

Brigadas-Salud-Materna — Data Pipeline

Overview

This repository contains a reproducible Python data pipeline that builds ADM2-level risk tables for maternal health brigades in Mexico. The pipeline integrates multiple datasets—conflict, health access, poverty, and population—to generate actionable, evidence-based risk scores. It supports humanitarian targeting by identifying municipalities where violence, access constraints, and structural vulnerabilities converge to elevate maternal health risks.

View full repository: <https://github.com/robertclayh/Brigadas-SaludMaterna>

Input Variables

- *Violence* (V_{30} , V_{3m}) (ACLED 30-/90-day event counts)
- *Access* (A) (CLUES facility density, inverse scaled)
- *Poverty* (MVI) (CONEVAL 2020 % pobreza)
- *Spillover* (S) (Queen-contiguity neighbor mean of violence)¹
- *Forecast* ($CAST$) (ACLED CAST, state-level)
- *Population* (WorldPop 2025 raster; WRA \approx 25% of total)²

Model Structure

Two composite indices quantify risk:

Descriptive Composite Risk (DCR100)

Reflects current structural and spatial risk.

$$DCR100 = 100 \times (0.35V_{3m} + 0.15S + 0.30A + 0.20MVI)$$

Predictive Risk Score (PRS100)

Captures near-term shifts using recent trends and CAST forecasts.

$$PRS100 = 100 \times (0.30V_{30} + 0.25\Delta V_{30} + 0.10S + 0.18CAST + 0.12A + 0.05MVI)$$

The overall **priority score** balances predictive and structural risk:

$$priority100 = 100 \times (0.6PRS + 0.4DCR)$$

Automation

At least one variable—ACLED event data—refreshes daily through automated API calls, along with CAST forecasts. If the API recency cap has not advanced, cached data are reused. Static layers (population, poverty, and facilities) rebuild only when missing.

Data Access Caveat

Following ACLED’s 2024–25 licensing change, disaggregated event data newer than 12 months are restricted. The pipeline and weights assume timely access; if such data remain unavailable, the model can pivot to use longer-term trends and state-level aggregates while retaining analytical validity.

Outputs

File	Description
adm2_risk_daily.csv	Main risk table, one row per ADM2
acled_events_violent_90d.csv	90-day violent event points
adm2_geometry.csv	ADM2 centroids for mapping
Google Sheet Dashboard	Viz-ready tables for Tableau

¹Queen-contiguity defines neighbors sharing a border or vertex (Arribas-Bel, *A Course on Geographic Data Science*, Lab E).
²Based on INEGI population statistics, women 15–49 \approx 25% of total; used as WRA proxy pending finer data.

Citations

Primary Data Sources

- Armed Conflict Location & Event Data Project (ACLED). (2025). *ACLED dataset*. <https://acleddata.com>
— Events and CAST forecasts accessed via API for Mexico, filtered for violent event types (rolling 30-day and 90-day windows).
- Secretaría de Salud, Dirección General de Información en Salud (DGIS). (n.d.). *Catálogo de Clave Única de Establecimientos de Salud (CLUES)*. http://www.dgis.salud.gob.mx/contenidos/sinais/s_clues.html
— Facility locations and attributes used to calculate access indicators through spatial joins.
- Consejo Nacional de Evaluación de la Política de Desarrollo Social (CONEVAL). (n.d.). *Pobreza a nivel municipio 2010–2020*. <https://www.coneval.org.mx/Medicion/Paginas/Pobreza-municipio-2010-2020.aspx>
— Municipal poverty data integrated as a structural vulnerability indicator.
- Instituto Nacional de Estadística y Geografía (INEGI). (n.d.). *Censo de Población y Vivienda*. <https://www.inegi.org.mx/programas>
— Population counts used to estimate total population and women of reproductive age (WRA), approximated at 25% of the total.
- WorldPop. (2025). *WorldPop Mexico 2025 (R2025A) population dataset*. <https://hub.worldpop.org/geodata/summary?id=74383>
— 100m resolution raster used to spatially aggregate population by ADM2.
- Humanitarian Data Exchange (HDX). (n.d.). *COD-AB Mexico Administrative Boundaries (ADM2)*. <https://data.humdata.org/dataset/ab-mex>
— Used as the spatial reference for joins, spillover computation, and centroid extraction.

Methodological Reference

- Arribas-Bel, D. (n.d.). *A course on Geographic Data Science — Lab E: Spatial Weights (Queen/Rook contiguity and transformations)*. https://darribas.org/gds_course/content/bE/lab_E.html
— Reference for constructing Queen-contiguity spatial weight matrices using PySAL.

Tools and Assistance

- GitHub Copilot. (n.d.). *GitHub Copilot for Visual Studio Code*. <https://github.com/features/copilot>
— Used for code autocompletion and function scaffolding during pipeline development.
- OpenAI. (2025). *ChatGPT (VS Code Integration)*. <https://help.openai.com/en/articles/10128592-how-to-install-the-work-with-apps-visual-studio-code-extension>
— Used for debugging assistance and formatting within the development environment.

AI Usage Disclosure

This deliverable was developed with GitHub Copilot and the ChatGPT VS Code integration for code autocompletion and debugging. All modeling decisions, analysis, and writing were performed by the author in compliance with SDS program guidelines.