

# Estimating the Underreporting of Gender-Based Violence in Mexico

Based on Digital Conversation Data

Carla Pilgram

MSc Applied Social Data Science



# Motivation

- “What gets counted counts” – Data Feminism
- Many victims don’t recognize what they experience as violence
- Without measuring the experiences that remain hidden, we are designing policies based on partial truths
- Violetta is a chatbot that prevents violence and has reached 300,000+ users *vi*
  - Most users only spoke up after seeing a paid media campaign and not because they were looking for help

# Research Question

To what extent can digital support interactions reveal the underreporting of gender-based violence in Mexico, and what role does paid-media play in surfacing these hidden cases?

# What This Project Will Deliver

- Estimate violence not captured in official records
- Identify patterns and severity using ML
- Measure effect of campaigns on visibility
- Provide evidence to inform public policy

# Methodology: Combining ML, Statistics & Digital Conversations

- Use digital conversation data from a structured chatbot (2020–2024)
- Analyze over 200,000 anonymized user interactions
- Combine machine learning and regression to uncover underreported violence
- Some classifications from Fundación Origen used as ground truth

# Predicting Violence Type & Risk with ML

- Input: Chatbot paths, time, region, campaign source
- Label: Type of violence / risk level (from Fundación Origen)
- Models: Random Forest, Logistic Regression, XGBoost
- Goal:
  - Automate classification of 100k+ cases
  - Detect high-risk users that didn't reach formal support
  - Scale measurement of “invisible violence”

# Regression to Estimate Underreporting

- Compare chatbot cases vs. official data (INEGI, SESNSP)
- Include paid media as explanatory variable
- Use regression models
- Explore pre/post campaign impact
- Goal: Quantify how much violence goes unreported – and when campaigns make it visible

# Ethical Approach

- All data used is anonymised and collected with user consent
- The project avoids extracting personal identifiers or sensitive metadata
- Prioritising lived experience and equity
- Aims to amplify marginalized voices often excluded from official data
- Models are designed for group-level insight, not individual profiling
- The entire analysis is developed with a **gender-sensitive and intersectional perspective**



# Why This Matters

- Invisible violence can't be addressed if it's not measured
- Digital tools allow victims to speak without fear
- Data and ML can reveal what institutions overlook
- We cannot fix what we cannot measure
  - But now, we can start to measure what was previously invisible