



BENEMÉRITA UNIVERSIDAD AUTÓNOMA DE PUEBLA

FACULTAD CIENCIAS DE LA COMPUTACIÓN

INTRODUCCION A LA CIENCIA DE DATOS

M.C. JAIME ALEJANDRO ROMERO SIERRA

REPORTE FASE 1: DESCUBRIMIENTO DEL

PROYECTO

ABRAHAM FUENTES LÓPEZ

21/10/2024

Reporte Fase 1: Descubrimiento del Proyecto

1. Título del Proyecto

"Análisis de Patrones de Criminalidad en la Ciudad de Los Ángeles (2020 a la Actualidad)"

2. Objetivo del Proyecto

El objetivo de este análisis es identificar patrones de criminalidad en la ciudad de Los Ángeles para mejorar la asignación de recursos policiales, implementar políticas de seguridad más eficaces y reducir la incidencia delictiva en áreas con mayor concentración de delitos. Se busca comprender los factores clave que afectan la frecuencia y tipo de delitos en la ciudad, con la finalidad de diseñar estrategias de prevención más precisas.

3. Descripción del Problema

Desde 2020, la ciudad de Los Ángeles ha enfrentado un aumento significativo en la criminalidad, lo cual ha generado preocupación tanto en las autoridades como en las comunidades locales. A pesar del despliegue de esfuerzos para mitigar esta situación, la respuesta de las fuerzas del orden se ha visto obstaculizada por la falta de comprensión detallada sobre los patrones de criminalidad en términos de ubicación, tiempo y características demográficas. La actual distribución de recursos policiales se ha basado en enfoques reactivos que no consideran de manera precisa las áreas y momentos de mayor incidencia delictiva.

La ausencia de un análisis profundo de los factores que influyen en la frecuencia y tipo de delitos limita la capacidad de los responsables de seguridad para implementar estrategias preventivas adecuadas. Este proyecto busca identificar los patrones de criminalidad más relevantes en la ciudad, a fin de optimizar la asignación de recursos y mejorar las políticas de seguridad pública. Al abordar estos vacíos en el conocimiento, se espera contribuir a una reducción efectiva de la incidencia delictiva y al fortalecimiento de la seguridad en las comunidades más afectadas de Los Ángeles.

4. Recursos Disponibles

- **Tecnología y Herramientas:**
 - Python (Pandas, NumPy) para manipulación y análisis de datos.

- Matplotlib y Seaborn para la visualización de datos.
- **Datos:**
 - **Número de crimen:** Identificador único del incidente.
 - **Fecha del informe:** Fecha en que se presentó el reporte.
 - **Fecha del suceso (Date_occurred):** Fecha del incidente.
 - **Hora del suceso (Time_occurred):** Hora del incidente.
 - **Área:** Ubicación general del incidente (aproximado por la cuadra).
 - **Código de delito (Crime_Code):** Clasificación del tipo de crimen.
 - **Edad de la víctima (Victim_age):** Edad de la persona afectada.
 - **Género de la víctima (Victim_sex):** Género de la víctima.
 - **Descenso de la víctima (Victim_descent):** Origen étnico/racial de la víctima.
 - **Premisas (Premis):** Tipo de lugar donde ocurrió el delito (residencia, vía pública, etc.).
 - **Arma (Weapon):** Tipo de arma si es que se empleó alguna.
 - **Estado del sospechoso (Status):** tipo de categoría del sospechoso (continúa la investigación, Desconocido, Adulto arrestado, etc)

5. Hipótesis Iniciales

1. **Hipótesis 1:** Los delitos violentos tienen mayor prevalencia durante la noche y en áreas con baja presencia policial.
2. **Hipótesis 2:** Las víctimas de delitos contra la propiedad tienden a ser personas mayores de 40 años en áreas residenciales.
3. **Hipótesis 3:** Las áreas con mayor densidad poblacional presentan un número más alto de delitos reportados en comparación con áreas menos pobladas.

6. Definición de Stakeholders Clave

- **Departamento de Policía de Los Ángeles (LAPD):** Principal beneficiario del análisis; utilizarán los hallazgos para mejorar la asignación de recursos y ajustar estrategias de vigilancia.

- **Gobierno de Los Ángeles:** Podrán usar la información para el diseño e implementación de políticas de seguridad pública más eficientes.
- **Comunidades locales:** Indirectamente beneficiadas por las mejoras en la prevención del crimen y la mayor seguridad en sus áreas de residencia.
- **Investigadores y analistas de criminalidad:** Utilizarán los datos y análisis para estudios más profundos de comportamiento criminal.

7. Preguntas Clave

1. ¿Cuáles son las áreas de Los Ángeles con mayor frecuencia de crímenes violentos?
2. ¿Qué tipos de delitos son más comunes en cada área?
3. ¿Existe una relación entre la hora del día y el tipo de crimen cometido?
4. ¿Cuáles son las edades y géneros más afectados por delitos específicos?
5. ¿Qué tipos de premisas son más propensas a ser escenario de ciertos delitos?
6. ¿Cómo varían los patrones de criminalidad en función del origen étnico de las víctimas?
7. ¿Hay picos en la actividad criminal durante ciertos meses del año?
8. ¿Las áreas con mayor presencia policial tienen una menor incidencia delictiva?
9. ¿Cómo afecta la densidad poblacional la frecuencia de delitos en diferentes áreas?
10. ¿Qué impacto tienen las intervenciones recientes del LAPD en la reducción de criminalidad?

8. Fuentes de Datos Identificadas

- **Datos históricos de incidentes criminales del LAPD** (2020 - Actualidad).
- **Datos geográficos de áreas de criminalidad:** Definición de las zonas policiales en Los Ángeles.
- **Datos demográficos de Los Ángeles:** Información sobre la distribución de la población por edad, género y etnia.

- **Registros de intervenciones policiales recientes:** Información sobre el despliegue de recursos y programas comunitarios.

9. Justificación del Proyecto

La criminalidad es un tema de gran importancia para el bienestar social y económico de cualquier ciudad. Este análisis proporcionará información crucial que permitirá a las autoridades optimizar la distribución de recursos policiales y diseñar políticas de prevención de delitos más efectivas. Al comprender mejor las dinámicas de los incidentes delictivos —quiénes son las víctimas, dónde ocurren los delitos y en qué condiciones— se podrán tomar decisiones más fundamentadas para reducir el crimen, mejorar la seguridad pública y aumentar la calidad de vida de los ciudadanos.

10. Cantidad y Tipo de Datos

- **Cantidad de Datos:** La base de datos contiene aproximadamente 937,536 registros de incidentes reportados entre 2020 y la actualidad. Sin embargo, no se dispone de un registro completo del 100% de los reportes, lo que puede influir en el análisis.

-

- **Tipo de Datos:**

Estructurados: Incluyen identificadores de incidentes, fechas y horas, áreas de ocurrencia, códigos de delitos, y datos demográficos de las víctimas (edad, género, origen étnico). También se registran características del lugar y el uso de armas.

Geoespaciales: Datos que permiten realizar análisis geográficos y mapas de calor para identificar puntos críticos de criminalidad.

Estados del sospechoso (Status): Categorías como "Adult Arrest", "Invest Cont", "Adult Other", "Juv Arrest", "Juv Other" y "UNK", que reflejan el estado procesal de cada incidente.