Project Title:

Analyzing Healthcare Access Inequality in Kenya

Author:

Collins Kipkirui Rutto

Date:

July 2025

---

1. Project Overview

This project explores the distribution of healthcare facilities across Kenya and examines disparities in access by county. Using visualization techniques in Tableau and supporting data analysis in Excel, the goal is to uncover insights into which counties may be underserved and the extent of inequality in healthcare infrastructure.

---

2. Dataset Description

- **Primary Dataset:** Kenya Master Health Facility List

- **Source:** Ministry of Health, Kenya

- **Additional Data:** County-level population estimates (manually integrated)

- **Variables Used:**

    o   County

    o   Facility Count

    o   Population

---

3. Visualizations and Explanations

a. Bar Chart – Facility Count by County

This chart shows the total number of healthcare facilities in each county. It highlights which regions have more facilities overall, with Nairobi and Kiambu among the top.

Sheet 1
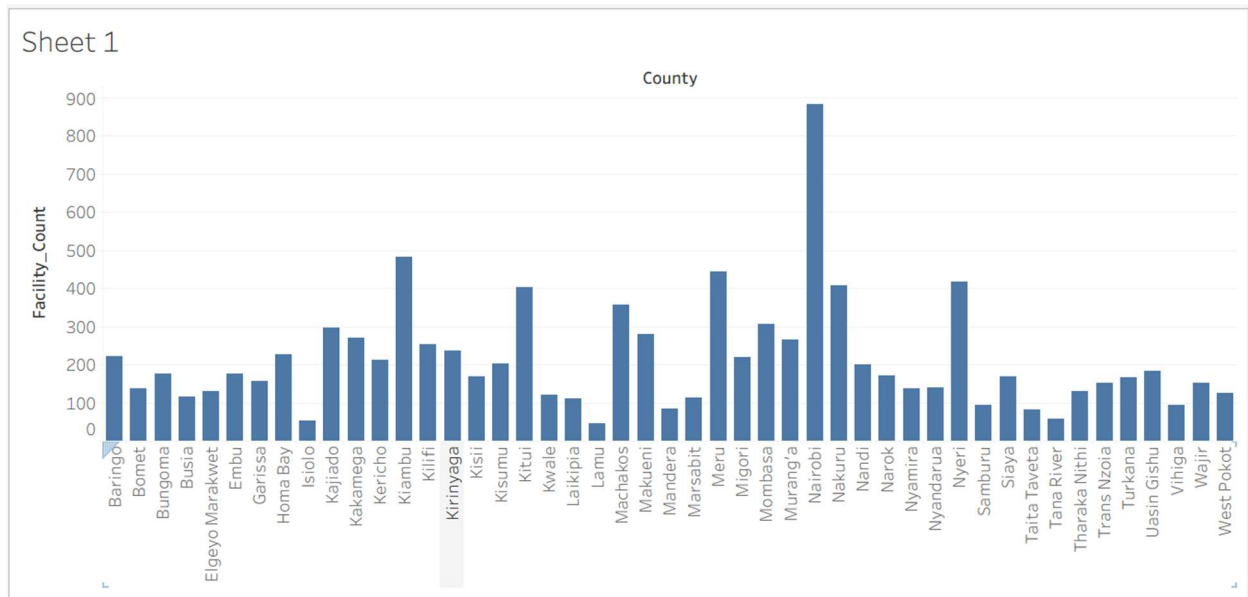
b. Map – Healthcare Facility Distribution in Kenya

The map visualizes the geographic spread of health facilities. Densely populated counties like Nairobi and Kisumu show high concentrations, while vast areas in the north show sparse coverage.


Sheet 3
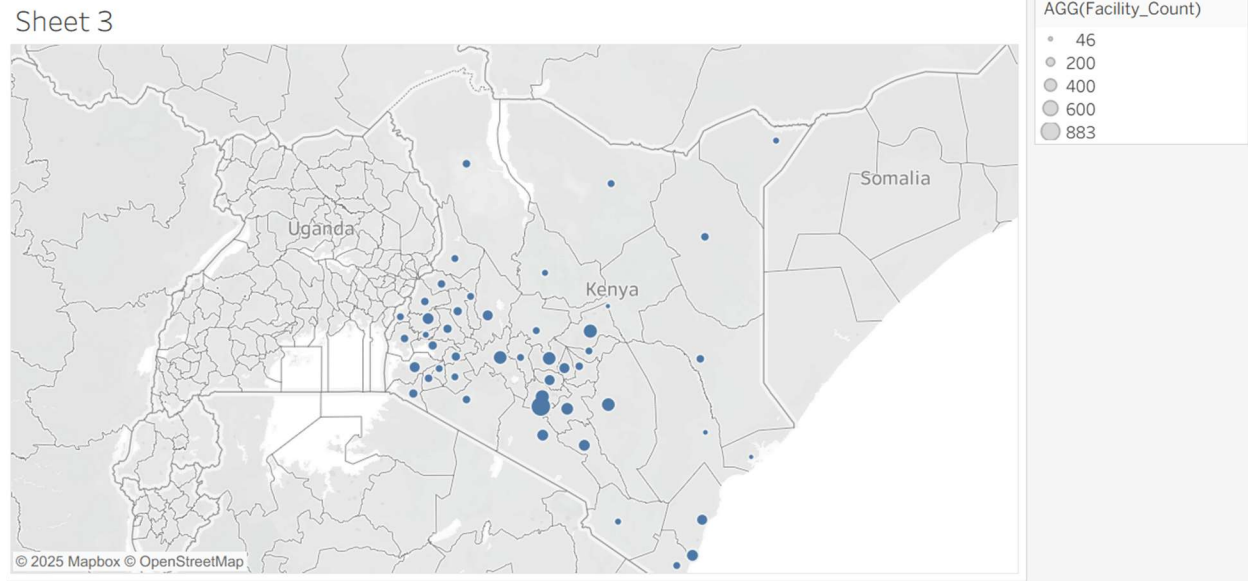
c. Scatter Plot – Population vs Facility Count

This plot reveals the relationship between population and facility count by county. A trendline is added to identify outliers—counties with disproportionately low or high facility numbers relative to population.
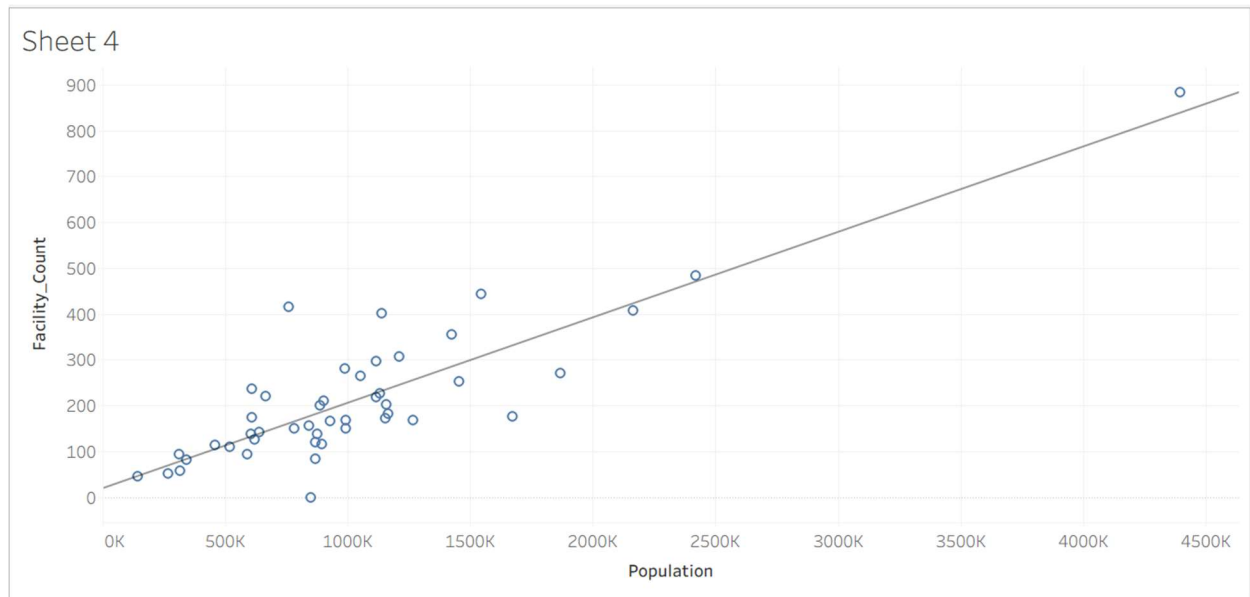
d. Choropleth Map – Population per Facility by County

This map highlights the ratio of population to healthcare facility per county. High ratios (darker shades) suggest potential strain on available services.
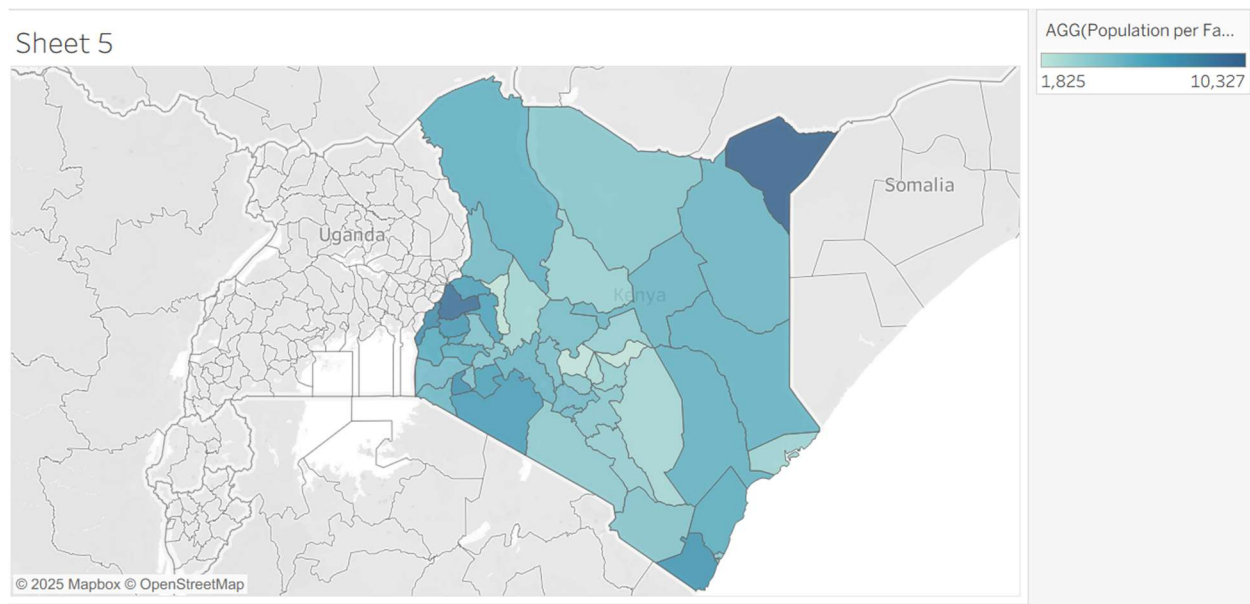

e. Bar Chart – Top 10 Counties by Population per Facility

A focused bar chart ranking the ten counties with the highest population per facility. These are the counties likely most in need of additional infrastructure.
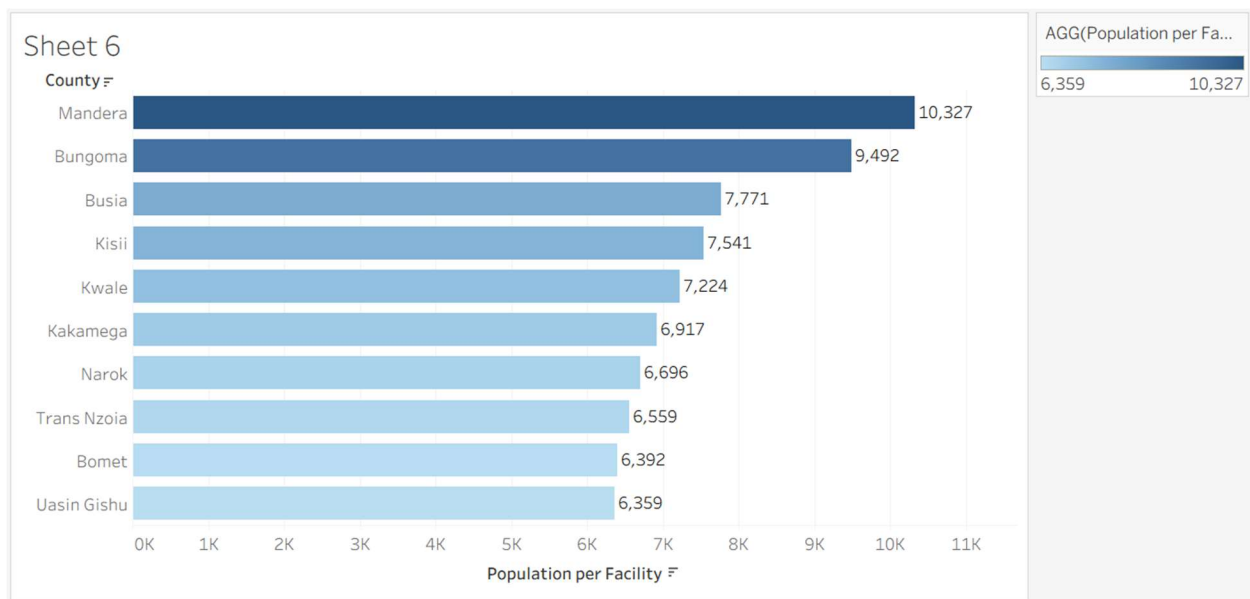
f. Dual Axis Trendline – County vs Population and Facility Count

This chart compares both total population and facility count side by side on dual axes for each county. It helps visualize how well healthcare infrastructure scales with population.
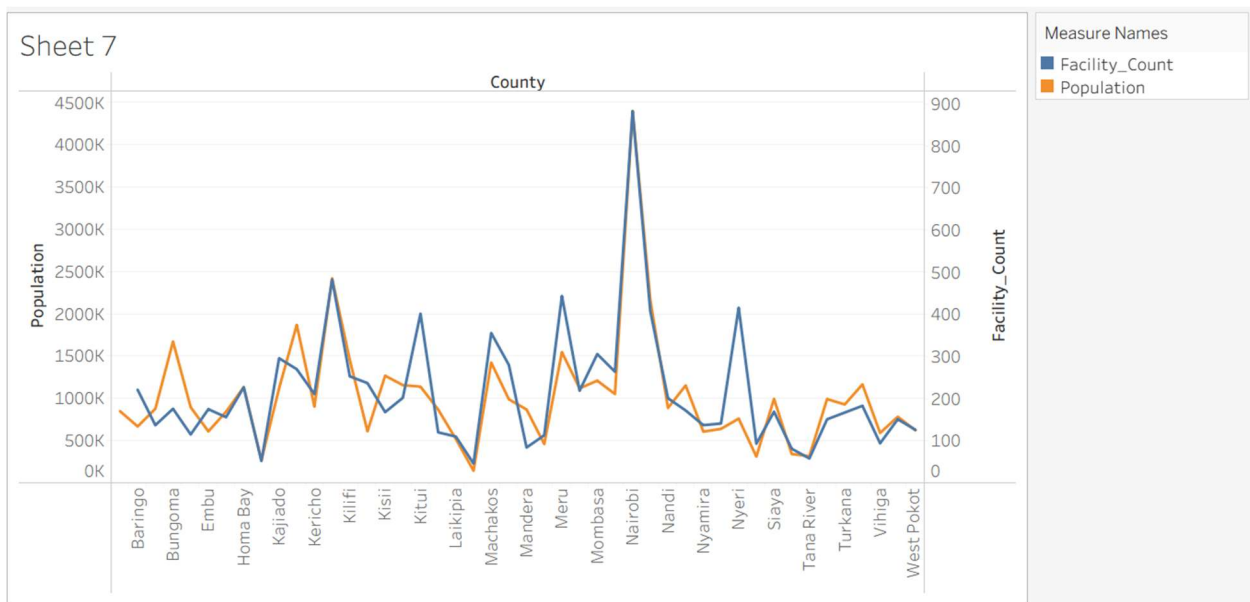
4. Regression Analysis

A linear regression analysis was performed to examine whether county population can predict the number of healthcare facilities across Kenya. The results showed that the model is statistically significant, with an **R-squared value of 0.76**, indicating that **76% of the variation in**

healthcare facility counts can be explained by population differences among counties. The regression equation derived from the analysis is:

Facility_Count = 27.41 + 0.000183 × Population

This suggests that, on average, for every additional person in a county, the number of healthcare facilities increases by approximately **0.000183** — or about **1 facility for every 5,460 people**. The **population coefficient is highly significant (p < 0.0000000000000016)**, confirming a strong positive relationship. Although the intercept (27.41) was not statistically significant (p = 0.15), it does not impact the overall validity of the model. The overall model F-statistic (F = 142.2) further confirms the model's strength.

These results indicate that population size is a powerful predictor of healthcare facility distribution in Kenya and should be considered in planning and policy decisions.

	A	B	C	D	E	F	G	H	I	J
3	*Regression Statistics*									
4	Multiple R	0.87155935								
5	R Square	0.7596157								
6	Adjusted R Square	0.75427383								
7	Standard Error	72.3226766								
8	Observations	47								
9										
10	ANOVA									
11		*df*	*SS*	*MS*	*F*	*Significance F*				
12	Regression	1	743788.2848	743788.3	142.2002	1.5854E-15				
13	Residual	45	235375.6301	5230.57						
14	Total	46	979163.9149							
15										
16		*Coefficients*	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*	
17	Intercept	27.4144802	18.80445242	1.457872	0.15182	-10.459631	65.2885916	-10.4596311	65.2885916	
18	Population	0.00018343	1.53819E-05	11.92477	1.59E-15	0.00015245	0.00021441	0.00015245	0.00021441	
19										

5. Conclusion

The regression analysis reveals a strong and statistically significant relationship between population and the number of healthcare facilities across counties in Kenya. With an R^2 of 0.76, the model explains approximately 76% of the variation in facility count using population alone. The p-value for the population variable is extremely small (1.59×10^{-15}), confirming its predictive power. This suggests that counties with larger populations tend to have more healthcare facilities. While the intercept is not significant and may not hold practical meaning in this context, the overall model confirms that population is a key driver in the distribution of health infrastructure. For deeper insights, future models could include additional variables like land area, poverty index, or rural/urban classification.