

## Documentació i codi del repositori de dades

**Títol del TFM:** Estudi sobre la relació entre el creixement de l'extrema dreta i la violència contra les dones.

Aquest directori conté tot el material computacional, les dades originals i els models estadístics utilitzats per a la investigació de la relació entre l'indicador del vot a l'extrema dreta i els indicadors de violència masclista.

### 1. Estructura del GitHub i fitxers:

**requirements.txt:** llistat detallat de llibreries de Python i les seves versions.

**Informe\_analític\_visual\_TFM\_Bfelip.pdf:** l'informe visual analític és un instrument de transferència dissenyat per fer la recerca socialment accessible i metodològicament rigorosa. El seu propòsit és doble: d'una banda, democratitzar l'accés a dades complexes (sèries temporals i models de regressió) mitjançant una narrativa visual coherent que permet identificar tendències i disparitats territorials; de l'altra, oferir una evidència empírica robusta que serveixi de contranarrativa davant els relats negacionistes.

Mitjançant un disseny que prioritza la claredat i l'ús de suports versàtils com PowerPoint, s'eliminen barreres d'entrada per a actors clau (polítics, periodistes i societat civil), transformant dades tècniques sobre extrema dreta i violència de gènere en coneixement estratègic i directament aplicable.

- **/data:** conté els conjunts de dades en format CSV.
  - originals/: fonts originals (INE, CGPJ, Ministeris), sense processar.
  - processats/: inclou el fitxer `rates_all_metrics_by_province_year_unified_final_FIXED.csv`, que és el dataset unificat i net utilitzat en els models definitius.
- **/notebooks:** fitxers Jupyter Notebook (.ipynb) organitzats per etapes:
  - `01_neteja_i_preparacio.ipynb`
  - `02_analisi_exploratoria_i_descriptiva.ipynb`

### 2. Requeriments tècnics

L'anàlisi s'ha desenvolupat sota l'entorn Python 3.8.8. Per garantir la reproductibilitat dels resultats, es recomana seguir aquests passos:

1. Instal·lar la versió de Python 3.8.8.
2. Instal·lar les dependències mitjançant el gestor de paquets pip: `pip install -r requirements.txt`

### 3. Notes sobre metodologia i rigor estadístic.

Per abordar la naturalesa multidimensional de les dades, s'han aplicat tècniques avançades:

- **Control de la taxa de falsos descobriments (FDR):** aplicat a les matrius de correlació per corregir l'efecte de les comparacions múltiples.
- **Imputació de dades (LOCF):** s'ha utilitzat Last Observation Carried Forward en les sèries temporals per mantenir la coherència visual i analítica davant de possibles buits en el registre provincial.
- **Models d'Efectes Fixos (FE):** utilitzats per explotar la variació *within* (intra-provincial) i aïllar els efectes de variables temporals i territorials constants.

### 4. Autoria i contacte

Autora: Begoña Felip Bengochea (bfelip@uoc.edu)

Data: gener 2026