

Nombres:

Patricia García Olmeda 1931549
Lindsey Zugey Alejandro Castillo 1676950
Jorge Adrian Gonzalez Peña 1821695

Fecha:

30/Oct/2020

San Francisco Crime Dataset

Primer Avance de PIA

Equipo # 09

1) Titulo de la Base de datos

Nombre: San Francisco Crime Dataset

URL: <https://www.kaggle.com/roshansharma/sanfrancisco-crime-dataset>

2) Descripción de la Base de Datos

Nuestra base de datos cuenta con un total de 10 columnas, en las cuales se incluye:

1. **IncidentNum:** Numero asignado al crimen sucedido.
2. **Category:** Categoría del crimen sucedido.
3. **Descript:** Descripción del crimen.
4. **DayOfWeek:** Día de la semana en que ocurrió el crimen.
5. **Date:** Fecha en que ocurrió el crimen.
6. **Time:** Hora en que sucedió el crimen.
7. **PdDistrict:** Distrito donde ocurrió el crimen.
8. **Resolution:** Tipo de sanción o castigo que se le dio al criminal.
9. **Adress:** Dirección del crimen sucedido.
10. **X:** Latitud de la ubicación del crimen.

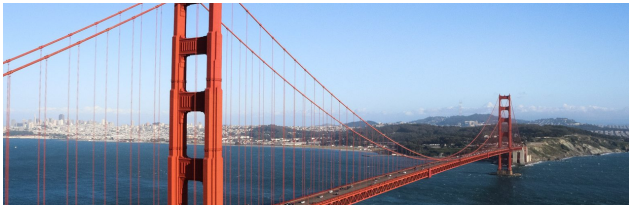
Nuestra base de datos cuenta con tres tablas de frecuencia para las columnas de Date, Time y X, donde nos facilitara poder encontrar patrones o similitudes entre los datos.

3) Justificación de Datos:

Decidimos utilizar esta base de datos ya que, además de llamar nuestra atención, la consideramos clara y bien redactada. Creemos que tiene los datos necesarios para poder trabajarla y poder cumplir nuestros objetivos los cuales presentaremos mas adelante.

Como beneficios nos resulto útil tener ya las ubicaciones y frecuencia de los crímenes sucedidos, ya que con esto tenemos un mejor orden y vista de lo que estamos trabajando y podremos llegar a mejores resultados.

4) Planteamiento del problema:



Conjunto de datos de delitos de San Francisco

Objetivo: Mejorar el sistema de seguridad de la ciudad de San Francisco y minimizar su tasa de criminalidad identificando lugar y causa de dónde se es más propenso sufrir delitos.

Problema planteado: La tasa general de delincuencia en San Francisco es un 151% más alta que el promedio nacional. Observamos que aunque la ley está presente el número total de delitos año tras año en la ciudad no ha cambiado. Y es que por cada 100,000 personas, hay 18.86 delitos diarios que ocurren en San Francisco.

Solución: Buscar relación entre los datos de la base de datos (distritos, direcciones, días de la semana, hora, categoría de delito y tipo de castigo) para determinar los lugares con más tendencia a sufrir delitos y entender ¿por qué? suceden. Esto permitirá distinguir locaciones clave para disponer de más seguridad policial, entre otros recursos.

5) Objetivo Final:

Nuestro objetivo final es encontrar , con ayuda de la base de datos, los momentos, lugares con mayor ocurrencia de delitos para así poder tomar acciones mas especificas y obtener un mejor resultado, tratando así de disminuir la tasa de delincuencia en San Francisco.

6) Planeación de la herramienta a utilizar:

Para lograr la solución a nuestro problema planteado usaremos la herramienta de Categoría y Visualización, organizaremos nuestros datos para tener una mejor visión sobre estos, observar si se encuentran patrones y así poder obtener los resultados de las zonas con mas crímenes y poder tener una mayor certeza sobre los futuros asaltos que podrían suceder y una mejor resolución ante el crimen.