



EMORY
COLLEGE
OF ARTS AND
SCIENCES

Department of Quantitative
Theory and Methods



The Carter Center Capstone DataBlitz Presentation

Students: Shin Adachi, Josh Gelfond, Jack Liu, and Caris Weilenmann

Mentor: Dr. Mark Bell

April 22, 2022

Our Team: Students



Shin Adachi

Class of 2022

Applied Mathematics
and Statistics, BS



Josh Gelfond

Class of 2023

Quantitative Science, BS
with a concentration in
Economics



Jack Liu

Class of 2022

Quantitative Sciences, BS
with a concentration in
Economics



Caris Weilenmann

Class of 2022

Applied Mathematics
and Statistics, BS

The Department of Quantitative Theory and Methods, QTM, helps students pursue an undergraduate education at the intersection of data science and liberal arts. Students learn to collect, analyze and report data as well as how to think critically about data presented to them.

Our Team: Mentors



**Dr. Mark Bell,
Ph.D.**

QTM Faculty Lead
Balentine



**Dr. Zhiyun Gong,
Ph.D.**

QTM Director of
Undergraduate
Research



The Carter Center

Ben Spears

Senior Program Associate

Hampton Stall, M.D.P.

Senior Program Associate

The Carter Center Mission: The Carter Center seeks to prevent and resolve conflicts, enhance freedom and democracy, and improve health.

Sudanese Mission: Carter Center objectives in Sudan are to (a) support the Federal Ministry of Health in controlling, eliminating and eradicating neglected tropical diseases and (b) assist in resolving conflict and transitioning to democracy.

Background

Sudan Current Affairs

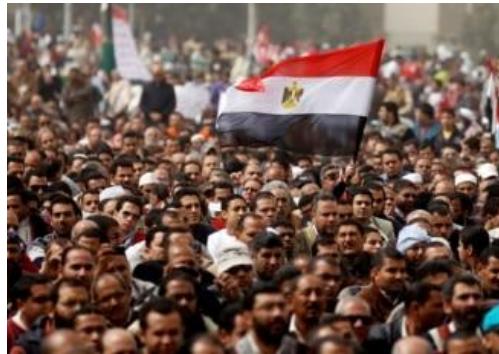
- Nationwide Protests Against Coup
 - Human rights abuses
 - Demonstrators killed
 - Alleged sexual assaults by security forces
- Economic Crisis
 - Currency devaluation (1USD = 640 SDG)
 - High commodity prices
- Increased Sanctions from US
 - Of note due to Sudan's recent removal from US State Sponsored Terror List under Trump Administration
- Food Crisis Looming (World Food Program)
 - Causes: War in Ukraine, economic crisis, poor harvests



||| Conflicts for Comparison

Egyptian Revolution, Arab Spring, January 25, 2011 - February 11, 2011

- Protests against poverty, government oppression
- Lead by youth organizations
- Resulted in President Mubārak stepping down, leaving military in charge
- Isolated range in ACLED: 9/1/2010 - 8/31/2011 (12 months)
 - 1,136 cases of political violence
 - 1,833 deaths



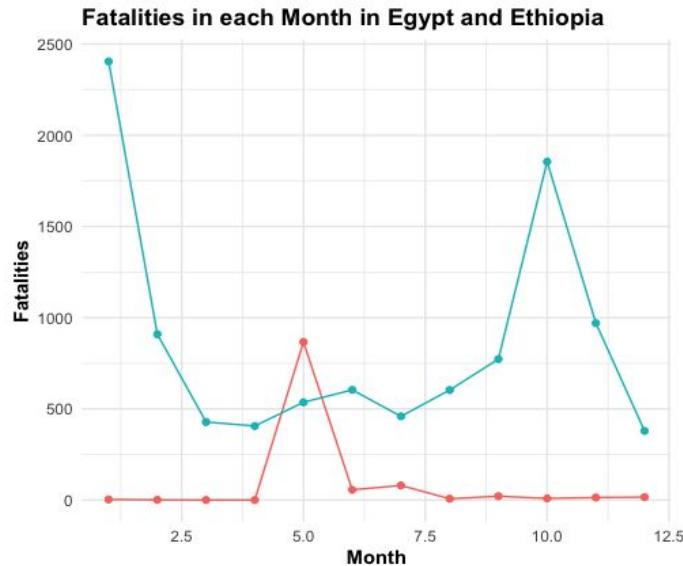
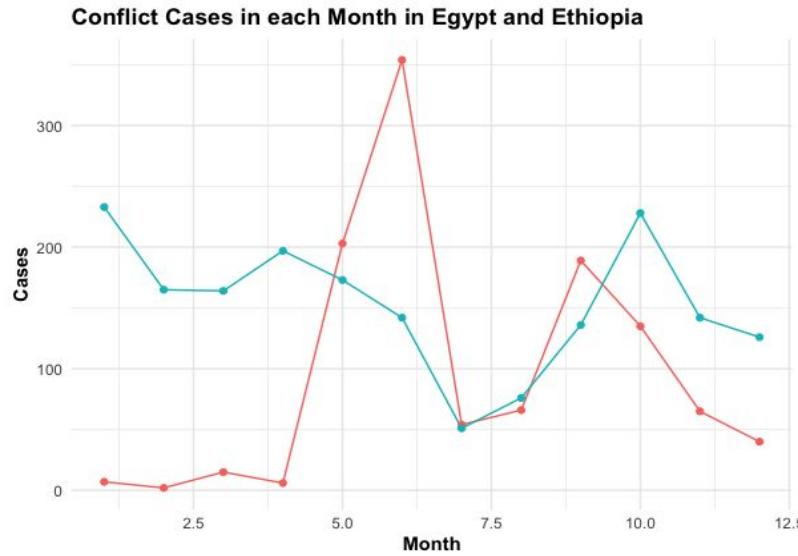
||| Conflicts for Comparison

Tigray War, Ethiopia, 2020-2021

- National government's military operation against regional ruling party
- Thousands of refugees have fled to Sudan
- Conflict is ongoing
- Isolated range in ACLED: 11/1/2020 - 10/31/2021 (12 months)
 - 1,074 cases of political violence
 - 10,327 deaths

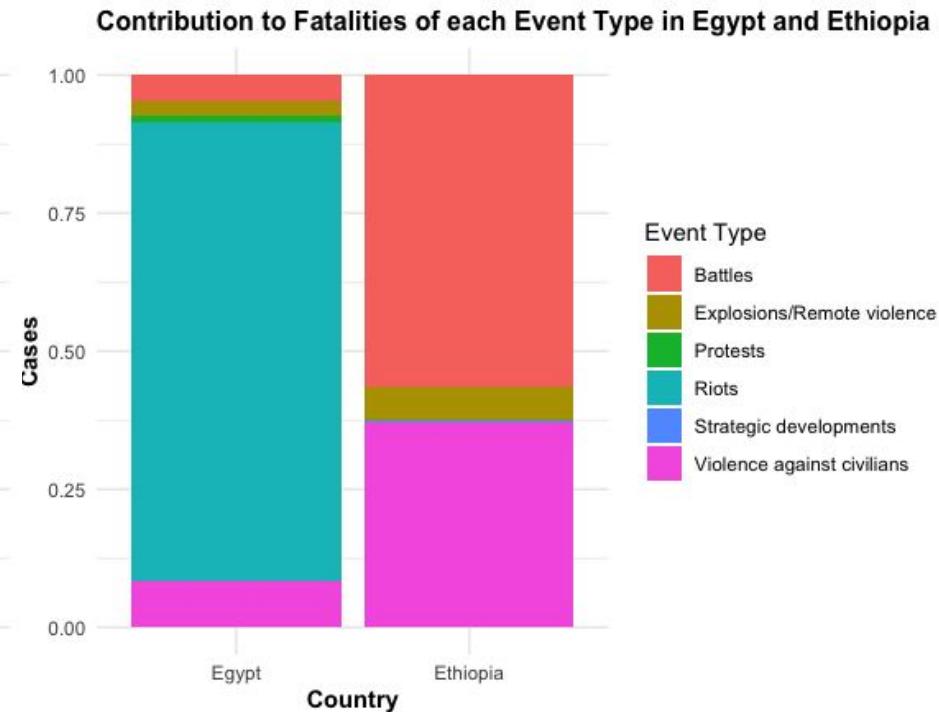
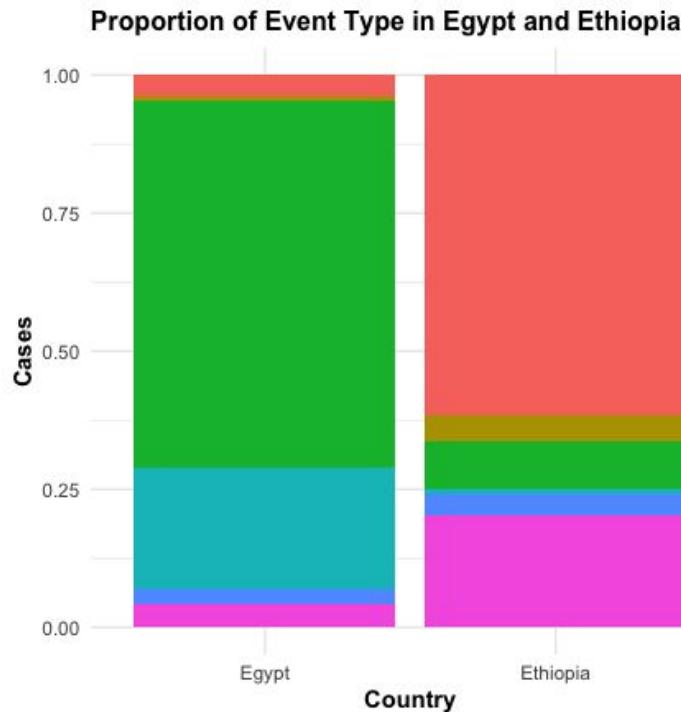


Conflicts and Fatalities - Egypt (2010) & Ethiopia (2020)



Country
— Egypt
— Ethiopia

Events Types - Egypt (2010) & Ethiopia (2020)



Event Type

- Battles
- Explosions/Remote violence
- Protests
- Riots
- Strategic developments
- Violence against civilians

Goals

- Create political violence risk assessment tool/model
 - Looking forward to elections in late 2022/early 2023
- Identify areas prone to conflict/conflict tracking
 - Potential for expansion beyond just South Darfur and Blue Nile regions
- Expand knowledge through data for conflict mitigation and prevention

NEMO Data Analysis

NEMO Sudan Youth Survey Data

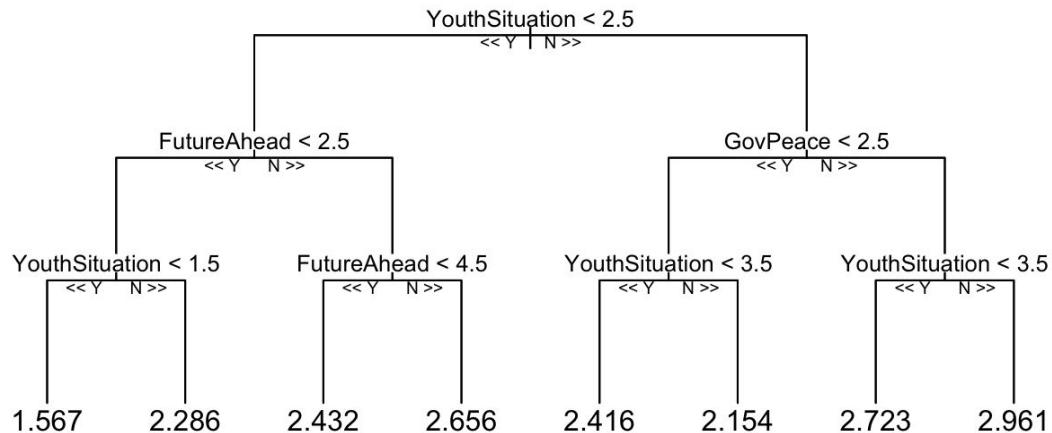
- 33 variables after de-identification and cleaning
 - 77 questions originally
 - 61 required
 - 16 nonrequired
 - 45 variables after de-identification
 - Removed 12 more from de-identified to isolate quantitative variables
- Answer types
 - Text (short or long)
 - Integer
 - Multiple-choice

Progress: Logistic Classification Model

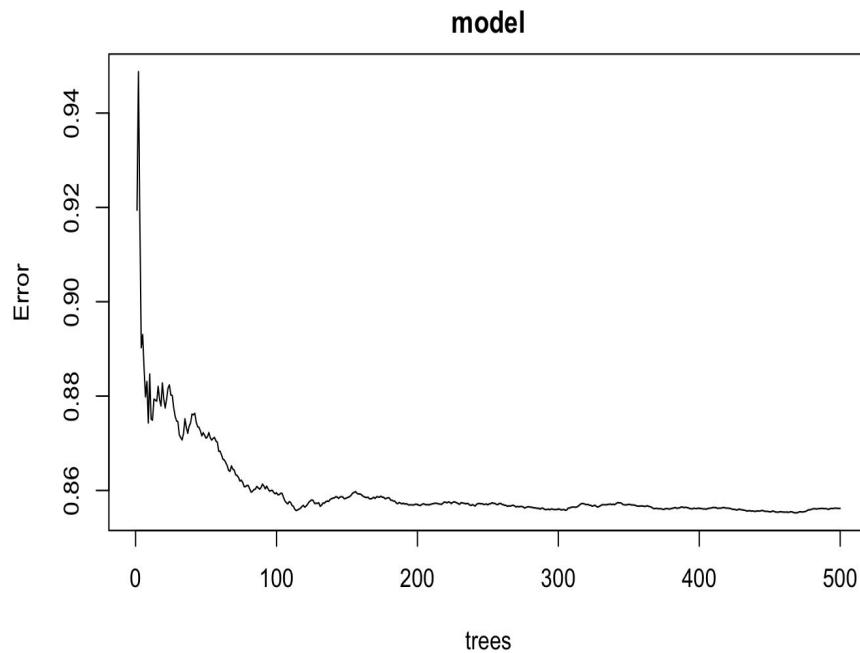
- Training (50%), Test (50%)
- Dependent Variable: Level of Conflict
- Logistic Regression
 - Test and train with Mean Squared Error to determine effectiveness
 - Bayes Estimator to differentiate selection of different levels
- Conflict Level Random Forest
 - Best variable selection
 - Proper Depth and Tree Count to avoid overfitting
 - Outputs level of conflict in different situations

 Random Forest Model

Random Forest Yielding Conflict Levels in Nemo Data



- Outputs are the expected Conflict level Response given responses in NEMO
- Allows for visualization of decision-making within model at different nodes given specific impactful variables

 Random Forest Model

- As the number of trees increases, the error rate decreases
- Want to avoid overfitting
- Extended data collection could allow for factor anticipation

Challenges & Limitations

- Lack of Data
 - Limitations of survey data
 - Limited open source data to support NEMO data
 - Lack of research related to variable selection for Sudanese conflict
- Differing Data Classes
 - How to layer and combine numerical and text responses effectively?

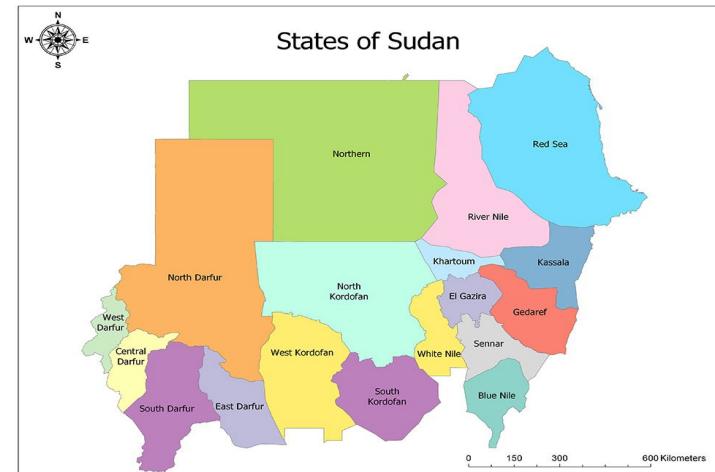
Open Source Data Analysis

||| The Armed Conflict Location & Event Data (ACLED)

- Publicly available data
- Armed conflict data from 1997 to 2022
- Variables: year, event type, actor, location, fatalities, etc.
 - 29 total variables

ACLED Goals

- Gain insights from descriptive statistics and visualizations
- Longitudinal Analyses
 - Limit to past 6 months and compare to known regions of activity:
 - Blue Nile
 - South Darfur

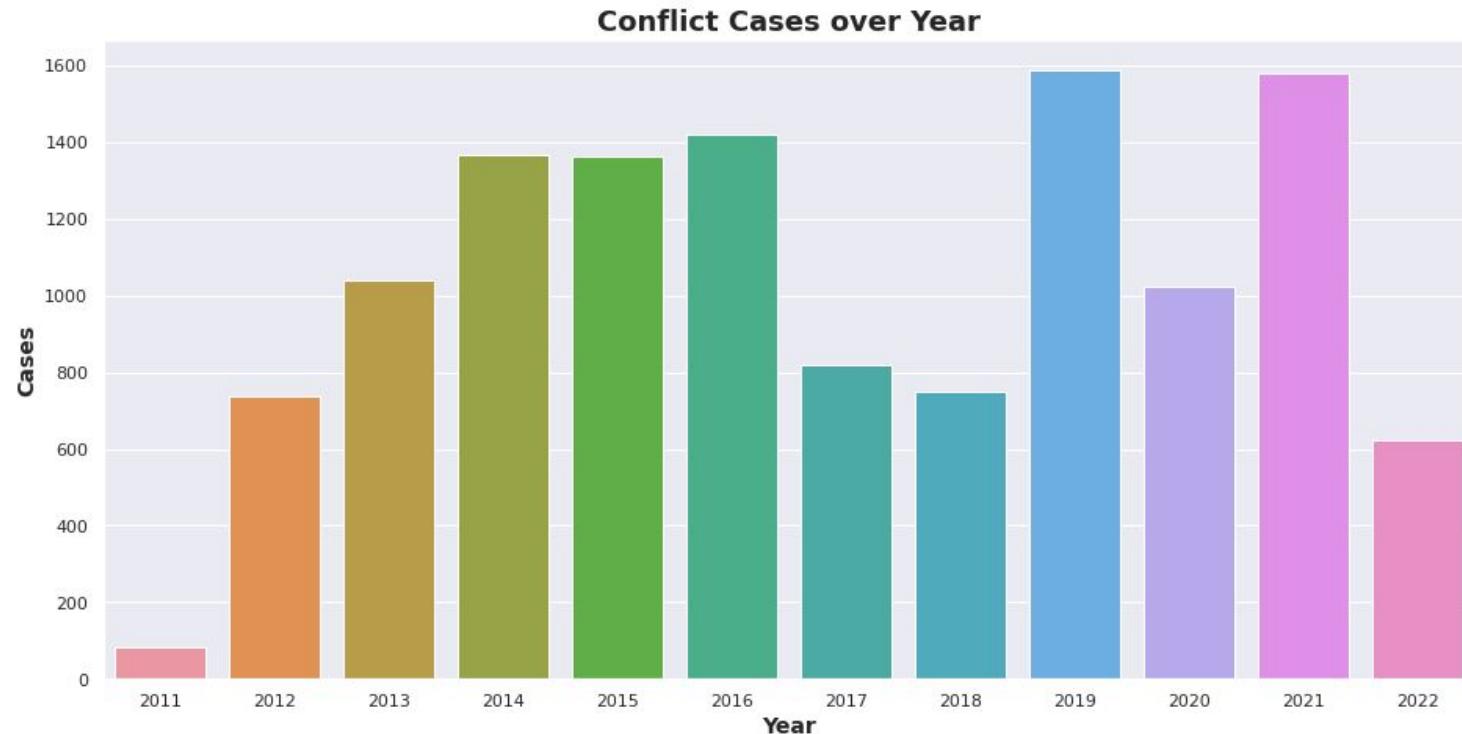


ACLED's Sudanese Data

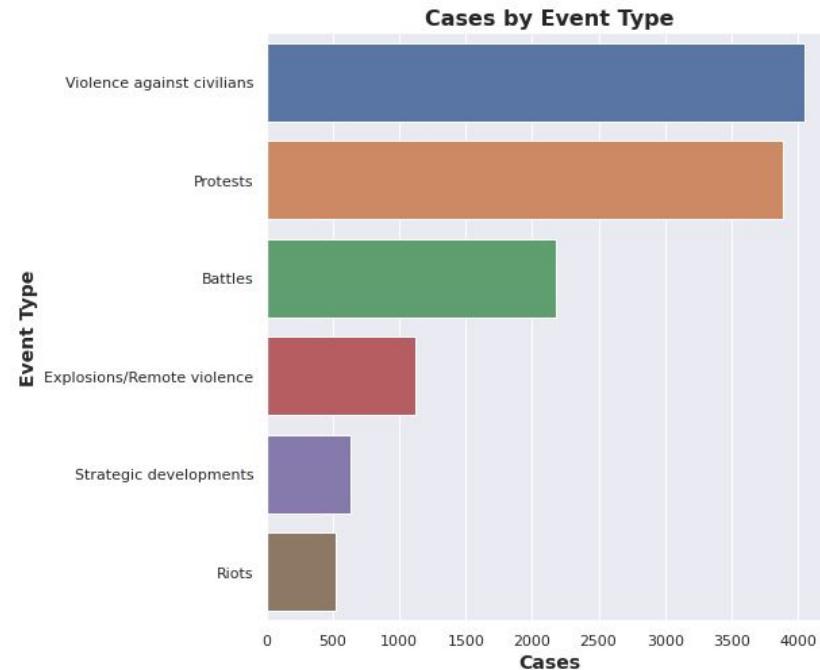
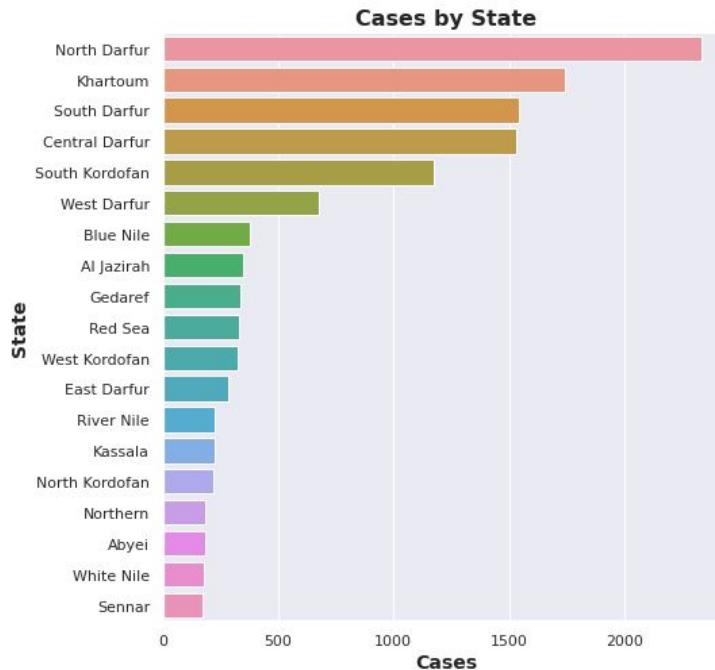
- Subset data after South Sudan's independence (July 9, 2011)
- 12,386 total cases
- 27,027 total fatalities
- Location - total 19 States, coordinates
- Event types - Battles, Protests, Strategic developments, Explosions/Remote violence, Violence against civilians, Riots



Conflicts Per Year Since Independence



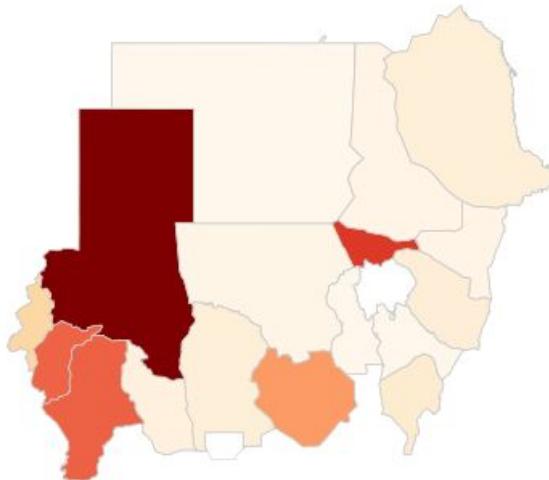
Cases by State and Event Type



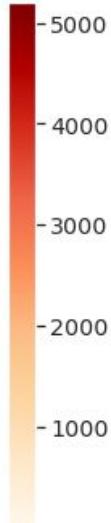
- Darfur states, Khartoum, & South Kordofan have large number of cases
- 'Violence against Civilians' & 'Protests' have the largest number of cases

Cases and Fatalities by State

Conflicts per State since 2011

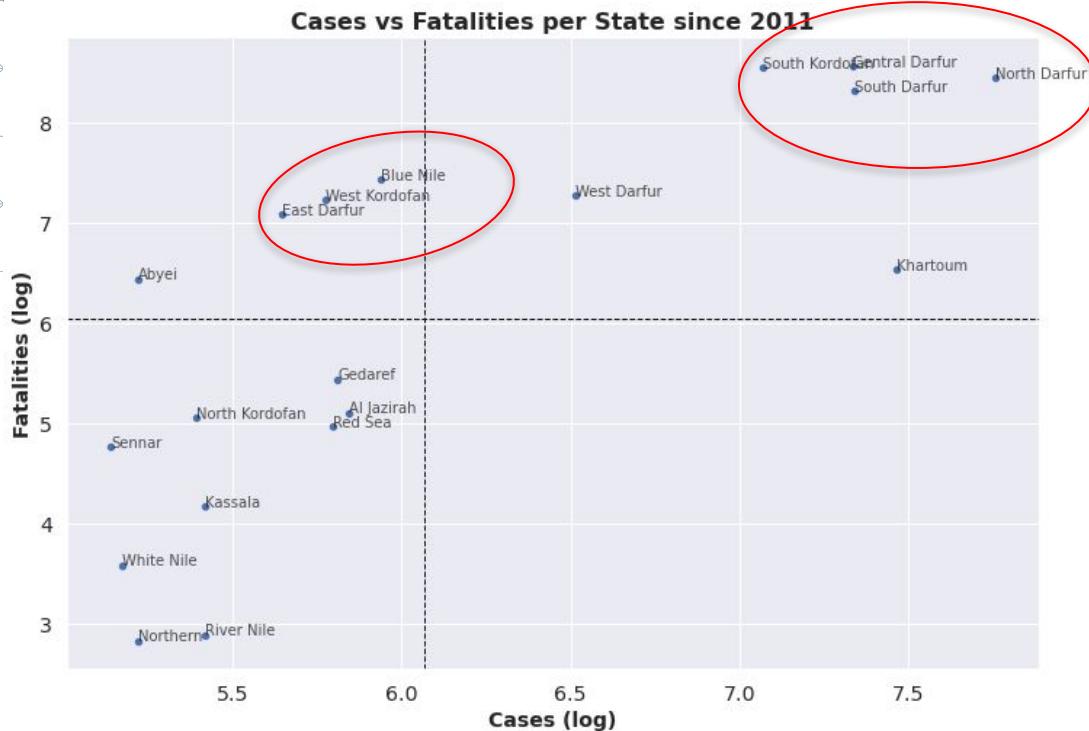


Total Fatalities per State since 2011



- **North Darfur, South Darfur, Central Darfur, and South Kordofan** have large cases and fatalities
- **Khartoum** has relatively higher cases but lower fatalities
- **Blue Nile, East Darfur, and West Kordofan** have high fatalities despite having relatively lower cases of conflict

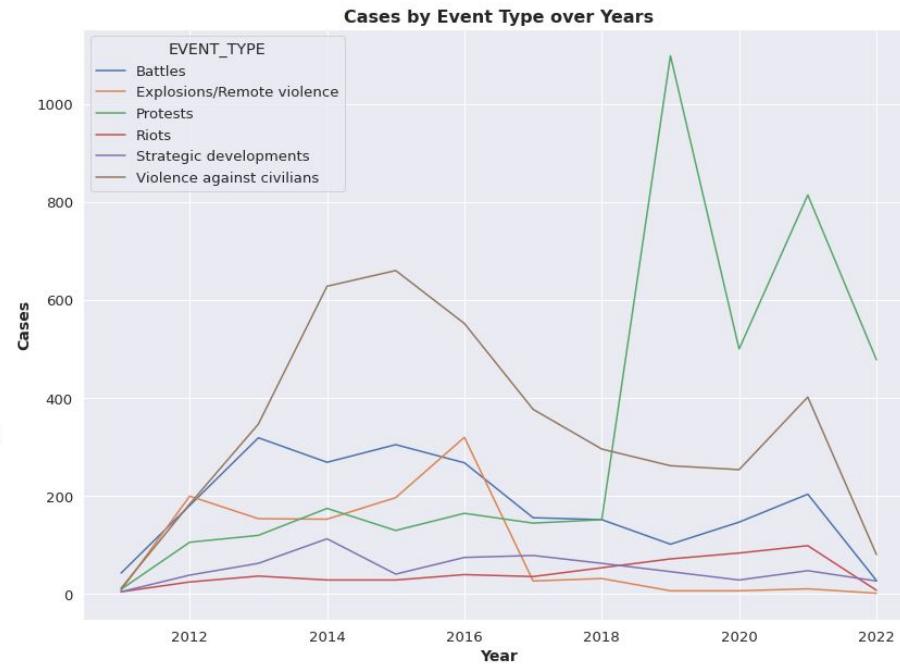
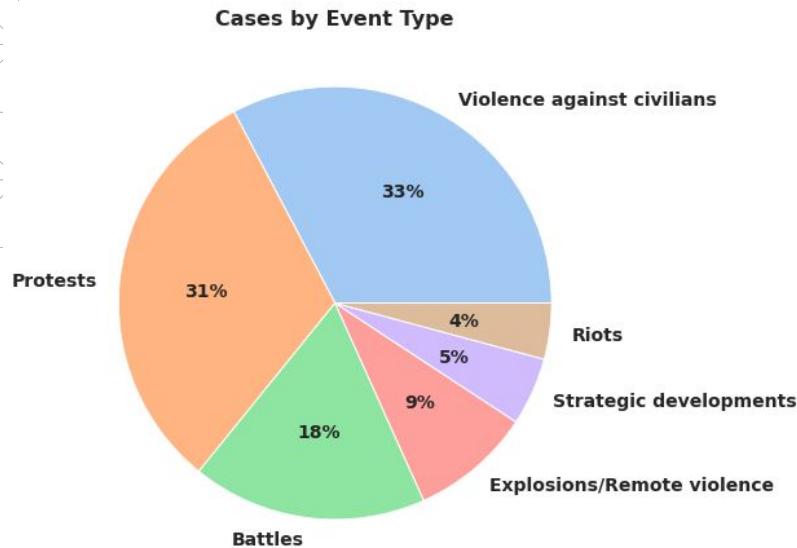
Suggested Visualization: Cases vs. Fatalities by State



Example insights if population data were available:

- **South Kordofan, Central Darfur, and North Darfur** have similar level of risk to **South Darfur**
- **West Kordofan** and **East Darfur** have a similar level of risk to **Blue Nile**
- The risk level of **West Darfur** and **Khartoum** is located between the circled two groups

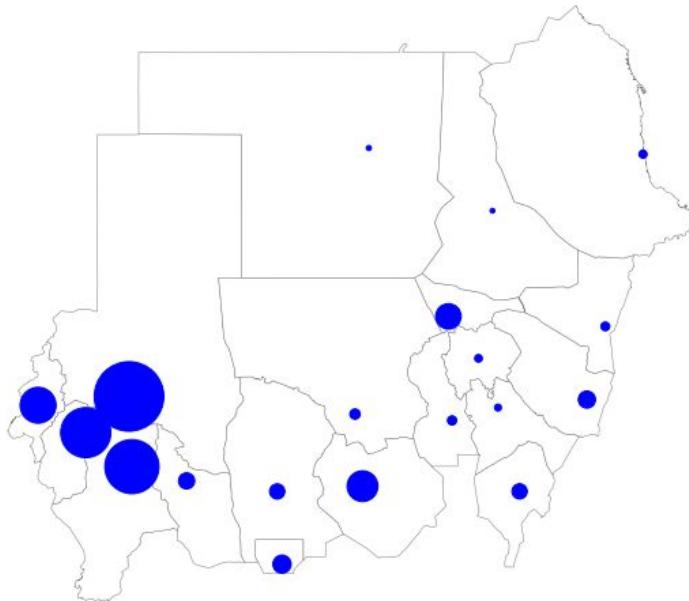
Types of Conflicts in Sudan



- Violence against Civilians has the largest cases, followed by Protests
- Increase in Protest since 2019

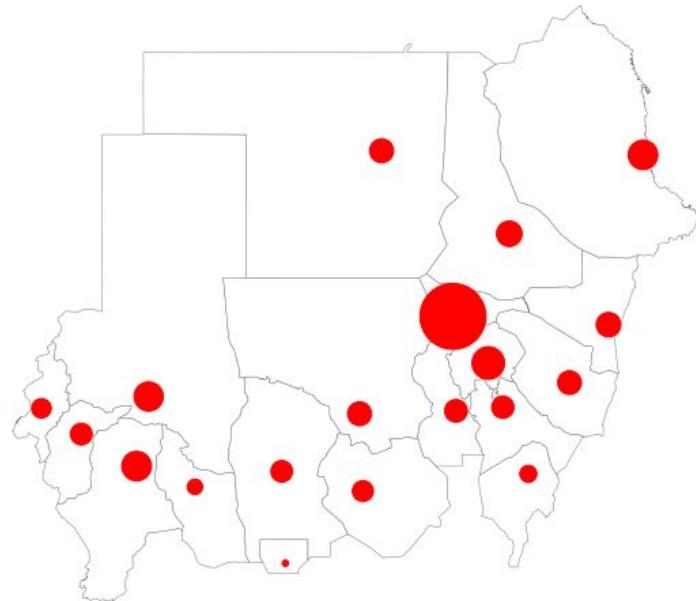
Violence against Civilians & Protests by State

Violence against Civilians by State since 2011



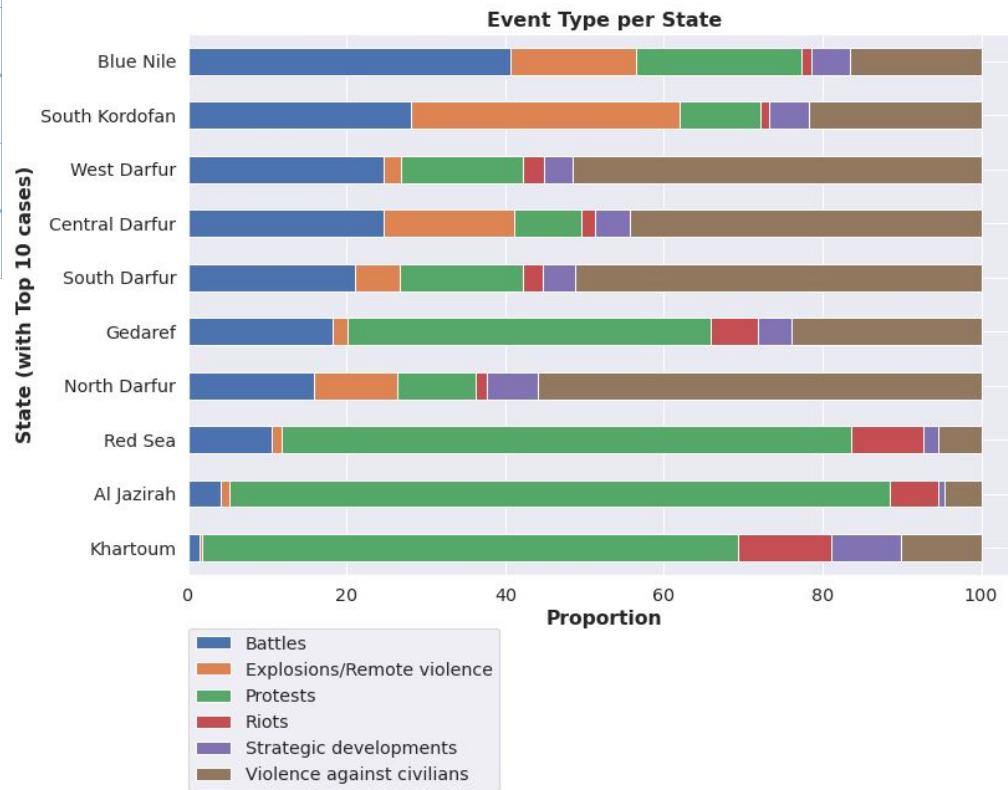
- Concentrated in **Darfur** states

Protests by State since 2011



- Equally distributed except **Khartoum**

Proportion of Conflict Type by State



- **Blue Nile, South Kordofan -** Battles and Explosions take up 50-60%
 - Large fatalities, Low cases
- **Khartoum -** More than 60% are Protests
 - Low fatalities, Large cases
- **West Darfur, Central Darfur, South Darfur, North Darfur -** Violence against civilians take up 50%

Challenges & Limitations

- Lack of population data
 - Limits insights due to inability to analyze per capita rather than count
 - States with larger populations likely to show as having higher degrees of conflict
- Limited data availability from other open source providers such as GDELT

Suggested Next Steps

Suggested Next Steps

- Further define variables in NEMO survey
 - Scale of 1-4 makes it difficult to compare to other data sets such as ACLED, by assigning ranges or classifications for what 1-4 represents allows for clearer understanding of results
- Increase or expand data collection
 - Allows for better and more reliable insights and lower error rate
- Implementation of qualitative responses into NEMO model



Suggested Next Steps

- Multifactor model
- Descriptive & Determinative Data



EMORY
UNIVERSITY



EMORY
COLLEGE
OF ARTS AND
SCIENCES

Department of Quantitative
Theory and Methods

*Thank you!
Questions?*





References

- Ahmado, N. (2021, February 18). *Ten years after Arab Spring, Egypt's press freedom dwindle*s. VOA. Retrieved April 21, 2022, from https://www.voanews.com/a/extremism-watch_ten-years-after-arab-spring-egypts-press-freedom-dwindles/6202188.html
- Britannica, T. Editors of Encyclopaedia (2022, January 19). Egypt
- Kirby, J. (2021, April 24). *"dying by blood or by hunger": The war in Ethiopia's Tigray Region, explained*. Vox. Retrieved April 21, 2022, from <https://www.vox.com/22370629/ethiopia-tigray-eritrea-war-ethnic-cleansing>
- Rashwan, N. (2019, August 1). *Four killed at Sudan protest, and tensions rise*. The New York Times. Retrieved April 21, 2022, from <https://www.nytimes.com/2019/08/01/world/middleeast/sudan-protest-killed.html>
- The Armed Conflict Location & Event Data Project . (2022, March 30). ACLED Africa Regional Data. ACLED. Retrieved April 20, 2022, from <https://acleddata.com/curated-data-files/>
- Tigray conflict*. Tigray Conflict | Human Rights Watch. (2022, April 6). Retrieved April 21, 2022, from <https://www.hrw.org/tag/tigray-conflict>
- Uprising of 2011. Encyclopedia Britannica. <https://www.britannica.com/event/Egypt-Uprising-of-2011>

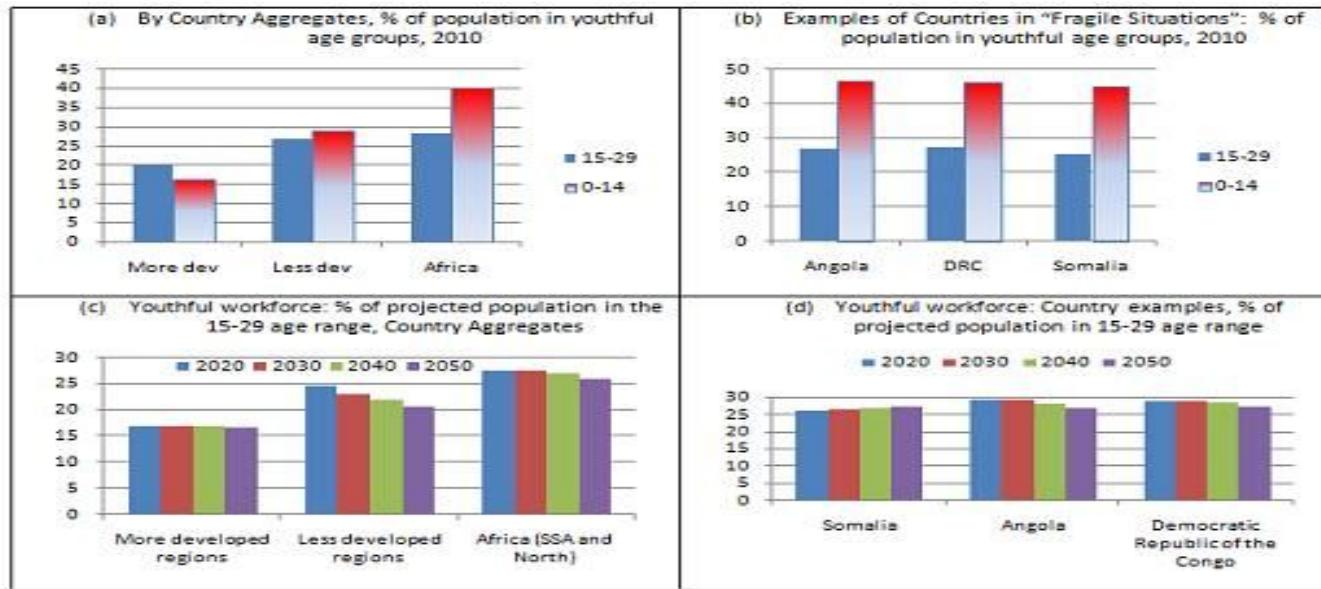
Appendix

Literature Review - Demographic Analyses

- Key Variables to Target:
 - Unemployment for educated youth
 - “Situation for youth”
 - Youth Bulge
 - Employment opportunities for youth
 - Migration opportunities for work
 - Education stress
 - Dependency ratio
 - Marriage status
 - Number of children
- Limitations
 - Limited data in Sudan
 - No recent census data
 - Data relies on projected percentages
 - Comparisons within Africa include countries with different levels of development

Literature Review - Demographic Analyses

Figures 1(a)-(d): Age Distribution of the Population, Aggregate and Country Examples, %



Data from United Nations, Department of Economic and Social Affairs, Population Division (2011). World Population Prospects: The 2010 Revision. Medium fertility scenario is used for the 2050 projections.

Literature Review - Economic Analyses

- Key Variables to Target:
 - National debt
 - Low economic growth
 - Male and female pay
 - Other income - female
 - Number of children
 - Socioeconomic status
 - Socioeconomic stress
 - Taxes on the elite
- Limitations
 - Limited data in Sudan
 - Unknown direction of correlation between economics and political violence

Literature Review - Health Analyses

- Key Variables to Target:
 - Death
 - Disability
 - Destruction of health services
 - Disruption of health programs
 - Militarisation
 - Refugees
 - Displaced peoples
 - Psychological stress
 - Patterns of distress
- Limitations
 - Most research focuses on the effect of political violence on mental health
 - Not enough research to support or refute using these variables