

Tipologia de Dades, Pràctica 1

Ramon Maria Gallart Escolà

Tipologia de Dades
Màster en Ciència de Dades
Mireia Calvo González
April 7, 2018

Índex

TÍTOL DEL DATASET	2
SUBTÍTOL DEL DATASET	2
IMATGE	2
CONTEXT	2
CONTINGUT	3
SEMPRE DISPONIBLES	3
OPCIONALS	3
AGRAÏMENTS	4
INSPIRACIÓ	4
LLICÈNCIA.....	5
CODI.....	6
DATASET	6
RECURSOS CONSULTATS	7

Títol del Dataset

Violència amb armes de foc als Estats Units d'Amèrica

Subtítol del dataset

Tirotejos massius

Imatge



Context

El conjunt de dades es correspon als tirotejos massius ocorreguts als Estats Units d'Amèrica tal i com l'FBI defineix un tiroteig massiu que és *“QUATRE persones o més que hagin estat disparades i/o mortes en un sol esdeveniment {incident}, en el mateix lloc i al mateix temps en general, no incloent el que ha disparat”*. Per tant, la diferenciació entre un tiroteig i un tiroteig massiu és únicament numèrica i no es té en compte si l'esdeveniment ha estat un accident, una baralla entre bandes o violència domèstica.

Contingut

Per cada tiroteig hi ha dades que estan sempre disponibles i algunes que poden o no estar-hi (però que acostumen a estar).

Els autors de la web obtenen les seves dades a partir de recursos policials i de mass-mèdia diàriament i registra incidents amb fins a 120 variables potencials dels incidents. Cada incident a la base de dades està enllaçat de forma independent amb la seva font o fonts que van ser emprades per generar l'informe de l'esdeveniment.

Sempre disponibles

- **Year:** any de l'esdeveniment
- **Month:** mes de l'esdeveniment
- **Day:** dia de l'esdeveniment
- **State:** estat en el que s'ha produït el tiroteig
- **City_or_county:** ciutat o comptat
- **Num_killed:** nombre de morts
- **Num_injured:** nombre de ferits
- **Incident_link:** enllaç amb informació de l'incident

Opcionals

- **Latitude:** latitud
- **Longitude:** longitud
- **Participants:** llista amb els participants i atributs dels participants. No tots els participants sempre tenen per què tenir els mateix nombre d'atributs
- **Characteristics:** conjunt de característiques que defineixen l'incident
- **Notes:** notes sobre l'esdeveniment
- **Guns_involved:** llista amb les armes de foc i els seus atributs que s'han vist involucrades en l'incident. No sempre tenen per què tenir els mateixos atributs
- **District:** districte en què l'incident ha succeït.

Agraïments

Gràcies a [1] per haver realitzat la tasca d'obtenció de les dades i d'haver-les posat a disposició del públic en general ([2]).

L'aplicació s'ha realitzat amb el llenguatge Python utilitzant tècniques de *Web Scraping* explicades a [3] i a [4].

Inspiració

Ja que fa més d'un any que estic vivint als Estats Units d'Amèrica he volgut realitzar la pràctica tocant un tema que és molt complex. Per una banda hi ha la 2a esmena [5] que garanteix als ciutadans americans el dret a dur armes i per l'altra els problemes que això comporta entre activistes per al dret a dur armes, els qui volen posar-hi límits i els tirotejos i morts que hi ha entre mig [6].

Gairebé cada dia hi ha aquí als USA almenys una notícia relacionada amb un incident en què almenys ha intervingut una arma de foc. Com un dels esdeveniments més recents tenim la massacre que es va produir a Florida el passat 14 de febrer on 17 persones, majoritàriament adolescents, van morir a trets mentre estaven a l'institut [7].

Amb el dataset producte de la pràctica es podrien realitzar estudis per encarar el problema i poder prendre decisions al respecte. Algunes de les idees a estudiar podrien ser, per exemple:

- Treure estadístiques descriptives (mitjana, desviació típica, etc.)
- Poder respondre a preguntes com per exemple:
 - Els tirotejos massius segueixen alguna tendència?
 - Augmenten o desminueixen amb el pas dels anys?
 - Quines són les zones, èpoques, etc. amb més incidents?
- Posar en el mapa els llocs on hi ha hagut els incidents de forma que generin un heat map [8].
- Poder buscar per zona, poble, estat, etc.
- Poder creuar les dades amb les legislacions aplicades en cada estat per saber si la legislació estatal afecta l'ús de les armes i, en cas de demostrar-ho, aplicar polítiques paliatives.
- Possibilitat de creuar les dades amb informació socio-laboral i econòmica del lloc dels incidents per trobar possibles correlacions entre violència per arma de foc i estatus social.

Llicència

Per a la llicència he escollit la GNU GPL v3 [9]. Crec que aquesta llicència va bé amb el seu contingut ja que, en general, permet:

- **Ús comercial:** el programari i els seus derivats poden ser usats amb finalitats comercials.
- **Distribució:** el programari pot ser distribuït.
- **Modificació:** igualment, el programari pot ser modificat.
- **Ús de patents:** la llicència garanteix drets de patent als contribuïdors del programari.
- **Ús privat:** el programari pot ser usat i modificat per ús privat.

Sota les condicions de:

- **Incloure el codi:** si el programari es distribueix, el codi sempre s'ha de posar disponible.
- **Donar a conèixer la llicència i el copyright:** una còpia de la llicència i del copyright sempre han d'acompanyar el programari.
- **Declarar els canvis:** s'han de documentar els canvis realitzats al codi.

I amb les limitacions de:

- **Responsabilitat:** aquesta llicència inclou una limitació de responsabilitat.
- **Garantia:** la llicència especifica explícitament que NO inclou cap garantia.

Codi

El codi està inclòs en aquest mateix repositori de Github sota el directori `src`.

Dataset

Un dataset d'exemple està inclòs en aquest mateix repositori de Github en el fitxer `out/20180407_output-2018.csv`.

Recursos consultats

- [1] «Mass Shootings | Gun Violence Archive». [En línia]. Disponible a: <http://www.gunviolencearchive.org/reports/mass-shooting>. [Consulta: 08-abr-2018].
- [2] «Gun Violence Archive». [En línia]. Disponible a: <http://www.gunviolencearchive.org/>. [Consulta: 08-abr-2018].
- [3] R. Lawson, *Web Scraping with Python*, 1st ed. Packt Publishing, 2015.
- [4] R. Mitchell, *Web Scraping with Python, 2nd Edition*, 2a ed. O'Reilly Media, Inc., 2018.
- [5] «Second Amendment to the United States Constitution - Wikipedia». [En línia]. Disponible a: https://en.wikipedia.org/wiki/Second_Amendment_to_the_United_States_Constitution. [Consulta: 08-abr-2018].
- [6] M. Valentine, «The Second Amendment Issue the Supreme Court Can't Avoid», *POLITICO Magazine*. [En línia]. Disponible a: <https://www.politico.com/magazine/story/2017/07/03/the-second-amendment-issue-the-supreme-court-cant-avoid-215331>. [Consulta: 08-abr-2018].
- [7] «Americans deserve to lose their 2nd Amendment right to bear arms», *latimes.com*. [En línia]. Disponible a: <http://www.latimes.com/opinion/readersreact/la-ol-le-florida-school-shooting-guns-20180215-story.html>. [Consulta: 08-abr-2018].
- [8] «Heat map», *Wikipedia*. 06-abr-2018.
- [9] «GNU Lesser General Public License v3.0», *Choose a License*. [En línia]. Disponible a: <https://choosealicense.com/licenses/lgpl-3.0/>. [Consulta: 08-abr-2018].