Andy Fan (Section 2: M/W 10:30-11:50), Git: programming-af

Juan Ulloa (Section 2: M/W 10:30-11:50), Git: jmulloa25

Will Sigal (Section 2: M/W 10:30-11:50), Git: willsigal

Data Links (includes datasets too large for “Data” folder)

ORIGINAL UNPROCESSED DATA:

https://drive.google.com/drive/folders/1Mfv-cvmm9iIgnQ7UKKiMuoTJo0Vjpcv4?usp=sharing

CLEANED DATA:

https://drive.google.com/drive/folders/1los_5JwF1rNVOaFtjH1B2Z8thgmBXHxC?usp=drive_link

Research Question:

We were interested in how gang violence/armed conflict can impact economic development. Specifically, our research question was how decrease in gang violence can effect the economic development indicators of El-Salvador and by extention Central America.

We were able to obtain ACLED data that allowed us to identify periods of discontinuity in levels of violence in El-salvador. The “Territorial Control Plan” implemented in 2019 by Nayib Bukele led to a drastic decrease in total violence, gang vaiolence, and violence involving civilians. Which allowed us to make comparisons of economic/development statistics before and after this period by also using World Bank development statstics.

Coding and Approach:

We utilized 2 ACLED violent events datasets and 2 World Bank Development indicators datasets to do our analysis. Our broad approach was to: 1.perform exploritory analysis on all datasets, identify useful data, trends, and graphs. 2.analyse trends in violence and make sure there is actually substaintial support for a reduction in violence related to policy events. 3.Seek to dissect and better understand the categories, actors, etc. of violence. 4.join ACLED and world bank data, to see how violence impacts development indicators.

coding descriptions

1.final_project_data_clean.qmd – cleans all the datasets (merges, na removes, pivots, remove redundant columns etc.) from raw data into usable data that the other qmds load

2.Economic_Analysis.qmd – data analysis using world bank development indicators such as GDP, homicides, graph correlation plots and indicators over time. Also graphing development indicators vs violence by merging World Bank and ACLED datasets, such as graphs of Birth/Death Rates vs #violent events from 2018-2023. (also containes some code related to shapefile cleaning used later in shiny app)

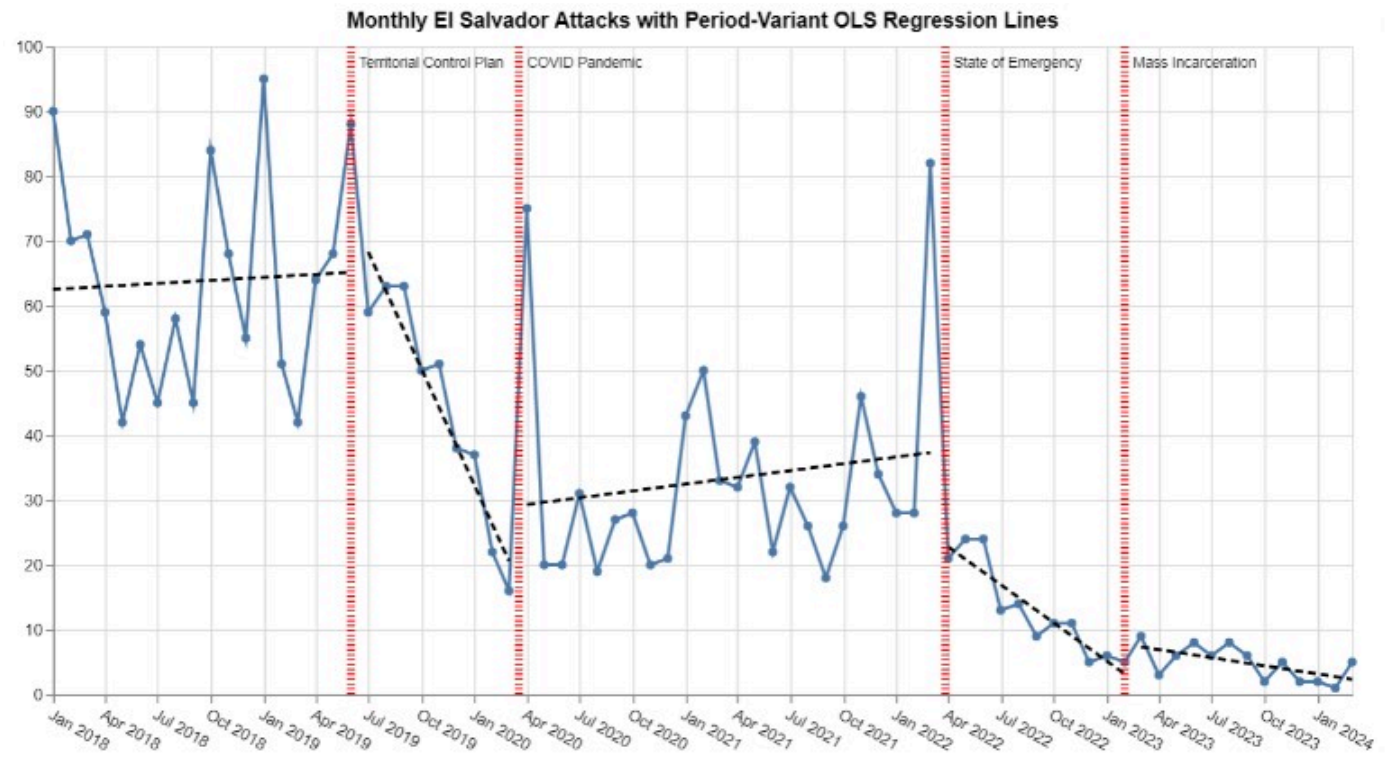
3.EDA_Analysis.qmd – data anslysis primarily on acled violecnce data and el-salvador specific data. Includes bar and pie charts that breakdown violent events by types, actors etc. Includes plots related to violence for both Central America and El-Salvador specifically. Includes natural language modeling using package 'gensim' to see how descriptions of events and their prevalence change over time. Also includes mapping for shiny app.

4.Time_Series_Analysis – graphs of violence in central america and el-salvador, including lines of best fit, shaded areas for impactful events etc.

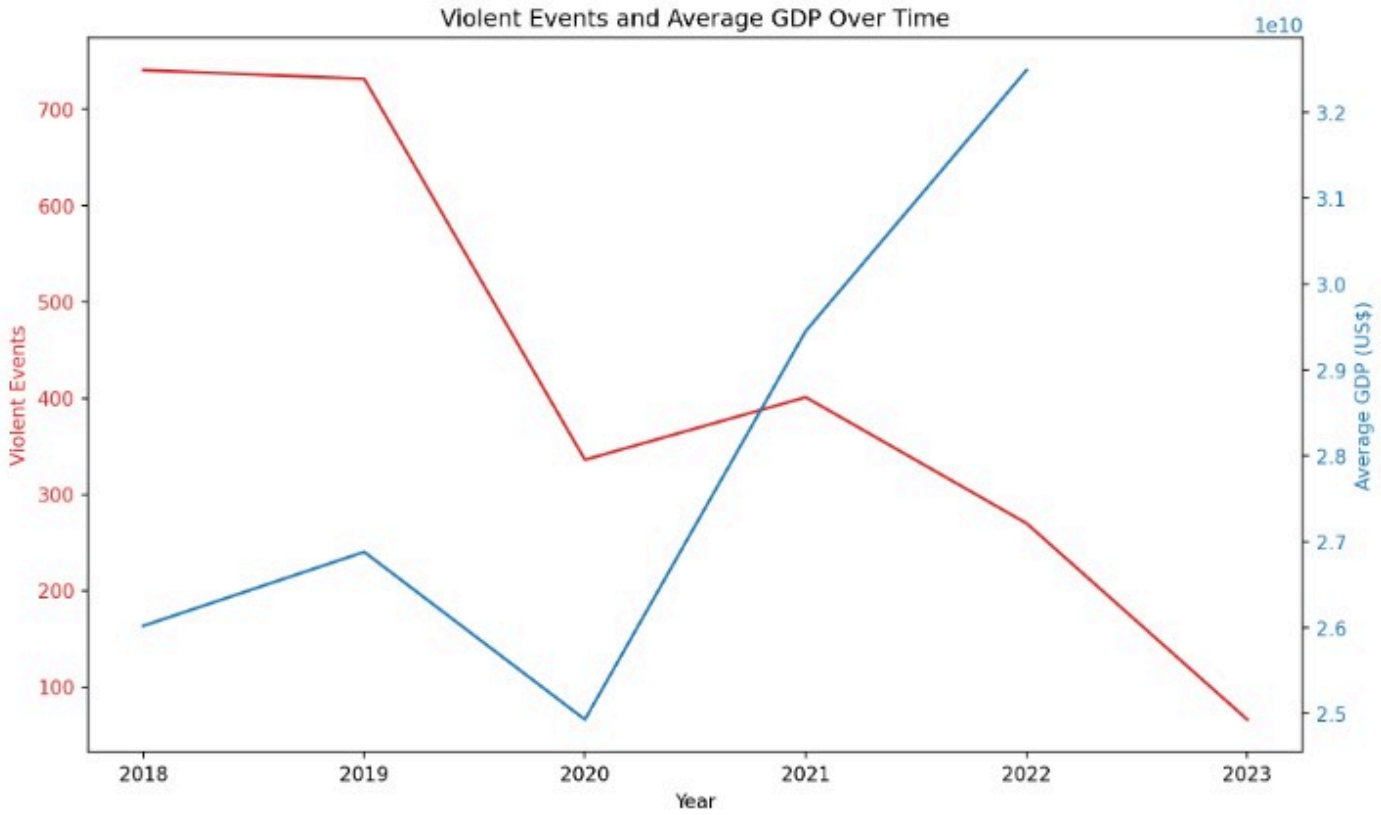
5.app.py – Our Shiny app: Contains dynamic plot/shapefile that shows changes in indicators over time. Country/Region and indicators all selectable with dropdown menu.

static plots

(We produced over a dozen static plots across all code, selected 2 to show as an example):



Graph of Violence across time in El-Salvador, includes key events and lines of best fit



Graph of relation between Central American GDP and total number of violent events

Directions for future work

We will definetly want to further explore how violence effects more types of economics indicators. There are hundereds of economic indicators provided by the world bank, we simply did not have time to examine most of them. Given more time we would like to explore this further.

We also want to see how different types of violence correlate with these indicators (eg does police violence affect certian indicators a different way than gang violence?)

Our analysis also identified an increase in violence in cournties surrounding El-Salvador whilst El-Salvador's own violence rate decreased. This may be a spillover effect that we would certianly like to investigate more.

We would also want to add regression analysis to our study, such that we can skee to infer potential causual relations instead of just correlation.