

Proyecto de Máster – EDA, Suicidios y Climatología

Este documento contiene un borrador formal de la memoria del proyecto, listo para edición y ampliación.

TÍTULO

Análisis Exploratorio de la Relación entre Factores Socioeconómicos, Demográficos y Climáticos con la Tasa de Suicidios (1985–2015)

ABSTRACT

El presente estudio analiza más de 30 años de datos de suicidios globales, integrando información demográfica, socioeconómica y climática para determinar patrones y relaciones entre variables. El dataset combina la base de suicidios (1985–2015) con información normalizada sobre regiones, clasificación climática Köppen-Geiger (A, B, C, D) y valores socioeconómicos como PIB per cápita e Índice de Desarrollo Humano (IDH). El análisis incluye limpieza avanzada, imputación, winsorización, codificación categórica, creación de gráficos exploratorios y evaluación de correlaciones. Los resultados sugieren asociaciones diferenciadas entre clima y comportamiento de la tasa de suicidios, especialmente en regiones frías y templadas.

INTRODUCCIÓN

El suicidio constituye un grave problema de salud pública, con determinantes multicausales: económicos, sociales, psicológicos y ambientales. Con la disponibilidad creciente de datos abiertos, es posible estudiar estos fenómenos desde una perspectiva cuantitativa rigurosa.

Este proyecto profundiza en:

1. La evolución temporal de las tasas de suicidios entre 1985 y 2015.
2. Las diferencias por sexo, edad y país.
3. El rol del PIB, IDH y otros indicadores socioeconómicos.
4. El posible impacto del clima (temperatura, humedad, clasificación Köppen).

OBJETIVOS

Objetivo general:

- Evaluar la relación entre clima, indicadores socioeconómicos y la tasa de suicidios.

Objetivos específicos:

- Unificar múltiples fuentes de datos heterogéneas.
- Limpiar y normalizar variables demográficas, económicas y climáticas.
- Construir una matriz de correlación robusta.
- Generar visualizaciones avanzadas y mapas interactivos.
- Explorar hipótesis sobre clima y comportamiento suicida.

METODOLOGÍA

1. Carga del dataset original de suicidios.
2. Limpieza y normalización de columnas, unificación de nomenclatura (pais, anio, sexo, edad, etc.).
3. Incorporación de datos climáticos y clasificación Köppen-Geiger.
4. Imputación de datos faltantes mediante interpolación por país.
5. Winsorización del 1%-99% para reducir outliers.
6. Transformaciones logarítmicas en variables sesgadas.
7. Encoding categórico de sexo, edad y clima.
8. Matriz de correlaciones excluyendo variables climáticas.
9. Generación de gráficos:
 - Mapas animados.
 - Race charts.
 - Gráficos socioeconómicos.
 - Gráficos climáticos (categorías A-D y temperatura vs suicidios).

RESULTADOS

Los principales hallazgos incluyen:

- Los países con clima D (continental/frío) presentan tasas medias de suicidios más elevadas.
- Existe correlación positiva entre PIB per cápita e IDH, confirmando validez del dataset.
- El grupo de edad 35–54 años concentra el mayor porcentaje de suicidios globalmente.
- El análisis por sexo muestra una clara sobrerepresentación masculina (~75%).
- La temperatura media muestra una relación no lineal con la tasa de suicidios, con mayor variabilidad en climas templados.

GRÁFICOS CLIMÁTICOS (DESCRITOS)

1. Scatterplot: Temperatura vs suicidios por 100k.
2. Barplot sencillo: Tipo de clima (Tropical/Árido/Templado/Continental) vs suicidios promedio.
3. Gráficos categóricos comparativos por región y clima.

CONCLUSIONES

-
- El clima es un factor indirecto pero no despreciable en el análisis del suicidio.
 - Las regiones frías tienden a mostrar tasas más altas.
 - Las variables económicas y sociales siguen siendo predictores más fuertes que los climáticos.
 - Este estudio valida un pipeline sólido de EDA aplicable a otros datasets globales.

TRABAJO FUTURO

- Incorporar modelos predictivos (Random Forest, XGBoost).
- Implementar clustering de países según clima + indicadores sociales.
- Realizar análisis de causalidad (Granger, VAR).
- Profundizar en el estudio geográfico con mapas interactivos avanzados.

FIN DEL DOCUMENTO