Universidad del Rosario Ingeniería de Datos Profesor: Javier Casas Brenda Nicole Arcos Varón David Andrés Lara Galindo Sebastián Rodríguez Patiño Sergio Andrés Casa Pinilla 2023-2



## **Entrega dos**

Bases de datos:

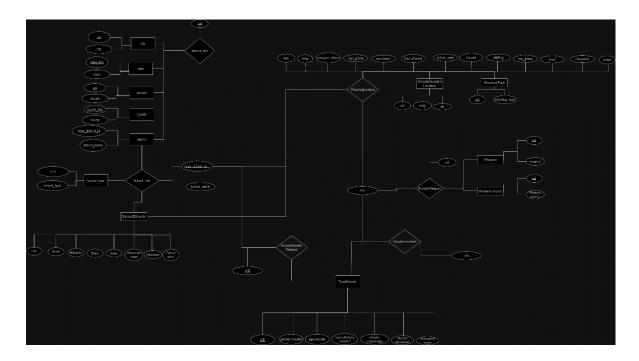
• Nombre de la base: School Shootings (Tiroteos Escolares)

Siguiendo con la línea de trabajo propuesta en la primera entrega, se corrigieron errores presentes en los diagramas originales. Las acciones tomadas fueron las siguientes:

- En la relación anterior, hay entidades como "Lugar", "evento", "shotter", "shotter2", "weapon", etc. Estas entidades se han desglosado en varias tablas más pequeñas como "School\_info", "school\_type", "SchoolEthnicity", "district\_info", "ShooterIncidentType", etc. Y tienen atributos como "nces\_school\_id", "school\_name", "nces\_district\_id", "school type" en la entidad "School info".
- Se eliminó el atributo 'year' de la tabla 'School', ya que la conexión se podía hacer mediante consultas utilizando otras tablas, Además, este atributo no aportaba a la solución de los escenarios de análisis.
- Antes había relaciones entre entidades como "shooter" y "weapon", "shooter" y "shotter2", etc. Y ahora las relaciones están representadas en entidades separadas como "ShooterWeapon", "ShooterIncidentRelation", "SchoolShooterRelation", etc.
- La entidad "race" con atributos como "w", "b", "h", "a", "ai", "hawaiian", "two or more". Se divide en la entidad "SchoolEthnicity" con nombres más descriptivos.

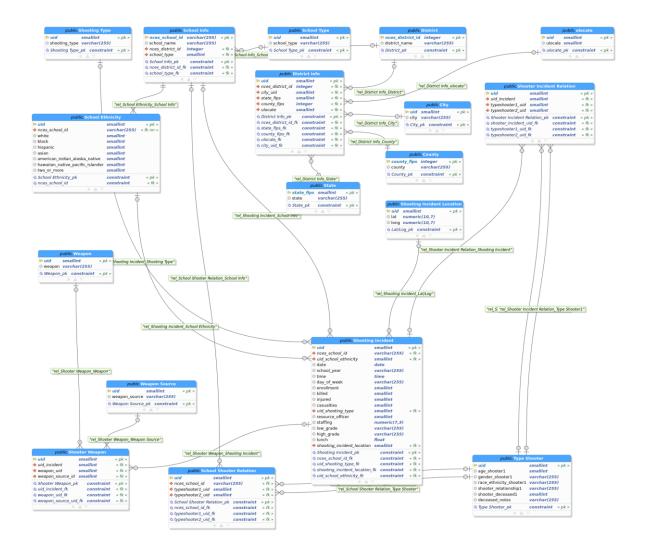
Los resultados fueron:

Diagrama entidad-relación



# Normalización Diagrama Relacional

Para este proceso, se utilizó la base de datos original en formato CSV y se empleó Excel para convertirla en una tabla. A continuación, se aplicó el proceso de normalización paso a paso para finalmente generar el diagrama relacional que se había adelantado previamente. El resultado obtenido es el siguiente:



# Carga de información en base de datos

Una vez que se normalizó tanto la tabla en Excel como el diagrama relacional, se transfirió toda la información a pgAdmin4 como una base de datos con el siguiente proceso:

- 1. Se definieron las tablas según las condiciones establecidas en el diagrama relacional previamente normalizado. Se tuvieron en cuenta todos los atributos, llaves primarias, llaves foráneas y restricciones como "unique" en caso de que estuvieran presentes.
- 2. Cada una de las tablas resultantes de la normalización se dividió en archivos separados en formato CSV UTF-8. Esto se hizo para evitar problemas al leer los archivos en pgAdmin4.
- 3. Una vez que se obtuvieron todos los archivos, se utilizó la función "copy" dentro de pgAdmin4. Esta función permite copiar los datos de un archivo CSV y pegarlos en una tabla de la base de datos. Se realizó una instrucción para cada tabla.
- 4. Finalmente, se ejecutaron todas las instrucciones para completar el proceso de carga de

información y obtener la base de datos creada con todos los datos cargados en SQL.

#### Análisis de escenarios

Con base en los datos que tenemos a nuestra disposición e información estadística obtenida de estudios anteriores realizados a la base de datos, se proponen los siguientes escenarios de análisis:

• Investigar las características de los perpetradores, como su edad, género, antecedentes de salud mental, armas utilizadas, relación con la víctima e historial de comportamiento puede ser fundamental para comprender las causas de los tiroteos en escuelas y desarrollar estrategias de prevención. Estos perfiles pueden ser utilizados por autoridades y profesionales para tomar medidas y mitigar el riesgo de futuros incidentes.

**Metodología:** Se van a analizar las siguientes tablas:

- -TypeShooter: Proporciona detalles sobre los perpetradores, incluyendo edad, género, antecedentes de salud mental y otros atributos personales.
- -ShooterIncident: Ofrece información detallada sobre cada tiroteo, incluyendo las armas utilizadas, la relación con la víctima y el historial de comportamiento del perpetrador en ese incidente.
- ShooterWeapon: contiene información sobre las armas utilizadas por los perpetradores en los tiroteos, lo que puede ayudar a comprender la naturaleza de la violencia

#### Utilidad

Facilita la comprensión de perfiles de perpetradores, ayudando a desarrollar estrategias preventivas específicas dirigidas a grupos de riesgo, lo que contribuye a la mitigación de futuros incidentes.

• ¿Cómo influyen los factores socioeconómicos y el entorno tanto familiar en la ocurrencia de tiroteos escolares? ¿Existe una relación temporal entre eventos de violencia armada y la implementación de ciertas políticas de control de armas?

**Metodología:** Se van a analizar las siguientes tablas:

- -School\_info: analizar la distribución de tiroteos en función del tipo de escuela (pública, privada, etc.) y el tamaño del distrito escolar.
- -SchoolEthnicity: evaluar si la diversidad étnica en las escuelas está relacionada con la frecuencia de tiroteos.

- -ShootingIncident: Contiene información detallada sobre cada incidente de tiroteo en las escuelas.
- TypeShooter: analizar las características personales de los agresores en tiroteos escolares. Examina variables como edad, género, relación con la víctima y correlaciona las características de los agresores con el desarrollo del incidente.

### **Utilidad**

Permite identificar posibles correlaciones entre condiciones socioeconómicas, entorno familiar y escolar, y la ocurrencia de tiroteos escolares. Y ayudará a comprender mejor las características personales y relaciones sociales de los agresores en tiroteos escolares, lo que puede informar estrategias de prevención y programas de intervención temprana en entornos educativos

• El impacto del uso de armas de familiares en los tiroteos y la relación con el fácil acceso a armas en Estados Unidos

**Metodología:** Se van a analizar las siguientes tablas:

- -ShootingIncident: Proporciona información detallada sobre cada tiroteo, incluyendo la fecha, ubicación, y detalles sobre las armas utilizadas y permite identificar si el arma utilizada pertenecía a un familiar del perpetrador.
- -ShooterWeapon: Contiene información específica sobre las armas utilizadas en cada tiroteo, incluyendo el tipo y origen de las armas y Permite rastrear si las armas pertenecen a familiares del perpetrador.
- TypeShooter: Proporciona detalles sobre los perpetradores, incluyendo su relación con la víctima y características personales y permite identificar si el perpetrador tiene acceso a armas de familiares y si esto influyó en el incidente.
- -Weapon: Ofrece información general sobre los tipos de armas y su origen y puede ser útil para comprender el fácil acceso a ciertos tipos de armas en Estados Unidos

#### Utilidad

Proporciona información valiosa para entender cómo el acceso a armas, especialmente aquellas de familiares, podría influir en la frecuencia y gravedad de los tiroteos. Esta comprensión puede orientar políticas y medidas preventivas para abordar específicamente este aspecto.

• Investigar en que estados, ciudades e instituciones es más probable que ocurran tiroteos, con el fin de encontrar los lugares más propensos y así las autoridades puedan tomar medidas adicionales para evitar este tipo de tragedias

Metodología: Se van a analizar las tablas las siguientes tablas:

- -ShootingIncident: Permite identificar los estados y ciudades donde son más probables los tiroteos.
- -SchoolInfo: Facilita el análisis de tiroteos específicamente en entornos escolares.
- -DistrictInfo: Proporciona información a nivel de distrito, incluyendo el nombre del distrito y la ubicación y puede ser útil para identificar patrones de tiroteos a nivel de distrito escolar.
- -City: Ayuda a examinar los tiroteos en relación con las características de diferentes ciudades.

#### Utilidad

Facilita a las autoridades locales y nacionales la identificación de áreas con mayor riesgo de tiroteos, permitiendo la implementación de medidas específicas y recursos adicionales para la prevención de tragedias.

# Módulos de Python:

https://colab.research.google.com/drive/1JNGdfaQCzXC-nJ1giQgRUUYAEO3YI4O4?usp=sharing

### **Anexos**

**Origen de la base de datos:** <a href="https://www.kaggle.com/datasets/joebeachcapital/school-shootings-data.csv">https://www.kaggle.com/datasets/joebeachcapital/school-shootings-data.csv</a>

### Repositorio:

https://github.com/errcr-exe/School-Shooting-Analysis