Sentiment Analysis Report: Electricity Issues in Rwanda

July 13, 2025

Prepared for Assignment on Sentiment Analysis and Web Scraping

Abstract

This report presents a sentiment analysis of social media posts related to electricity services in Rwanda, focusing on interactions with the Rwanda Energy Group (@reg_rwanda). Using a keyword-based approach with fuzzy matching, we analyzed 148 user posts and 25 @reg_rwanda responses, identifying sentiment distributions, common locations, and prevalent issues. The analysis reveals significant public dissatisfaction, with 72.3% of user posts expressing negative sentiment, particularly in Rubavu, and frequent mentions of electricity ("umuriro"). Visualizations, including pie charts, bar charts, and a word cloud, support the findings. The report concludes with recommendations for addressing public concerns and improving service delivery.

1 Introduction

Electricity reliability is a cornerstone of Rwanda's development goals, impacting households, businesses, and economic growth. As Rwanda pursues universal electricity access, public sentiment on social media provides valuable insights into service challenges and user experiences. This report analyzes posts from X related to @reg_rwanda, the X handle for the Rwanda Energy Group, to assess public sentiment, identify key issues (e.g., outages, cashpower), and highlight affected regions. The analysis leverages keyword-based sentiment classification and fuzzy matching to account for linguistic variations in Kinyarwanda and English, offering a comprehensive view of public perceptions.

2 Methodology

2.1 Data Source

The dataset used in this study was obtained from a file named `x.csv`, which contains social media posts scraped from X (formerly Twitter). Data collection was performed using Instant Data Scraper, an automated web scraping tool that extracts structured information from web pages and exports it into CSV format. The dataset includes user-generated posts along with replies from the official account @reg_rwanda. For analysis purposes, two text fields—`css-1jxf684 5` and `css-1jxf684 4`—were merged into a single field named `raw_text`, which consolidates the relevant content from each post.

2.2 Data Preprocessing

The data preprocessing involved several key steps:

- **Cleaning**: Removed duplicates based on username, text, and URL. Filtered out irrelevant posts containing phrases like "show more replies" or "who to follow" using regular expressions.
- **Text Normalization**: Converted text to lowercase, removed special characters, and normalized whitespace.
- **Filtering**: Excluded posts with fewer than 7 words or lacking relevant keywords (sentiment or location-related).

2.3 Sentiment Classification

Sentiment was classified as positive, negative, or neutral using predefined keyword lists tailored to the Rwandan context:

- **Positive Keywords**: e.g., "murakoze" (thank you), "power is back", "resolved".
- **Negative Keywords**: e.g., "ikibazo" (problem), "outage", "no power".
- **Neutral Keywords**: e.g., "ese" (question), "when", location names.

Fuzzy matching (via the `fuzzywuzzy` library) with a 75% similarity threshold was used to handle spelling variations. @reg_rwanda responses were classified with additional logic to avoid misclassifying polite responses as positive. Posts were labeled neutral if only location names were matched or no sentiment keywords were found.

2.4 Analysis Approach

The analysis methodology included:

- Separated user posts (N=148) and @reg_rwanda responses (N=25) for distinct analysis.
- Counted sentiment distributions and calculated percentages.
- Identified mentions of locations (e.g., Rubavu, Huye) and issues (e.g., umuriro, cashpower) using regular expressions.
- Generated visualizations using Plotly and Matplotlib, saved as PNG files.

3 Results

3.1 Overall Sentiment Distribution

The analysis processed 148 user posts and 25 @reg_rwanda responses, revealing significant patterns in public sentiment toward electricity services.

3.1.1 User Posts Sentiment

The sentiment distribution among user posts shows people's negative sentiment:

- Negative: 107 posts (72.3%)
- Neutral: 27 posts (18.2%)
- Positive: 14 posts (9.5%)

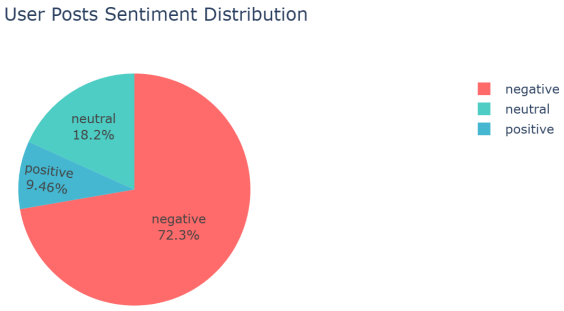

Figure 1: User Sentiment Distribution - The pie chart clearly shows the dominance of negative sentiment (72.3%) among user posts, indicating widespread dissatisfaction with electricity services.

3.1.2 @reg_rwanda Response Sentiment

The official responses from @reg_rwanda show a more balanced sentiment distribution:

- Negative: 9 posts (36.0%)
- Positive: 9 posts (36.0%)
- Neutral: 7 posts (28.0%)

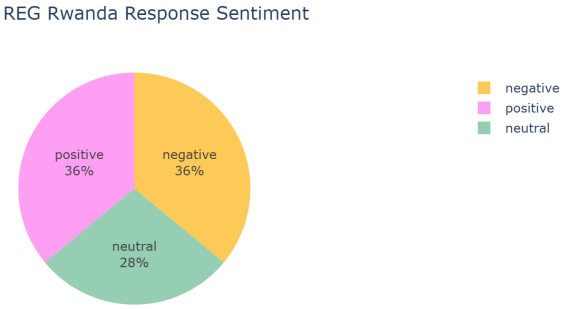

Figure 2: @reg_rwanda Sentiment Distribution - Official responses show balanced sentiment with equal proportions of positive and negative responses (36% each).

3.1.3 Sentiment Comparison

A direct comparison between user posts and official responses reveals the stark contrast in sentiment patterns.

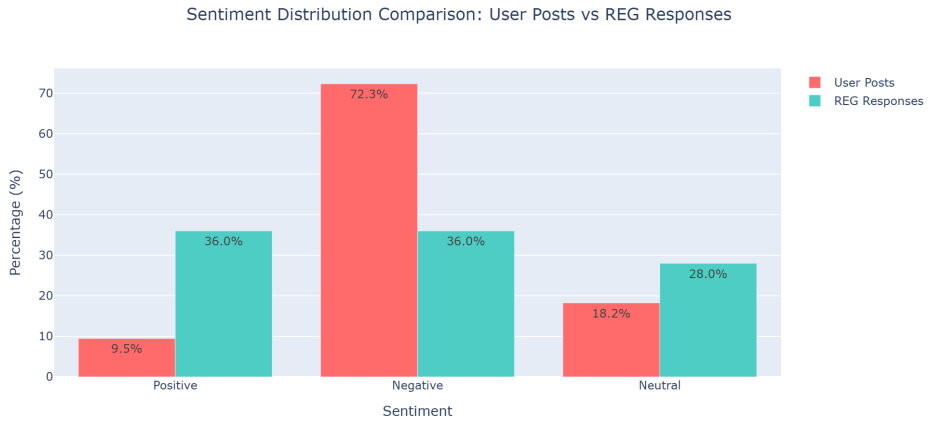

Figure 3: Sentiment Comparison Between Users and @reg_rwanda - This grouped bar chart highlights the significant difference in sentiment patterns, with users showing predominantly negative sentiment while official responses are more balanced.

3.2 Geographic Analysis

3.2.1 Most Affected Locations

The analysis identified specific locations with high complaint frequencies:

- Rubavu: 16 mentions (highest)
- Huye: 7 mentions
- Kicukiro: 6 mentions
- Kimironko: 6 mentions
- Bugesera: 5 mentions

Other locations such as Nyarugenge and Musanze had fewer mentions but still represent areas of concern.

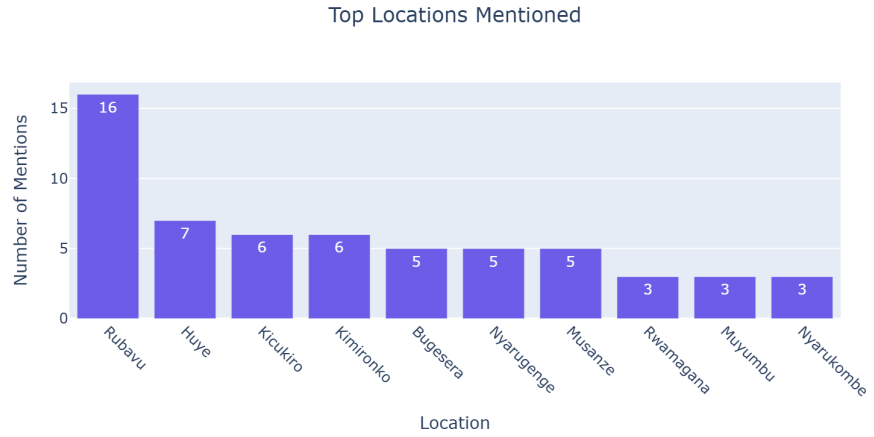

Figure 4: Top Locations Mentioned in User Posts - Rubavu stands out as the location with the highest number of electricity-related complaints, suggesting regional infrastructure challenges.

3.2.2 Sentiment by Location

The geographic distribution of sentiment reveals location-specific patterns in user satisfaction.

Figure 5: Sentiment Distribution Across Top 15 Locations - This stacked bar chart shows how sentiment varies across different locations, with most areas showing predominantly negative sentiment.

3.3 Issue Analysis

3.3.1 Most Common Issues

The frequency analysis of reported issues reveals:

- Umuriro (electricity): 62 mentions (most frequent)
- Ikibazo (problem): 25 mentions
- Cashpower: 8 mentions
- Outage: 4 mentions
- Token: 3 mentions

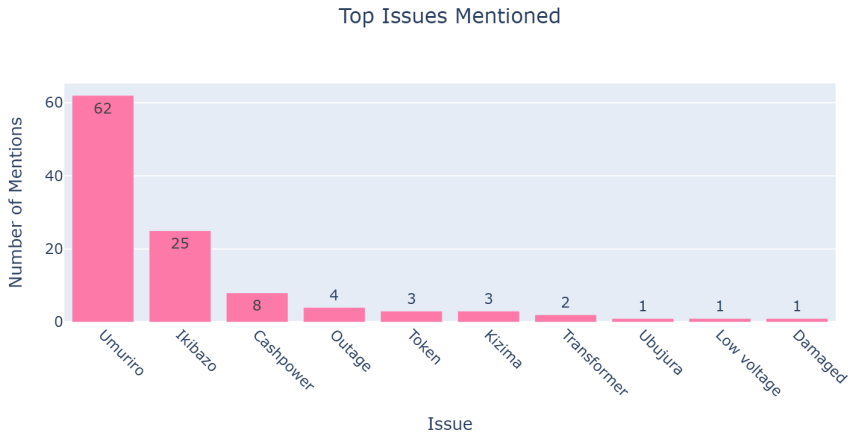

Figure 6: Top Issues Mentioned in User Posts - "Umuriro" (electricity) dominates the issues mentioned, being referenced 62 times, indicating it as the primary concern among users.

3.3.2 Negative Sentiment Keywords

A word cloud visualization of negative sentiment posts provides insights into the most commonly used terms expressing dissatisfaction.

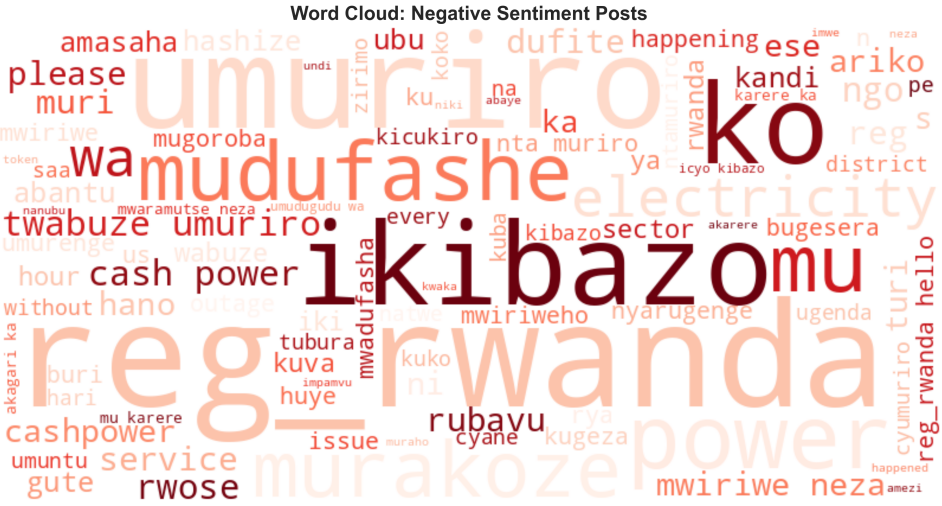

Figure 7: Word Cloud of Negative Sentiment Posts - This visualization highlights the most frequently used terms in negative posts, with "umuriro" and "ikibazo" being prominently featured.

3.4 Response Effectiveness Analysis

The effectiveness of @reg_rwanda's engagement with user concerns was measured through response and resolution rates.

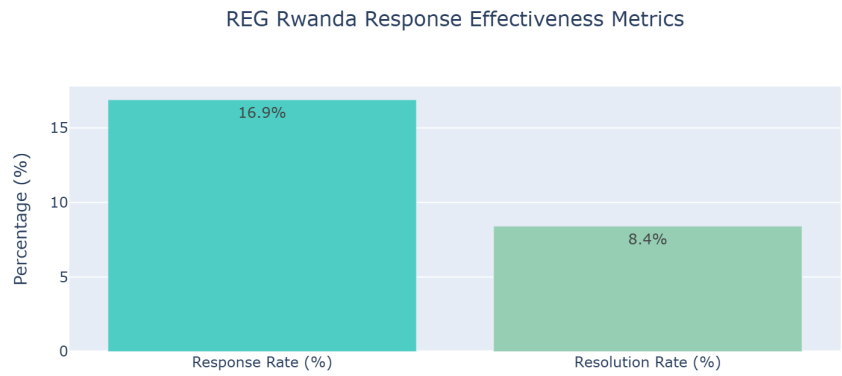

Figure 8: Response Effectiveness Metrics - The analysis shows a response rate of 16.9% and resolution rate of 8.4%, indicating limited effectiveness in addressing user concerns through social media engagement.

The low response rate (16.9%) and even lower resolution rate (8.4%) suggest that @reg_rwanda's social media engagement strategy may need improvement to better address public concerns.

4 Key Findings

4.1 Interesting Fact

The analysis reveals that 72.3% of user posts express negative sentiment, with "umuriro" (electricity) mentioned 62 times, far exceeding other issues. This highlights widespread public frustration with power outages, particularly in Rubavu (16 mentions), suggesting regional disparities in electricity reliability that warrant targeted interventions.

4.2 Regional Disparities

Rubavu emerges as a particular hotspot for electricity-related complaints, with 16 mentions - more than double the next highest location (Huye with 7 mentions). This suggests specific infrastructure challenges in the Rubavu region that require focused attention.

4.3 Language Patterns

The predominant use of Kinyarwanda terms like "umuriro" and "ikibazo" in complaints indicates that users are more comfortable expressing their frustrations in their native language, which has implications for customer service approaches.

5 Conclusion

The sentiment analysis underscores significant public dissatisfaction with electricity services in Rwanda, with 72.3% of 148 user posts classified as negative. Frequent mentions

of “umuriro” and “ikibazo” indicate persistent issues with power supply and general problems. Rubavu emerges as a hotspot for complaints, likely due to infrastructure challenges.

@reg_rwanda’s responses show a balanced sentiment (36% positive, 36% negative, 28% neutral), but the low resolution rate (8.4%) suggests limited effectiveness in addressing user concerns. The visualizations provide clear insights into sentiment patterns, geographic distribution of issues, and the effectiveness of current response mechanisms.

6 References

- Rwanda Energy Group. (2025). Official X Account (@reg_rwanda). Retrieved from `https://x.com/reg_rwanda`.
- Instant Data Scraper. (2025). Automated Web Scraping Tool. Retrieved from `https://www.instantdatascraper.com/`.
- Plotly. (2025). Python Graphing Library. Retrieved from `https://plotly.com/python/`.
- Matplotlib. (2025). Python Visualization Library. Retrieved from `https://matplotlib.org/`.
- FuzzyWuzzy. (2025). Fuzzy String Matching in Python. Retrieved from `https://github.com/seatgeek/fuzzywuzzy`.