**BENEMÉRITA UNIVERSIDAD**

AUTÓNOMA DE PUEBLA



**REPORTE FASE 1: DESCUBRIMIENTO DEL PROYECTO**

Asignatura: INTRODUCCION A LA CIENCIA DE DATOS

Carrera: INGENIERIA EN CIENCIA DE DATOS

Nombre del alumno: JOANA MONTSERRAT CARMONA NOGUEZ

Nombre del profesor: JAIME A. ROMERO SIERRA

## No. De cuenta: 202466401

## Fecha de entrega: 22 DE OCTUBRE DE 2024

**Reporte Fase 1: Descubrimiento del Proyecto**

**1.Título del Proyecto:**

**"Análisis del Tasas de criminalidad en México.”**

**2. Objetivo del Proyecto:**

Usar la información para ayudar a las autoridades a planificar estrategias de seguridad, identificando tendencias o ciclos de aumento de delitos en ciertos meses o bajo determinadas modalidades.

**3. Descripción del Problema:**

A partir de los datos sobre los meses en los que se cometieron los delitos se puede ayudar a identificar períodos o temporadas críticas en los que ciertos tipos de delitos son más probables. Esto se puede utilizar para determinar si hay un patrón cíclico, como un aumento del robo durante las temporadas festivas, ciertos tipos específicos de violencia que son característicos de distintos períodos, etc. Clasificar los delitos por su tipo y subtipo también permitirá identificar en qué momentos son más probables ciertos tipos de actividades delictivas. Por ejemplo, quizás los delitos contra la propiedad, como los robos, los daños intencionales y el fraude ocurren con más frecuencia en ciertos momentos, mientras que los delitos contra la integridad de la persona, como las lesiones y los homicidios, son predominantes en otros. Al analizar los bienes jurídicos que un crimen ataca, como la vida, la propiedad, la seguridad pública, etc, las agencias de seguridad podrán comprender dónde se concentran los riesgos en momentos específicos o bajo ciertos tipos de delitos. Esto permitirá a los responsables tomar medidas preventivas, como enviar patrullas policiales, lanzar campañas de concienciación, patrullar en zonas con un alto nivel de delitos en ciertas épocas.

**4. Recursos Disponibles**

Se utilizarán herramientas como Python y Pandas para procesar y manipular los datos de manera eficiente, mientras que Matplotlib y Plotly se emplearán para generar visualizaciones iniciales que faciliten el análisis.

Descripcion de Datos:

* **Año**: Este campo compila el año en el que se recopilaron los datos sobre los delitos cometidos en México. Permite analizar las tendencias a lo largo del año.
* **Clave\_Ent (Entidad clave**: Este código numérico es usado en lugar del nombre completo de cada entidad federativa (estado). Ayuda a identificar rápidamente para el usuario la entidad en cuestión.
* **Estado**: Nombre de la entidad federativa (estado) de México en la que ocurrieron los delitos. Geolocaliza la actividad delictiva.
* **Bienes jurídicos afectados**: Este campo describe el objeto jurídico protegido por la ley que ha sido violado el delito (por ejemplo, la vida, que se considera un derecho humano). Este campo posibilita categorizar delitos de acuerdo con el impacto en derechos fundamentales.
* **Tipo de delito**: Esta es una categoría general del delito cometido: asesinato, robo, secuestro. Muestra una única idea general sobre cómo se cometen crímenes.
* **Otro tipo de delito**: especifica con más detalle el tipo de delito (por ejemplo, asesinato intencional, robo a mano armada). Ofrece una visión más detallada del crimen que se está cometiendo.
* **Manera en que se cometió el delito**: Esta categoría describe las circunnancias y el modo en que el delito fue cometido (por ejemplo, con violencia, con arma de fuego). Es útil para estudiar crimen según sus características individuales.
* **Meses del año**: Muestra el número de delitos registrados en cada mes del año para una entidad, bien jurídico, tipo y subtipo de delito. Permite ver la dirección tomada por el crimen y las fluctuaciones a lo largo del año, síntoma de posibles auge o disminución en actividad delictiva.
* **Sexo/Averiguación previa**: Puede indicar el sexo de la víctima o el presunto criminal, o bien la naturaleza de la averiguación previa (expediente abierto en un proceso de investigación del crimen). Esta columna puede tener diferentes tipos de información según el diseño de la base de datos.
* **Rango de edad**: Clasifica a las víctimas o presuntos responsables de los crímenes según su grupo de edad. Estos rangos pueden ser definidos por intervalos (por ejemplo, 0-17, 18-29, 30-44, etc.), y ayudan a identificar patrones en la incidencia de crímenes en diferentes grupos etarios.

**5. Hipótesis Iniciales**

*Hipótesis 1*: "En México, el delito más frecuente relacionado con la violencia es el homicidio, y la mayoría de estos crímenes son cometidos con armas de fuego”.

*Hipótesis 2*: " En México, marzo es el mes en el que se registra el mayor número de delitos".

*Hipótesis 3*: “Aguascalientes es el estado con los niveles más altos de violencia en México”.

**6. Definición de Stakeholders Clave**

* Gobierno/ Autoridades de seguridad pública y justicia: Utilizarán los resultados para diseñar políticas y estrategias de prevención y combate al crimen, así como para asignar recursos de manera más eficiente.
* Ciudadanía: Los resultados del análisis pueden traducirse en mejores políticas de seguridad que beneficien a la población.
* Medios de comunicación: Pueden utilizar los resultados del análisis para informar al público sobre tendencias delictivas, áreas peligrosas y posibles causas de aumento de la violencia, ayudando a sensibilizar y educar a la sociedad sobre la situación.
* Políticos y legisladores: Utilizan los hallazgos del análisis para tomar decisiones informadas al momento de formular leyes y regulaciones que aborden los problemas del crimen y la violencia, promoviendo reformas que mejoren la seguridad y el bienestar de la sociedad.

**7. Preguntas Clave**

* ¿En qué año(s) han aumentado más los delitos?
* ¿Qué estados registran más delitos en ciertos meses?
* ¿Qué derechos (como la vida o la propiedad) se ven más afectados en ciertos meses?
* ¿Qué tipos de delitos tienden a aumentar en cada mes?
* ¿Hay subtipos específicos de delitos que crecen en algunos meses?
* ¿Qué formas de cometer delitos son más comunes en algunos meses?
* ¿En qué meses se concentran más delitos?
* ¿Qué estados y tipos de delitos muestran ciclos parecidos a lo largo del año?
* ¿Cuáles son los picos de delitos según la forma en que se cometen en cada mes?
* ¿En qué meses aumenta el impacto en derechos específicos como la vida o la propiedad?
* ¿Existen estados que tienen patrones similares de aumento o disminución de delitos en los mismos meses?
* ¿Qué subtipos de delitos deben ser prioridad en los meses con más actividad delictiva?

**8. Fuentes de Datos Identificadas**

* Datos de criminalidad
* Datos demográficos y socioeconómicos por entidad
* Factores socioeconómicos y políticos temporales
* Estadísticas de migración interna y externa
* Factores socioeconómicos y políticos temporales

**9. Justificación del Proyecto**

El análisis de las tasas de criminalidad en México es fundamental para ayudar a las autoridades a planificar estrategias de seguridad más efectivas. Al estudiar cuándo y dónde ocurren los delitos con mayor frecuencia, se pueden identificar patrones que permitan anticiparse a picos delictivos, como robos que aumentan en la temporada festiva o actos violentos que se intensifican en ciertos meses. Esto no solo permite mejorar la asignación de recursos, como incrementar la presencia policial en momentos clave, sino también lanzar campañas preventivas que puedan reducir la incidencia delictiva.

Además, al desglosar los delitos por tipo, subtipo y cómo se cometen, las autoridades pueden enfocarse en aquellos que afectan más gravemente derechos como la vida, la propiedad o la seguridad pública. Este tipo de análisis brinda una perspectiva más clara para ajustar las acciones según la realidad de cada región y tipo de crimen, lo que resulta en estrategias más precisas y eficaces.

Finalmente, el análisis no solo beneficia a las autoridades, sino también a la ciudadanía, que verá reflejadas mejores políticas de protección en su vida diaria. Asimismo, los legisladores podrán utilizar estos hallazgos para crear leyes más adecuadas a las necesidades actuales, y los medios de comunicación podrán sensibilizar mejor a la población sobre los riesgos en diferentes momentos del año.

**10.- ¿Cuántos datos y que tipo son?**

RangeIndex: 101830 entries, 0 to 101829

Data columns (total 21 columns):

# Column Non-Null Count Dtype

--- ------ -------------- -----

0 Año 97626 non-null object

1 Clave\_Ent 97626 non-null object

2 Entidad 97626 non-null object

3 Bien jurídico afectado 97626 non-null object

4 Tipo de delito 97626 non-null object

5 Subtipo de delito 97626 non-null object

6 Modalidad 97626 non-null object

7 Enero 97626 non-null object

8 Febrero 97626 non-null object

9 Marzo 97626 non-null object

10 Abril 97626 non-null object

11 Mayo 97626 non-null object

12 Junio 97626 non-null object

13 Julio 86763 non-null object

14 Agosto 86757 non-null object

15 Septiembre 86741 non-null object

16 Octubre 86678 non-null float64

17 Noviembre 86747 non-null object

18 Diciembre 86755 non-null object

19 Sexo/Averiguación previa 97626 non-null object

20 Rango de edad 97626 non-null object

dtypes: float64(1), object(20)

memory usage: 16.3+ MB