**Healthcare Spending Impact Analysis:  
Optimizing Health Investment Strategies Across Africa**

Team: AfriHealth

Capstone Project:  
CSIS 503: Data Science and Analytics  
Summer 2025

Instructor:

Noble Anumbe, Ph.D.

August 23, 2025

By: Group 7  
Jaji Jimoh Oludare – jimoho@osiriuniversity.org  
Tatenda Chidewu – tachidewu@osiriuniversity.org  
Rediet Tedla – retedia@osiriuniversity.org  
Kingsley Njoku – kinjoku@osiriuniversity.org  
Daniel Cole – dancole@osiriuniversity.org  
Peter Fondoh – pfondoh@osiriuniversity.org

  
**Osiri University,**

**Lincoln, Nebraska, USA**

jjj

**1. Project Title**

***Healthcare Spending Impact Analysis: Optimizing Health Investment Strategies Across Africa.***

**2. Introduction**

Healthcare systems across Africa face the dual challenge of limited resources and high disease burdens. While many countries invest in health, the efficiency and impact of that spending vary widely. This project analyzes how different forms of healthcare financing — government, private, out-of-pocket, and external — affect health outcomes and identifies high-efficiency investment strategies.

**Objective:**  
To examine how various types of healthcare spending — government, private, out-of-pocket (OOP), and external — affect health outcomes across African countries. The analysis seeks to identify high-efficiency health systems and recommend investment strategies that maximize health gains per dollar spent.

**Relevance:**  
Using data from WHO and World Bank, findings aim to provide evidence-based insights to help policymakers optimize health investments and advance UN Sustainable Development Goal 3 (Good Health and Well-Being) by identifying scalable, efficient health system models.

**3. Dataset Description**

* **Sources:** WHO Global Health Expenditure Database (GHED) and World Bank Health, Nutrition, and Population (HNP) Statistics
* **Size & Structure:** 1,054 rows × 22 columns, covering 47 African countries from 2000–2022
* **Key Variables:**
  + CHE\_GDP (Current Health Expenditure as % of GDP)
  + GGHE\_pct\_GDP (Government Health Expenditure)
  + OOP\_pct\_CHE (Out-of-Pocket expenditure)
  + Life\_Expectancy, Infant\_Mortality, DPT\_Immunization
  + GDP\_per\_capita, Population
* **Preprocessing:**
  + Merged 12 WHO indicators
  + Standardized country names
  + Applied *Weighted Mean 66POP* for regional aggregation
  + No imputation — missing values handled per World Bank rules

**4. Methodology**

**Data Pipeline:**

1. **Acquisition:** WHO + World Bank datasets
2. **Cleaning & Merging:** Unified african\_health\_spending\_outcomes.csv
3. **EDA:** Correlation matrices, scatter plots, bar charts, time-series trends
4. **Modeling:** Linear regression to predict life expectancy
5. **Efficiency Analysis:** Efficiency score = Actual – Predicted life expectancy
6. **Regional Aggregation:** Weighted Mean 66POP for robust summaries – this implies that no aggregate is shown if countries with missing data represent more than one third of the total population of the group.
7. **Policy Recommendations:** Based on high-efficiency case studies

**Techniques & Tools:**

* **EDA:** Correlation analysis, visual trends
* **Modeling:** Linear regression to isolate effects of different spending types
* **Tools:** Python (pandas, numpy, scikit-learn, matplotlib, seaborn) in Jupyter Notebook
* **Rationale:** Regression quantifies spending impact; efficiency scoring ranks countries; weighted mean ensures representative regional results

**5. Key Findings & Results**

* **Model Performance:** R² = 0.243, RMSE = 5.57 years
* **Efficiency Leaders:** Algeria and Cabo Verde outperform predictions by large margins (Algeria: +11 years above predicted)
* **Regional Patterns:** Southern Africa spends most but has lower life expectancy than East Africa (likely due to HIV prevalence and inequity)
* **Strongest Predictor:** DPT immunization rate — each 1% increase adds ~0.225 years to life expectancy
* **Government Spending Effect:** Negative coefficient — possibly reverse causality (sicker populations need more spending)
* **Out-of-Pocket Spending:** Higher OOP correlates with higher infant mortality and reduced access

**6. Insights & Recommendations**

**Key Insights:**

* Efficiency > Spending — smart investments outperform high expenditure
* Preventive care yields strong health returns
* Reducing OOP improves equity and financial protection
* Data-driven monitoring enhances policy targeting

**Policy Recommendations:**

* **Replicate High-Efficiency Models:** Algeria’s universal coverage, Cabo Verde’s strong primary care
* **Invest in Prevention:** Immunization, maternal care, disease surveillance
* **Reduce OOP:** Move toward universal health coverage (UHC)
* **Improve Data Quality:** Strengthen national health accounts for better tracking
* **Adopt Weighted Mean 66POP:** For regional reporting consistency

**7. Limitations**

* Model explains only 24.3% of variation — many non-spending factors (governance, equity, disease burden) excluded
* Government spending’s negative coefficient may reflect reverse causality
* Missing data for some countries, especially pre-2010
* No imputation — results approximate, not definitive
* Linear model assumes no interaction effects; causality cannot be confirmed

**8. References & Acknowledgements**

**Data Sources:**

* World Health Organization. (2024). *Global Health Expenditure Database*. <https://apps.who.int/nha/database>
* World Bank. (2024). *Health, Nutrition and Population Statistics*. <https://databank.worldbank.org/source/health-nutrition-and-population-statistics>
* World Bank. (2024). *Aggregation Rules Documentation*. <https://databank.worldbank.org/source/health-nutrition-and-population-statistics?spm=a2ty_o01.29997173.0.0.56f65171sgIC6Q>

**Tools:** Python (pandas, scikit-learn, matplotlib, seaborn)

**Team Roles:**

* Jaji Jimoh Oludare — Data Wrangler
* Chidewu Tatenda — EDA Lead
* Rediet Tedla — Modeling Lead
* Daniel Cole — Visualizer
* Kingsley Njoku — Storyteller & Presenter

**Presentation Video link**:

<https://drive.google.com/file/d/16m9Uzl0hKniq-gtd2yfOtYZjogbCOCql/view?usp=sharing>

**github link :** [jaji-capstone/project\_AfriHealth: Capstone project analyzing healthcare spending impact in Africa and how to optimize Health Investment Strategies Across Africa](https://github.com/jaji-capstone/project_AfriHealth)

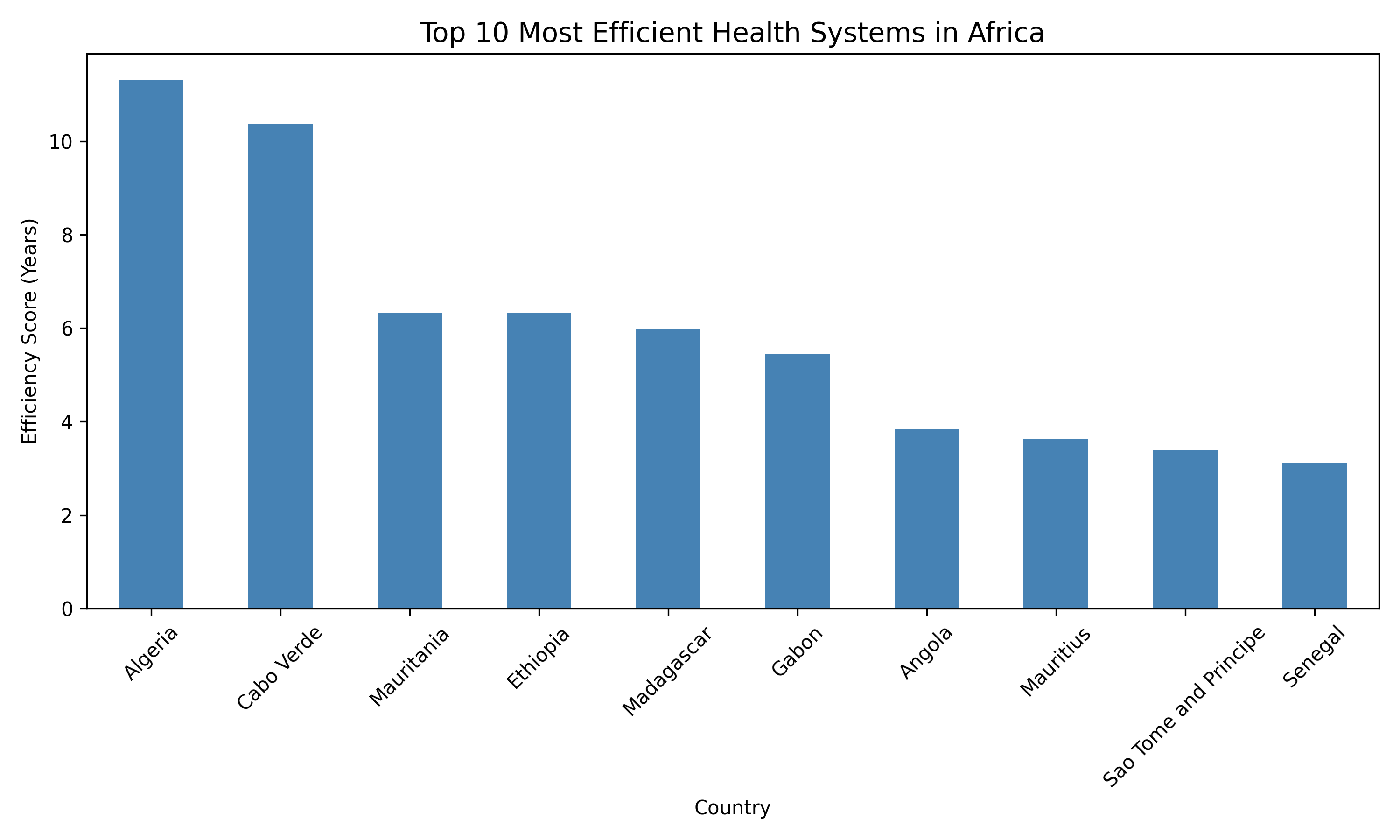
**APPENDIX: IMPORTANT TABLES AND GRAPHS**

***Fig: 1* -> 📊 Major Trends & Visuals - Top 10 Most Efficient Health Systems**

*(Efficiency Score = Actual − Predicted Life Expectancy)*

|  |  |  |
| --- | --- | --- |
| **Rank** | **Country** | **Efficiency Score  (Years)** |
| 1 | Algeria | +11.31 |
| 2 | Cabo Verde | +10.36 |
| 3 | Mauritania | +6.33 |
| 4 | Ethiopia | +6.32 |
| 5 | Madagascar | +5.99 |
| 6 | Gabon | +5.44 |
| 7 | Angola | +3.84 |
| 8 | Mauritius | +3.64 |
| 9 | São Tome and Principe | +3.39 |
| 10 | Senegal | +3.11 |

***Fig: 2 ->*  📊 Top 10 Most Efficient Health Systems**

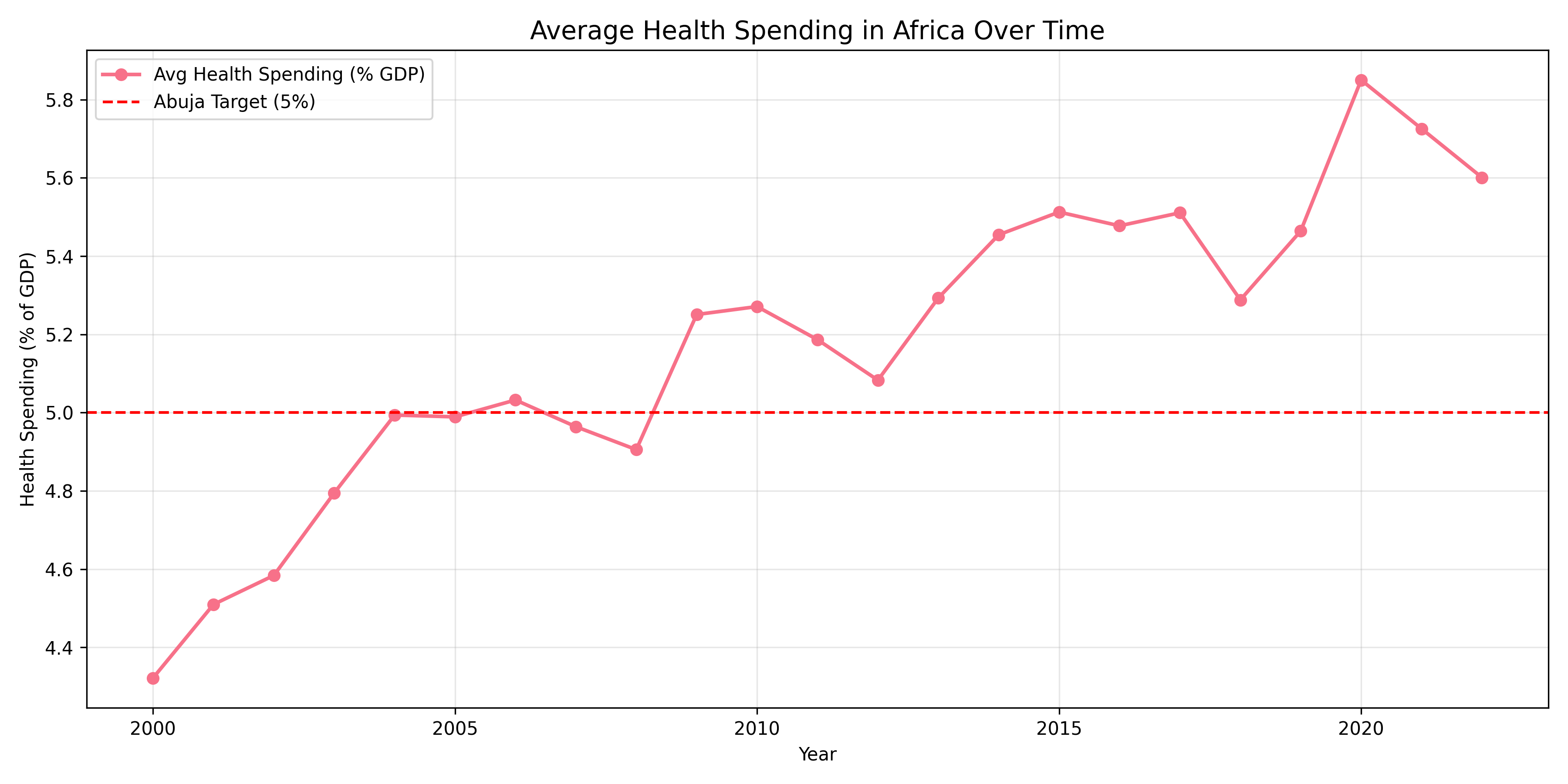


***Fig: 3 ->*  📊 Regional Trends (2020, Weighted Mean 66POP**

|  |  |  |
| --- | --- | --- |
| **Region** | **Life Expectancy** | **CHE (% GDP)** |
| North Africa | 73.3 | 5.60% |
| Southern Africa | 64.7 | 8.40% |
| East Africa | 64.8 | 4.10% |
| West Africa | 55.4 | 3.60% |
| Central Africa | 62.1 | 4.20% |

📉 *Southern Africa spends the most but underperforms — likely due to HIV and inequality.*

***Fig: 4 ->*  📊 Average Health Spending in Africa Over Time**



***Fig: 5 ->*  📊 Regional Trends (2020, Weighted Mean 66POP)**

|  |  |  |
| --- | --- | --- |
| **Region** | **Life Expectancy** | **CHE (% GDP)** |
| North Africa | 73.3 | 5.60% |
| Southern Africa | 64.7 | 8.40% |
| East Africa | 64.8 | 4.10% |
| West Africa | 55.4 | 3.60% |
| Central Africa | 62.1 | 4.20% |

***Fig: 6 ->*  📊 Government Health Spending Vs Life Expectancy**

