pmachieng / **P_3_project**

<> Code    Issues    Pull requests    Actions    Projects    Wiki    Security    Insights    Settings

## P_3_project  Public

**1 Branch**    **0 Tags**    Go to file    t    Go to file    +    Add file    About    Code    ...

| | | | |
|---|---|---|---|
| pmachieng | final_commit | 9437da8 · 11 minutes ago | |
| .ipynb_checkpoints | final_commit | 11 minutes ago | |
| Gaduel_Phase 3-Pr... | updated working progress | 1 hour ago | |
| README.md | finish | 23 minutes ago | |
| Sudan_1997-2024... | fist commit | 2 days ago | |
| Untitled.ipynb | final_commit | 11 minutes ago | |
| index.ipynb.ipynb | final_commit | 11 minutes ago | |
| notebook.pdf.pdf | final_commit | 11 minutes ago | |
| presentation.pdf.pdf | final_commit | 11 minutes ago | |

### About

end of phaseII

📖 Readme

〰 Activity

☆ 0 stars

👁 1 watching

⑂ 0 forks

### Releases

No releases published
Create a new release

### Packages

No packages published
Publish your first package

### README

# P_3_project

**1: Research Objectives**

1.1: Project objectives

*To quantify the prevalence of political violence events in Africa and Sudan in particular over the past decade.*

*To analyze the trend in the number of fatalities resulting from political violence events in Sudan.*

*To investigate the correlation between the frequency of political violence events and the number of fatalities.*

*To identify the types of political violence that result in the highest number of fatalities.*

*To provide policy recommendations based on the findings to reduce fatalities associated with political violence in Sudan.*

1.2: Hypothesis

Null Hypothesis ($H_0$)

*"There is no correlation between the frequency of political violence events and the number of fatalities in Sudan."*

Alternative Hypothesis ($H_1$)

*"There is a positive correlation between the frequency of political violence events and the number of fatalities in Sudan."*

1.3: Data Collection

## Languages

● **Jupyter Notebook** 100.0%

*Data on political violence events and fatalities was collected from*
[*https://acleddata.com/curated-data-files/#regional*](https://acleddata.com/curated-data-files/#regional)

Data Analysis

*Data will be analysed using statistical methods to to investigate the correlation coefficients and strong linear relationship using regression analysis.The results will be tested using statsmodels and scikitlearn.*

*Interpretation: Determine if the results support the hypothesis and assess the strength and significance of the correlation.*

1.4: Description of the ACLED Data Columns at a Glance

**Column name Column description Values**

**event_id_cnty**: A unique alphanumeric event identifier by number and country acronym. This identifier remains constant even when the event details are updated. E.g. ETH9766

**event_date**: The date on which the event took place. Recorded as Year-Month-Day. E.g. 2023-02-16 year The year in which the event took place. E.g. 2018

**time_precision**: A numeric code between 1 and 3 indicating the level of precision of the date recorded for the event. The higher the number, the lower the precision. 1, 2, or 3; with 1 being the most precise.

**disorder_type** The disorder category an event belongs to. Political violence, Demonstrations, or Strategic developments.

**event_type** The type of event; further specifies the nature of the event. E.g. Battles For the full list of ACLED event types, see the ACLED Event Types table.

**sub_event_type**: A subcategory of the event type. E.g. Armed clash

**actor1**: One of two main actors involved in the event (does not necessarily indicate the aggressor). E.g. Rioters (Papua New Guinea)

**assoc_actor_1**: Actor(s) involved in the event alongside 'Actor 1' or actor designations that further identify 'Actor 1'. E.g. Labor Group (Spain); Women (Spain)

**inter1**: A text value indicating the type of 'Actor 1' (for more, see the section Actor Names, Types, and 'Inter' Codes). E.g. Rebel group

**actor2**: One of two main actors involved in the event (does not necessarily indicate the target or victim). E.g. Civilians (Kenya)

**assoc_actor_2**: Actor(s) involved in the event alongside 'Actor 2' or actor designation further identifying 'Actor 2'. E.g. Labor Group (Spain); Women (Spain)

**inter2**: A text value indicating the type of 'Actor 2' (for more, see the section Actor Names, Types, and 'Inter' Codes). E.g. State forces

**interaction**: A text value based on a combination of 'Inter 1' and 'Inter 2' indicating the two actor types interacting in the event (for more, see the section Actor Names, Types, and 'Inter' Codes). E.g. Rebel group – Civilians

**civilian_targeting**: This column indicates whether the event involved civilian targeting. Either 'Civilians targeted' or blank.

**iso**: A unique three-digit numeric code assigned to each country or territory according to ISO 3166. E.g. 231 for Ethiopia

**region**: The region of the world where the event took place. E.g. Eastern Africa

**country**: The country or territory in which the event took place. E.g. Ethiopia

**admin1**: The largest sub-national administrative region in which the event took place. E.g. Oromia

**admin2**: The second largest sub-national administrative region in which the event took place. E.g. Arsi

**admin3**: The third largest sub-national administrative region in which the event took place. E.g. Merti

**location**: The name of the location at which the event took place. E.g. Abomsa

**latitude**: The latitude of the location in four decimal degrees notation (EPSG:4326). E.g. 8.5907

**longitude**: The longitude of the location in four decimal degrees notation (EPSG:4326). E.g. 39.8588

**geo_precision**: A numeric code between 1 and 3 indicating the level of certainty of the location recorded for the event. The higher the number, the lower the precision. 1, 2, or 3; with 1 being the most precise.

**source**: The sources used to record the event. Separated by a semicolon. E.g. Ansar Allah; Yemen Data Project

**source_ scale**: An indication of the geographic closeness of the used sources to the event (for more, see the section Source Scale). E.g. Local partner-National

**notes** A short description of the event. E.g. On 16 February 2023, OLF-Shane abducted an unidentified number of civilians after stopping a vehicle in an area near Abomsa (Merti, Arsi, Oromia). The abductees were traveling from Adama to Abomsa, Arsi. fatalities The number of reported fatalities arising from an event. When there are conflicting reports, the most conservative estimate is recorded. E.g. 3. No information on fatalities is recorded as 0 reported fatalities.

**tags** Additional structured information about the event. Separated by a semicolon. E.g. women targeted: politicians; sexual violence

**timestamp** An automatically generated Unix timestamp that represents the exact date and time an event was uploaded to the ACLED API. E.g. 1676909320

1: INTRODUCTION & BUSINESS UINDERSTANDING • Sudan has been a violent hotspot in Africa due to political instability since 1997.

• With the increasing level of violence, there is need to understand, the drivers of violence in Sudan, the role of various actors in violent conflict.

• This project uses the ACLED Data to analyze Sudan violence and predict the impact on fatalities using regression model and random forest.

• Peacebuilding interventions require thorough understanding of local context such as influence of actors, affected categories of people.

• The analysis from this project will advise peacebuilding partners on how to address the underlying causes of violence in Sudan.

• It will draw conclusions aimed at choosing appropriate peace building program such as strengthening rule of law and security institutions

2. EDA & KEY FINDINGS

This section involves inferential analysis based on key variables

It will answer the following questions:

What is the most common type of violence in Sudan?

• From the visuals, battles constitute the largest proportion of violent incidents with more than 10,000 counts of fatalities per year.

• Riots are the least type of violent incidents in Sudan

• Most of the battles are basically armed clashes.

Which actors are responsible for most of violent counts in Sudan?

• The state forces are the primary perpetrators in Sudan conflict linked with over 38% of violent counts

• Other actors are identity militia and protestors

Where is violence concentrated in Sudan; and what are the top 10 regions in Sudan with the highest count of fatalities? • North Darfur region is the most affected hotspot in Sudan with over 16000 fatalities

• Upper Nile is the second most affected region by war

• West Darfur is the least affected region with low fatalities count.

When did Sudan record the highest number of fatalities? • The highest fatalities count in Sudan was recorded in 2023, and 2013 with over 10,000 deaths

• The quietest period was between 2005 to 2010 with lowest record of violent deaths

• There was an upward trend in violence between 1999 to 2002, 2010 - 2013 , and 2020 - 2023.

• There was a sharp decline of fatalities between 2004 -2010, 2014 -2022. 3: MODELLING & EVALUATION Linear regression model and Random Forest regressor were used to evaluate the metrics for my analysis. In both models, the R2_score was used to evaluate the effectiveness of my prediction.

• The $R^2$ (R-squared) score is a measure of how well the independent variables (features) in my model explain the variability of the target variable ("Fatalities").

• The Linear regression model explain only 43.47% of the variance in the training dataset's target variable ("Fatalities")

• The explanation is moderate suggesting some relationships are not clearly explained but has room for improvement

• The Random Forest model explains 86.18 per cent of the variance in fatalities within training dataset that indicate the model fits well with training dataset.

• The model fails to generalize to unseen data, which makes it unreliable for predicting "Fatalities".

4: CONCLUSIONS

• Armed clashes are the most violent events in Sudan

• The most violent period in Sudan is between 2012 – 2014, and 2022 – 2023

• Most of the violence in Sudan is concentrated in North Darfur, Upper Nile, Jonglei and South Darfur region

• The State Security Forces are the main perpetrators of violence in Sudan conflict.

5: RECOMMENDATIONS

• More peacebuilding efforts should be done in North Darfur, Upper Nile and Jonglei regions

• Peacebuilding effort in Sudan should be directed to rule of law and security