

¿Quieres ir a E.U?

INTEGRANTES Equipo 6

Gonzalez Barron B. A. **1620548**
Lince Alvarez J. E. **1745066**
Cavazos Rodriguez G. A. **1742463**

INTRODUCCIÓN

Datos de Violencia Armada

Este proyecto tiene como objetivo principal el análisis de los datos recaudados en el data set Gun Violence Data con un registro de más de 260K de incidentes de violencia armada en EE. UU. Entre el 2013 y 2018.

Los datos se descargaron de gunviolencearchive.org Gun Violence Archive (GVA) es una corporación sin fines de lucro formada en 2013 para proporcionar acceso público gratuito en línea a información precisa sobre la violencia relacionada con armas de fuego en los Estados Unidos.

Hipótesis al observar los datos del data set

¿El número de armas de fuego involucradas influyen en el estatus final del participante?

OBJETIVO

Al estudiar este conjunto de datos se pretende poder hacer deducciones con base al comportamiento de estos, en el aspecto de qué lugares son más peligrosos o propensos en los que las personas puedan sufrir de un atentado, además con base a sus características predecir el estatus del participante. Una vez teniendo los datos que agrupamos datos de tal manera que podamos usar dichas agrupaciones para la obtención de información que nos ayudara a solventar nuestro objetivo principal plasmamos Tops y comportamiento por medio de gráficas de distintos tipos que nos ayudaron a mostrar datos que responda las preguntas planteadas.

RECURSOS

Base de Datos (Gun Violence Gun)

Para esta investigación la base de datos fue obtenida a través del sitio gunviolencearchive.org del cual se descargaron más de 260K registros de incidentes que se sucedieron entre los años del 2013 al 2018 y quien subió la base de datos a la plataforma Kaggle fue el usuario James Ko. Como se mencionó anteriormente utilizamos la plataforma Kaggle para la obtención de la base de datos en formato CSV.

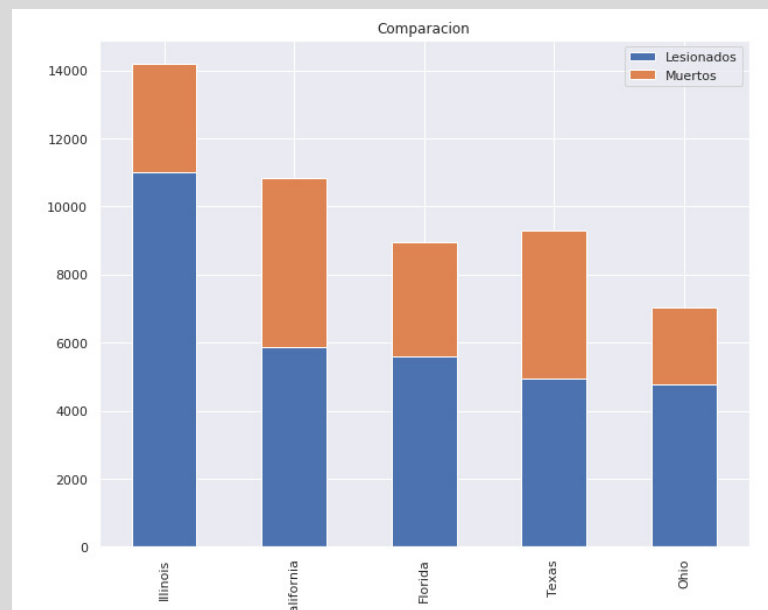


Lenguaje de Programación y entorno de trabajo



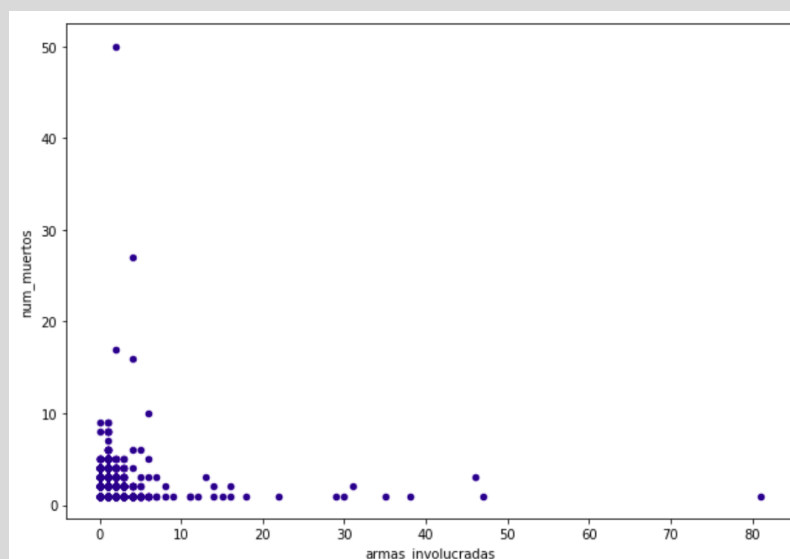
RESULTADOS

A pesar de contar con una base de datos muy extensa y robusta, se pudo encontrar información esperada que se encontraba inmersa dentro de el gran conjunto de datos que representaba el data set que hablamos elegido con mas de 260K registros de incidentes. Dicha información la pudimos visualizar mediante distintas gráficas que muestran el comportamiento de los datos según las características estudiadas en cada una de ellas.



(Gráfica 1) Estados más peligrosos según el índice de mortalidad

En E.U. existen 5 estados que destacan por ser los más peligrosos del país con respecto el numero de lesionados y de muertos por incidente(Mas de uno por caso) en el periodo del CO. Como nota extra, estos estados también son los que más incidentes presentaron.

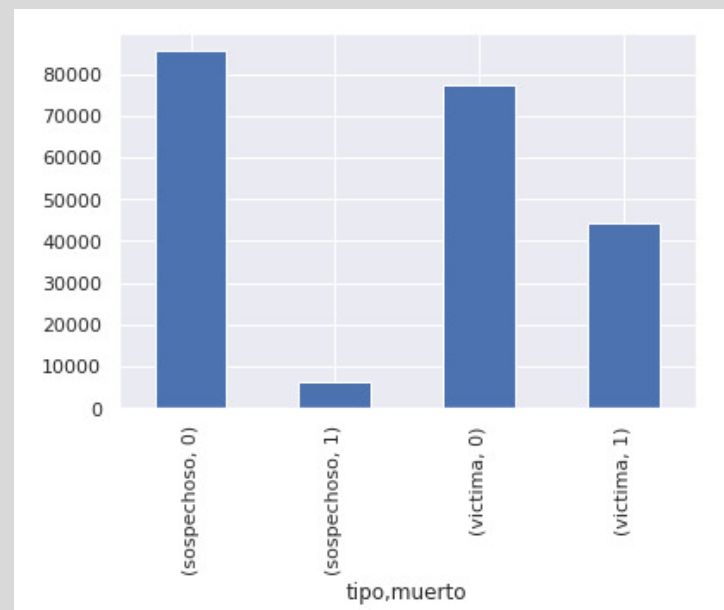


(Gráfica 2) Número de muertos respecto a las armas de fuego involucradas

Vemos que existen incidentes los cuales hay un gran número de armas involucradas pero el numero de muertos es muy bajo, sin embargo existen casos en los que el número de armas involucradas es bajo y el número de muertos es mucho más alto; también se observan algunos casos aislados los cuales son los que están alejados del grupo donde se encuentran más concentrados los datos.

Además poniendo en comparación los incidentes generales con los incidentes de los estados más peligrosos no hay una gran diferencia de la cantidad de armas de fuego utilizadas.

	Estados Peligrosos	General
Sin Armas	52%	56%
1 Arma	42.6%	38.9%
2+ armas	5.4%	5.1%



(Gráfica 3) Número de participantes por tipo y estatus

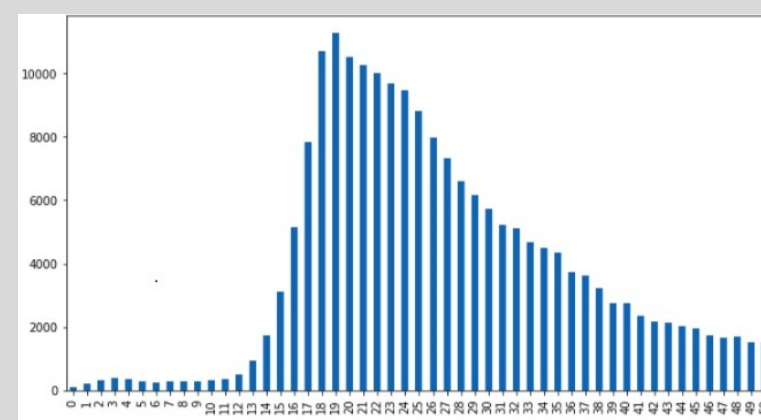
Podemos observar que el número de victimas muertas es mayor al número de sospechosos muertos por lo el indicador nos dice los participantes que son de tipo víctima tienden a morir.

	Victimas	Sospechosos
Arrestado	1%	75%
Muerto	36%	6%
Herido	52%	4%
Desconocido	11%	15%



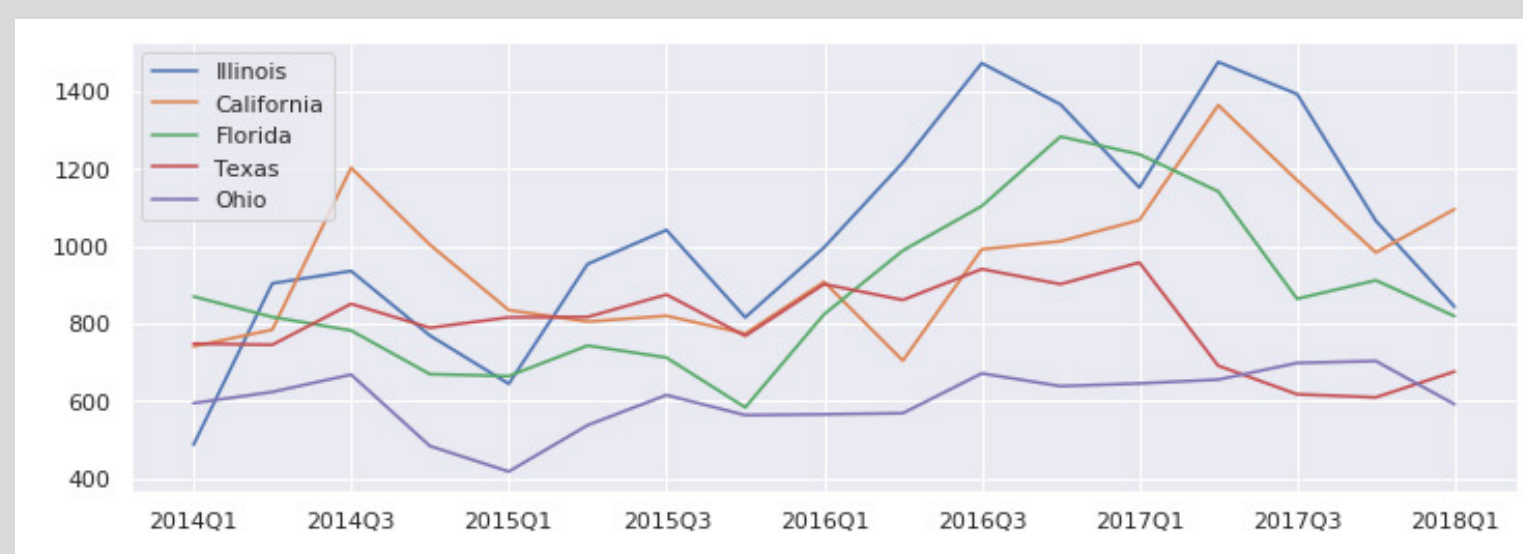
(Gráfica 4)

A pesar de que el hay mas mujeres victimas en proporción con los hombres no se ve una diferencia en la tasa de mortalidad.



(Gráfica 5) Las edades de los participantes

Sin importar el tipo del participante, ya sea sospechoso o víctima, podemos observar que las edades se concentran entre los 16 y los 30 años, viendo que los participantes en su mayoría sin discriminar su tipo son jóvenes en los incidentes en todo los Estados Unidos.



(Gráfica 6) Tendencia de incidentes en el Top 5 de Estados más peligrosos en E.U.

El comportamiento de los datos en esta gráfica en los estados más peligrosos nos muestra la tendencia, en los que se puede apreciar que hay alzas y bajas de incidentes en cada uno de los trimestres a lo largo de la línea del tiempo. Además en esta tendencia podemos observar que los trimestres o el periodo más crítico para los 5 estados en su mayoría esta entre los trimestres del 2016 al 2017.

METODOLOGÍA

Se empezó por identificar las columnas que necesitaríamos para nuestro análisis, entre las cuales estaban la fecha de cuando sucedió el hecho, número de muertos, la edad de los participantes, etc.

Entre las columnas que se eliminaron son 'arma_robada' o 'sources' que en su mayoría eran datos nulos o no se ocuparían, además se eliminó el nombre de los participantes para mantener el anonimato. Se eliminan los datos nulos por cero de la columna armas_involucradas para que se pueden hacer búsquedas con este campo.

El CO se dividió en dos, teniendo un CO para los incidentes con 239677 y otro CO para los participantes, teniendo en total 213754 registros, esto sucede debido a que no siempre, el registro de los incidentes tenía la información completa de los participantes, pero si que tenía la información de muertos, heridos etc. Por eso no se eliminaba ese incidente.

También, en los participantes se encontró que en genero solo hay dos casos, Male (Hombre) y Female (Mujer), además en el tipo, hay dos casos, los cuales son Victim (víctima) y Subject-Suspect Y en estatus son varios casos pero solo se toman tres fundamentales, Arrestado, Muerto, Herido.

CONCLUSIONES

No pudimos encontrar una relación sólida entre las características de los participantes y el estatus final de los mismos (Herido, Muerto ó Arrestado) ya que las características disponibles que teníamos como edad género y tipo no parecía tener un patrón que nos ayudara a determinar su situación final; ya que por ejemplo, a pesar de que había menos mujeres víctimas que los hombres, la proporción de muertos y los sobrevivientes son parecidos.

Encontramos que los estados más peligrosos son: Illinois, California, Florida, Texas y Ohio, siendo estos los que más incidentes presentan, los que mas muertos y heridos por incidente tienen, pudiendo concluir que estos son los estados más arriesgados o conflictivos.

También no encontramos un patrón entre las armas de fuego y el número de muertos por incidente ya que por un lado los incidentes con muchas armas no sobrepasaban los 4 muertos o incidentes con 1 o 2 armas de fuego tenían hasta 50 muertos siendo estos casos de masacre.

TRABAJO A FUTURO

En esta investigación no pudimos que aplicar una técnica de predicción que pudiera encontrar el estatus final de un participante dadas las características del mismo y del incidente; debido a que no encontramos alguna relación que nos pudiera guiar para poder predecirlas.

En un futuro se planea buscar una base de datos que nos proporcione características que nos faciliten la aplicación de una técnica de predicción, por ejemplo: la raza a la que pertenece el participante (para tomar en cuenta los factores racistas de actos delictivos), orientación sexual, entre otros.