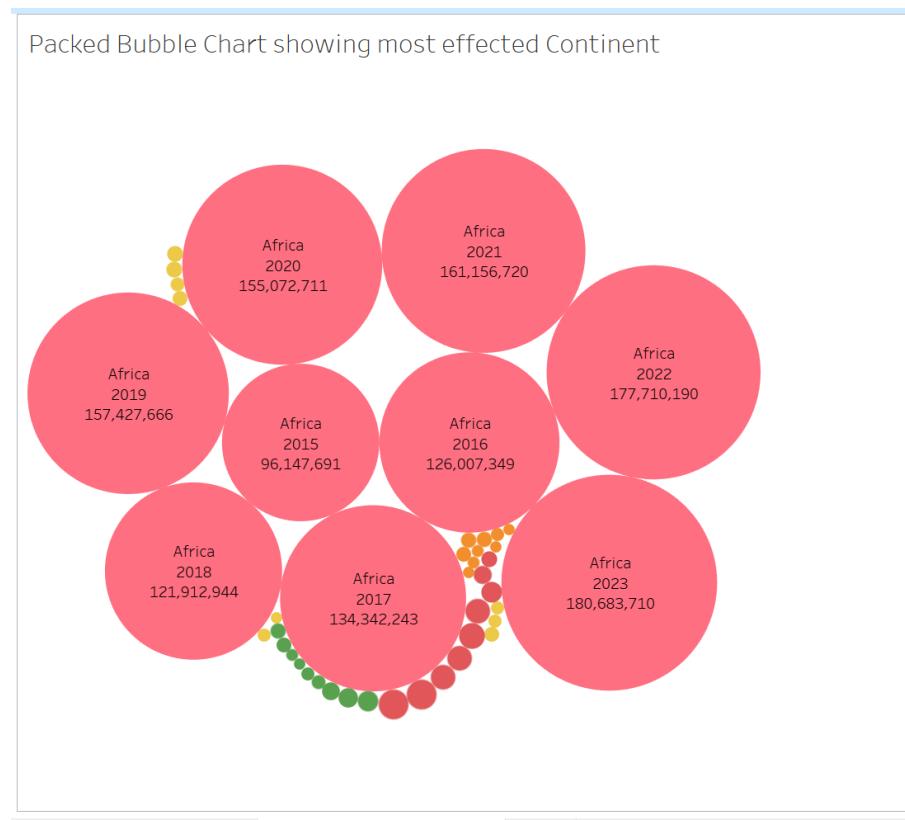
- Increased spread of the disease, or
- Improved detection and reporting

Given diagnostic technology has remained stable, the trend suggests a **real increase in cases**, highlighting the need for preventive interventions.

2. Most Affected Continent & Death Rates

A packed bubble chart identified Africa as the most affected continent.

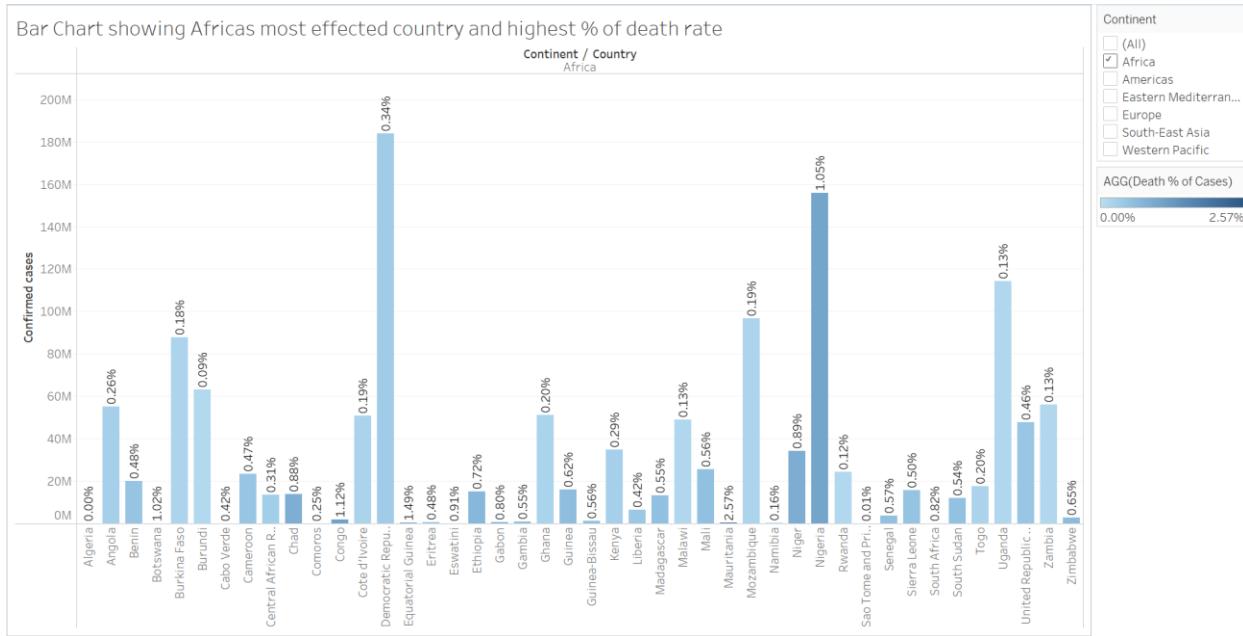
Figure 3: Packed Bubble Chart showcase most effect Continent over the years



A bar chart compared African countries by:

- Confirmed cases
- Death rate (% of confirmed cases)

Figure 4: Bar Chart most effected cases of which % of death



Calculated Field:

Death % = Deaths / Confirmed Cases

Figure 5: Calculated Table to find % of death of total cases

Death % of Cases

```
SUM([Deaths]) / SUM([Confirmed cases])
```

The calculation is valid. 1 Dependency ▾ Apply OK

Key Findings:

- Democratic Republic of Congo: highest confirmed cases
- Nigeria: high cases with elevated death rate

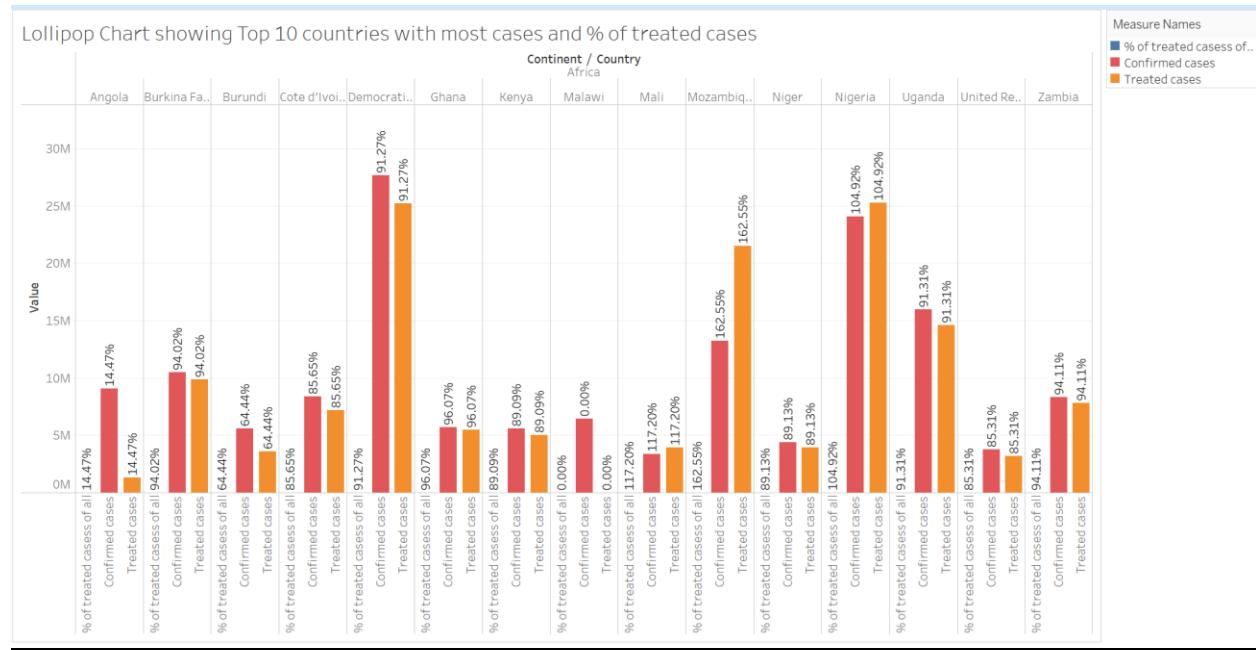
- Mauritania: lower case count but **highest death percentage**

Insight: Case volume alone does not indicate healthcare effectiveness — death rate highlights resource gaps.

3. Confirmed vs. Treated Cases (Top 15 Countries)

A **side-by-side bar chart** compared confirmed cases with treatment coverage.

Figure 6: Side by side Chart to show no of confirmed cases as to percentage of treated cases



Insight:

- Top 15 countries (mostly African) show ~**90% treatment coverage**
- Indicates strong treatment response, despite rising case numbers

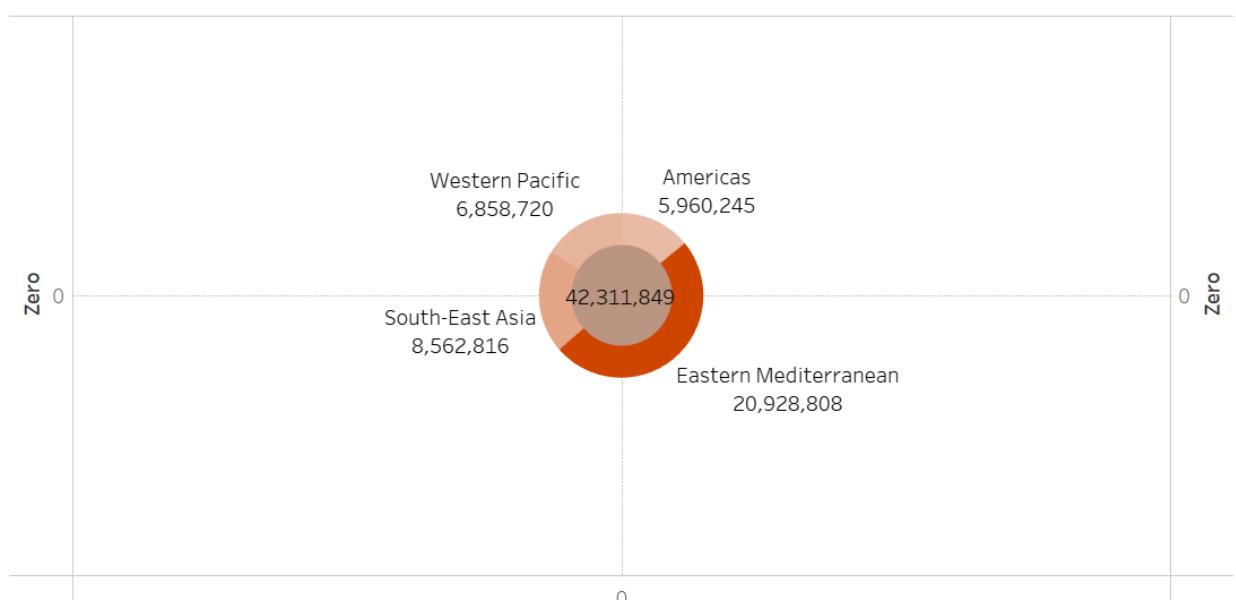
Conclusion: Treatment is effective, but **prevention is the main gap**.

4. Regional Risk Analysis (Excluding Africa)

A **donut chart** analyzed malaria distribution across other regions.

Figure 7: Donut Chart to show proportion of Malaria cases around the world

Other continents with most cases



Findings:

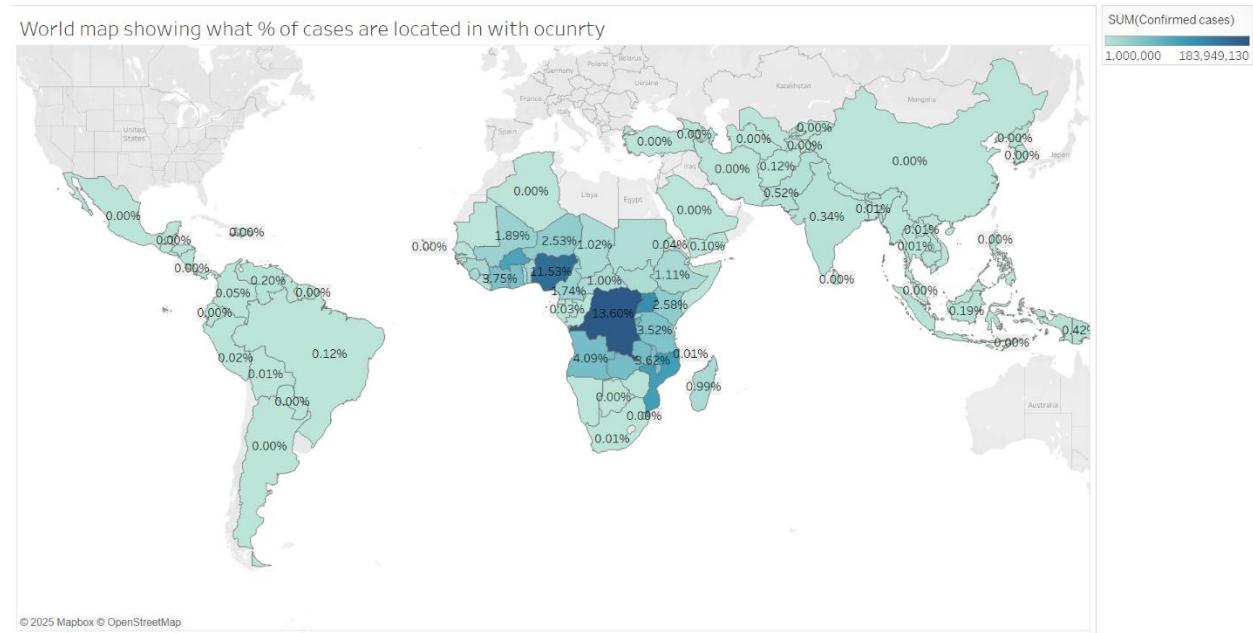
- Eastern Mediterranean & South-East Asia have the highest case proportions
- Strong correlation with **hot, humid, tropical climates**

Insight: Climate plays a major role in malaria prevalence.

5. Country Contribution to Global Cases

A **world map** visualized each country's share of global malaria cases.

Figure 8: World map showing % of cases around the world



Key Observation:

- India & Pakistan show significantly higher case percentages than neighboring countries
- Highlights opportunities to adopt successful regional prevention strategies

Tableau Dashboard & Story

All visualizations were consolidated into:

- **Interactive Tableau Dashboard** for stakeholder analysis
- **Tableau Story** to guide users through insights step-by-step

Figure 9: Tableau Dashboard ([Link](#))

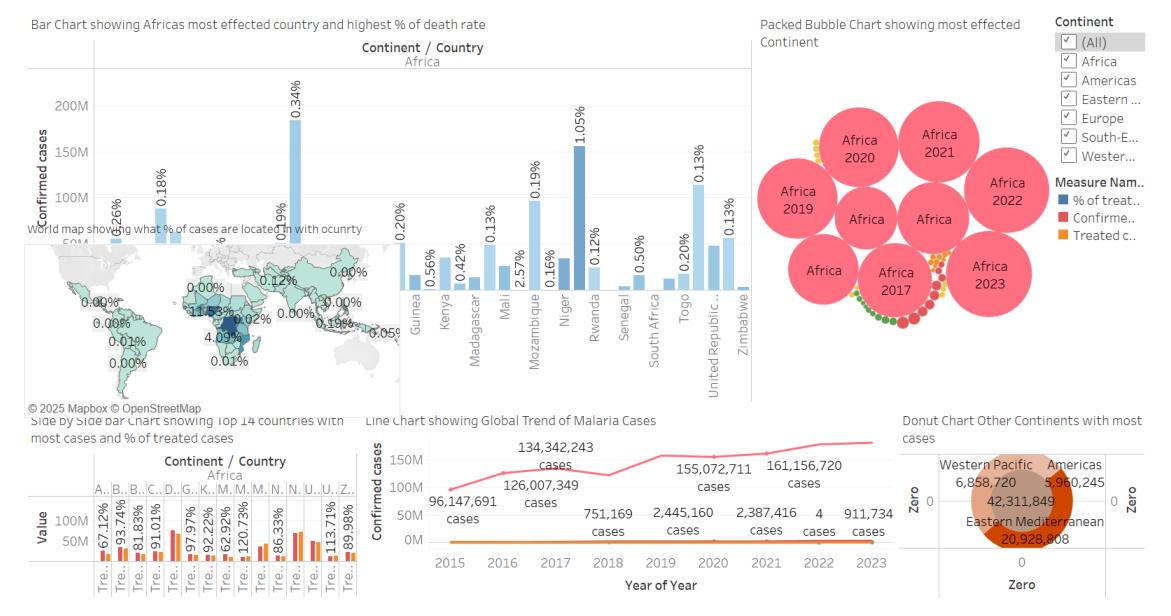
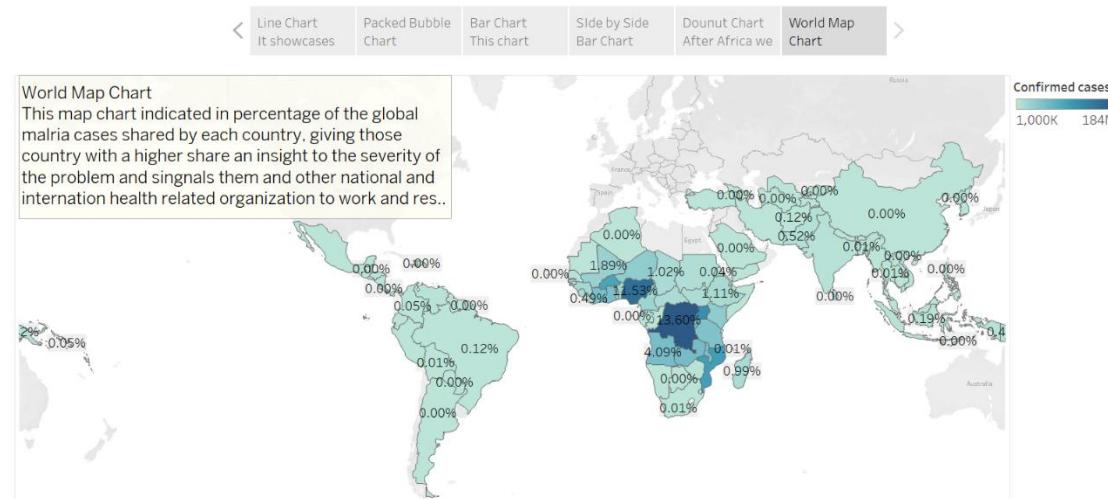


Figure 10: Tableau Story ([Link](#))

Story 1



Key Insights Summary

1. Malaria cases are increasing globally
2. Africa remains the epicenter, but Asia also shows rising risk
3. Treatment coverage is relatively strong (~90%)
4. Prevention efforts lag behind treatment
5. Climate and geography strongly influence case distribution

Recommendations

1. Shift focus from treatment → prevention
2. Improve drainage systems to eliminate mosquito breeding grounds
3. Large-scale mosquito control programs
4. Community education on night-time protection
5. Cross-border collaboration for regional containment

Future Work (Machine Learning Extension)

1. Build a **predictive model** using:
 - 1.1. Historical malaria cases
 - 1.2. Weather and temperature data
2. Forecast outbreak risk by region
3. Enable **early intervention planning**

Links (Datasets and Tableau Dashboard & Story)

Links to Datasets

<https://www.who.int/data/gho/data/indicators/indicator-details/GHO/number-confirmed-malaria-cases>

<https://www.who.int/data/gho/data/indicators/indicator-details/GHO/estimated-number-of-malaria-deaths>

<https://www.who.int/data/gho/data/indicators/indicator-details/GHO/number-of-malaria-cases-treated-with-any-first-line-tx-courses-including-artemisinin-based-combination-therapies-acts>

Dashboard Link

https://public.tableau.com/views/Dashboard1_17496442403670/Dashboard1?:language=en-US&:sid=&:redirect=auth&:display_count=n&:origin=viz_share_link

Story Link

https://public.tableau.com/views/Story1_17496443958410/Story1?:language=en-US&:sid=&:redirect=auth&:display_count=n&:origin=viz_share_link