**Project 2 ETL- Gun Violence**

Group Members: Matt Yurich, Allison Dorion, Swapna Kasturi

For our project, we decided to create a database and provide information on Gun Violence to address whether gun control or gun education is a more viable solution. Gun violence is huge issue for debate in the United States currently and we wanted to provide information and data to help provide a possible solution.

**Extract:**

We used 3 different datasets found using the platform Kaggle and led us to the Gun Violence Archive. The data in each file contains the following information:

* Accidental Death
* Mass Shootings
* Accidental Injuries

The sources that we found our datasets are:

<https://www.kaggle.com/gunviolencearchive/gun-violence-database>

<https://www.gunviolencearchive.org/reports>

<https://www.gunviolencearchive.org/mass-shooting>

<https://www.gunviolencearchive.org/accidental-deaths>

We extracted our data using csv files found in the Kaggle platform and formatted it using the pgAdmin 4.

Some of the following columns and data used are:

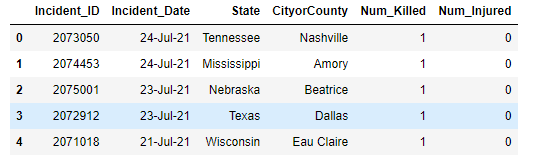
* Incident ID
* Incident Date
* State
* City/County
* Number Killed
* Number Injured

**Transform:**

To transform the data we extracted from Kaggle to study we performed the following:

* Used pandas in Jupyter Notebook to load the three CSV files we used.
* The CSV files were then transformed into panda dataframes
* The data was cleaned and included only the columns we wanted to be used in our database and the column names were made more clear.

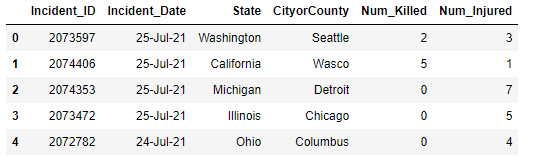
**Accidental Deaths:**



**Accidental Injuries**:

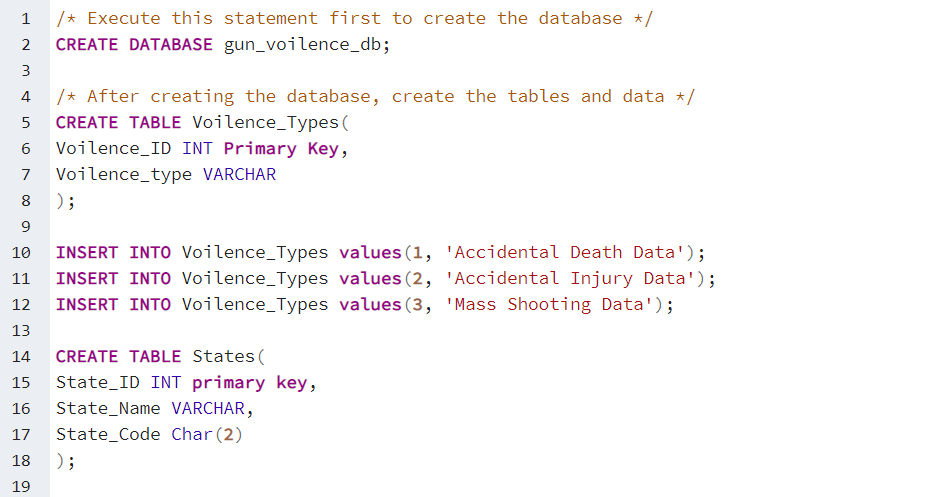


**Mass Shootings:**

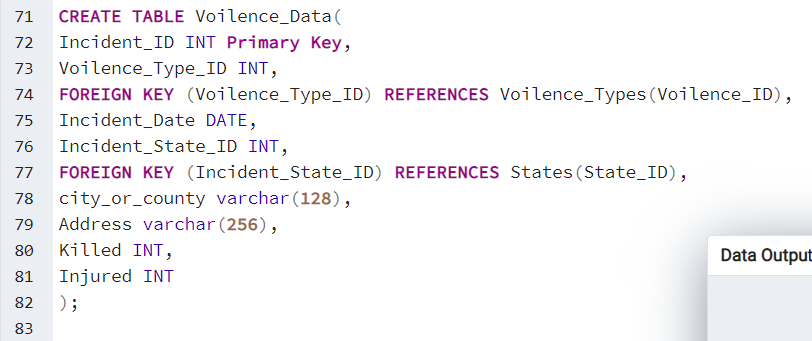


**Load:**

Once the CSV files were loaded into the data frames in Jupyter Notebook and transformed, the notebook was then connected to the Postgres database using PG Admin 4 to store the clean data sets. We created the initial table schema that was uploaded to the PostGres database and generated the original tables into the database. All the clean data frames created with all relevant information were loaded into the new schema tables.







**Summary:**

After extracting, transforming and loading the data, there is now a database that covers some of the topics of gun violence. Using our database, accidental deaths, accidental injuries, and mass shootings data can be used to educate and help create solutions for the continued debate on gun violence.