# KILLING USA

## HOMICIDIOS Y ASESINOS EN SERIE CON SQL (Y UN POCO DE PYTHON)

ALONSO VALBUENA - BOOTCAMP DATA ANALYTICS IMMUNE INSTITUTE

# INTRODUCCIÓN

Desde que leí "A sangre fría" de Truman Capote me ha fascinado el true crime. A día de hoy soy un ávido consumidor de podcast, documentales y blogs sobre crímenes y sus resoluciones. Es por eso que me interesaba realizar un proyecto en el que ponerme, por un momento, en la piel del detective, buscar patrones, señalar sospechosos, resolver casos.
Aunque este trabajo está muy lejos de ser la última de mis incursiones en el mundo de los datos sobre crímenes, quería probar a zambullirme de lleno y sin salvavidas en el fango. Ha sido muy gratificante.

# METODOLOGÍA

Este trabajo está dividido en dos partes:

- En la **primera parte**, analizo un .csv de el proyecto **Murder Accountability Project** (MAP), que contiene más de 800.000 filas con datos individuales de **homicidios** en USA, desde **1976** hasta ahora. El Murder Accountability Project es una organización sin fines de lucro que se dedica a recopilar y analizar datos sobre homicidios en los Estados Unidos. La mayoría de sus datos son del FBI pero ponen especial énfasis en la oficialidad de todos ellos.
  El .csv se puede encontrar aquí y, entre otras cosas, también tenemos un diccionario que nos ayuda a encontrar el contenido de cada columna. Operaremos principalmente con estas:
  - **State:** El estado en que ocurre el homicidio
  - **Solved:** Puede ser "Yes" o "No" para indicar si el homicidio está o no resuelto
  - **year_crime**: El año en que ocurre el crimen. En el .csv que nos ocupa va de 1976 a 2021, la última actualización.
    También tenemos una columa **Month**, pero no la usaré en este estudio en concreto.
  - **VicAge/OffAge:** Estas dos columnas contienen la edad de la víctima y el agresor. Los valores van de 0 a 99 (99 vale para todas aquellas de >99 años) y, en los casos en los que la edad es indeterminada (por no conocerse el agresor o por no poder identificar con exactitud la edad de la víctima), el valor usado es 999.
  - **Circumstances:** Representa una miríada de valores que clarifican las circunstancias del homicidio, como "Lovers Triangle", "Robbery", "Juvenile gang killings", etc.

- VicSex/OffSex: Estas dos columnas indican el sexo de la víctima y el agresor. 3 Valores
  - Male
  - Female
  - Unknown
- VicRace/OffRace: Similar a las anteriores. Valores:
  - Native American
  - Black
  - White
  - Asian
  - Pacific islander
- Weapon: El arma usada en el crimen. Los valores son:
  - Firearm, type not stated
  - Handgun – pistol, revolver, etc.
  - Rifle
  - Shotgun
  - Other gun
  - Knife or cutting instrument
  - Blunt object – hammer, club, etc.
  - Personal weapons, including beating
  - Poison, does not include gas
  - Pushed or thrown out of window
  - Explosives
  - Fire
  - Narcotics or drugs, sleeping pills
  - Drowning
  - Strangulation or hanging
  - Asphyxiation – includes death by gas
  - Other or type unknown weapon

En esta primera parte, en la parte final vinculo la tabla del MAP con otra, encontrada aquí, sobre presidentes de los USA para encontrar los **homicidios** totales **por legislatura**.

- En la segunda parte, **vinculo** esa primera tabla con datos sobre asesinos en serie encontrados en internet que transformé a tabla con una **mezcla de web scrapping** para crear la estructura inicial de la tabla **y limpieza y relleno a mano** de columnas que requerían una lectura más subjetiva de los métodos y edades de las víctimas de cada asesino en serie.
  - El código para el web scrapping lo usé primero sobre las diferentes entradas de esta página: https://en.wikipedia.org/wiki/Category:American_serial_killers para crear una tabla que contuviera dos columnas: nombre del asesino y URL de su página personal. Luego, volví a iterar sobre esa tabla para extraer, de cada página individual, los años que estuvo cometiendo crímenes y los estados en los que actuó. Acabé con una tabla con Nombre, URL, rango fechas, y estados de cada uno de los asesinos. El código completo se puede encontrar aquí.
  - Tras limpiar los datos resultantes (algunos caracteres daban problemas y el rango de fechas fue dividido en fecha_inicio y fecha_fin de los crímenes), me puse a rellenar a mano las columnas restantes, extrayendo datos tanto de sus páginas de wikipedia como de esta y esta otra web, bastante completas y útiles. Acabé con las columnas **full_name_sk** (el nombre del serial killer), **URL** (la URL a su página de wikipedia), **start_crimes/end_crimes** (columna fecha inicio / columna fecha fin de sus crímenes), **state** (estado principal de sus crímenes), **state2** (estado secundario – si lo hubiera - ), **victimrace** (raza más común de las víctimas y, en caso de que no hubiera una, "in" por indiferente), **victimsex** (sexo más común de las víctimas, mismas condiciones que la categoría anterior), **victimagemin/victimagemax** (columnas de edad mínima y máxima de sus víctimas), **weapon/weapon2/weapon3** (3 columnas con el método de asesinato más usado, hasta 3 métodos si los hubiera), y **victims** (numero estimado de víctimas).
  - Ciertas condiciones y limitaciones que encontré:
    - Eliminé todos los asesinos que mataron hasta antes de 1976 ya que no eran cubiertos por las estadísticas del MAP
    - Eliminé todos (o casi todos) los asesinos que actuaron en más de 3 estados, ya que ampliando tanto la zona de búsqueda y a este nivel de análisis iban a dar muchos falsos positivos al emparejarlos con la tabla del MAP.
    - Eliminé todos los asesinos con un perfil de víctima o de modus operandi aleatorio, ya que no hay forma de cotejarlos a este nivel (si mataban indiscriminadamente o con métodos muy variados)
    - Eliminé todos los asesinos en serie que actuaban sobre todo actuando como enfermeros contra pacientes, ya que están muy localizados en la institución en la que trabajaran.
    - Eliminé todos aquellos para los que no hay suficientes datos de acceso público como para precisar el perfil de víctima o M.O.

Este proyecto podría haber sido el triple de largo, pero consideremos esto un testeo. Empecemos.

# PARTE 1. ESTADÍSTICAS DE ASESINATOS.

Siendo el schema **map_usa** y las tablas **shr76_21** (los datos de MAP), **us_presidents** (la tabla con los presidentes) y **sk_dt** (asesinos en serie).

## →¿CUÁNDO SE MATÓ MÁS?

Saquemos un **top 5**:

```
SELECT year_crime as Year, COUNT(*) AS total_murders
FROM `map_usa`.`shr76_21`
GROUP BY Year
ORDER BY total_murders DESC
LIMIT 5 ;
```

| Year | total_murders |
|---|---|
| 1993 | 24337 |
| 1992 | 23793 |
| 1994 | 23246 |
| 1980 | 23092 |
| 1991 | 22657 |

A excepción del año 80, **los primeros años de los 90 parecen ser los más sangrientos**. A día de hoy, los criminólogos no se ponen de acuerdo en el por qué comienzan a bajar a partir de mediados de los 90. Algunos lo achacan al fin de los combustibles con plomo (el plomo, que se quedaba en el aire por su combustión, es conocido por aumentar los comportamientos violentos), la guerra contra las drogas, la posibilidad del aborto (que hizo que hubiera menos hijos no deseados que pueden acabar teniendo familias desestructuradas), la aparición de los videojuegos (que evitaron que los jóvenes estuvieran tanto en la calle causando problemas).... No hay una teoría concreta.

## →¿DÓNDE SE MATA MÁS? ¿Y MENOS?

Vamos a crear dos views, dos **top 10** (los estados en los que más y en los que menos se mata) para luego hacer **subconsultas**:

```
CREATE VIEW low_10 AS
SELECT State, COUNT(*) AS total_murders
  FROM `map_usa`.`shr76_21`
 GROUP BY State
 ORDER BY total_murders ASC
 LIMIT 10 ;
```

```
CREATE VIEW top_10 AS
SELECT State, COUNT(*) AS total_murders
  FROM `map_usa`.`shr76_21`
 GROUP BY State
 ORDER BY total_murders DESC
 LIMIT 10 ;
```

¿Qué resultado dan?

| State | total_murders |
|---|---|
| California | 125258 |
| Texas | 81731 |
| New York | 61235 |
| Florida | 49516 |
| Illinois | 37381 |
| Michigan | 37042 |
| Pennsylvania | 31739 |
| Georgia | 27915 |
| North Carolir | 27495 |
| Ohio | 26809 |

*Figura 1: top_10*

| State | total_murders |
|---|---|
| North Dakota | 522 |
| Vermont | 551 |
| South Dakota | 731 |
| Wyoming | 877 |
| New Hampsh | 899 |
| Montana | 1030 |
| Maine | 1191 |
| Rhode Island | 1533 |
| Idaho | 1650 |
| Delaware | 1783 |

*Figura 2: low_10*

←←←←←←←←←←¿Qué pasa si comparamos estos resultados con el ranking de estados más y menos poblados de USA (referencia aquí)?

**top_10** da un ranking que coincide al 100% (están los mismos estados), pero **low_10** tiene una ausencia notable: Alaska debería estar por población (es uno de los menos poblados), pero en su lugar se encuentra Idaho.
En cualquier caso, el orden en los ranking por asesinatos no coincide con el orden en el ranking de población (el 3º estado con más asesinatos no es el 3º estado más poblado, por ejemplo).

Aunque no he podido hacerlo para este estudio, sería interesante realizar una tabla de asesinatos totales por porcentaje de población total desde el año 1976.

## →¿A QUÉ ETNIAS SE MATA? ¿CÓMO?

Sobre esas views realizaremos dos subconsultas para dar un perfil de a qué etnias se mata más y cómo.

```
SELECT VicRace, Weapon, COUNT(*) as total_asesinatos
  FROM `map_usa`.`shr76_21`
 WHERE State IN (
                 SELECT State
                   FROM low_10
                   )
 GROUP BY VicRace, Weapon
 ORDER BY COUNT(*) DESC
 LIMIT 10 ;
```

```
SELECT VicRace, Weapon, COUNT(*) as total_asesinatos
  FROM `map_usa`.`shr76_21`
 WHERE State IN (
                 SELECT State
                   FROM top_10
                   )
 GROUP BY VicRace, Weapon
 ORDER BY COUNT(*) DESC
 LIMIT 10 ;
```

| VicRace | Weapon | total_asesinatos |
|---|---|---|
| White | Handgun - pistol, revolver, etc | 2524 |
| White | Knife or cutting instrument | 1301 |
| White | Personal weapons, includes beating | 911 |
| White | Rifle | 692 |
| White | Other or type unknown | 664 |
| Black | Handgun - pistol, revolver, etc | 645 |
| White | Firearm, type not stated | 608 |
| White | Shotgun | 496 |
| White | Blunt object - hammer, club, etc | 484 |
| Black | Firearm, type not stated | 449 |

*Figura 3: Raza/arma low 10*

| VicRace | Weapon | total_asesinatos |
|---|---|---|
| Black | Handgun - pistol, revolver, etc | 131354 |
| White | Handgun - pistol, revolver, etc | 113407 |
| White | Knife or cutting instrument | 40233 |
| Black | Knife or cutting instrument | 31350 |
| Black | Firearm, type not stated | 25759 |
| White | Firearm, type not stated | 17966 |
| White | Personal weapons, includes beating | 17400 |
| White | Other or type unknown | 15510 |
| White | Blunt object - hammer, club, etc | 13410 |
| White | Shotgun | 12949 |

*Figura 4: Raza/arma top 10*

Aunque en los estados **más poblados** es más **equitativo**,
en los estados **menos poblados** los principales asesinados son **blancos**.

Esto tampoco debería sorprender, ya que la población de los estados menos poblados es principalmente blanca (link aquí).

Por tanto no se encuentra ninguna anomalía estadística sino que todo se corresponde a lo esperado.

## →GÉNERO

Sobre género, miramos quién es más asesinado. ¿Hombres o mujeres?

```
SELECT VicSex, COUNT(*) as total_victimas
  FROM `map_usa`.`shr76_21`
 WHERE VicSex != "Unknown"
 GROUP BY VicSex
 ORDER BY COUNT(*) DESC ;
```

| VicSex | total_victimas |
|---|---|
| Male | 658240 |
| Female | 189425 |

Vemos que la mayoría de víctimas (casi el triple) son **masculinas**, pero esto es algo inconcluso: la estadística dice que la mayoría de asesinatos de hombres se producen por otros hombres. ¿Qué ocurre si filtramos sólo para ver víctimas asesinadas por alguien de otro género?

```
SELECT VicSex, OffSex, COUNT(*) as total_victimas , ROUND(AVG(VicAge),1) as media_edad_victimas
  FROM `map_usa`.`shr76_21`
 WHERE VicSex != OffSex AND VicSex != "Unknown" AND OffSex != "Unknown"
 GROUP BY VicSex, OffSex
 ORDER BY COUNT(*) DESC ;
```

| VicSex | OffSex | total_victimas | media_edad_victimas |
|---|---|---|---|
| Female | Male | 131880 | 43,9 |
| Male | Female | 52150 | 46,1 |

↑

Para sorpresa de nadie, hay casi el triple de mujeres asesinadas por hombres que hombres asesinados por mujeres.

La media de edad (una variable que he incluido para dar algo más de análisis) de las primeras es, además, algo menor.

## →VARIACIÓN PORCENTUAL

Ahora entramos en un código algo más largo. Primero vamos a agrupar por año y asesinatos totales, luego vamos a crear una view y sobre esa view, usando la función LAG() vamos a crear una columna extra que nos indique la diferencia porcentual de año en año.

```
CREATE VIEW y_a AS
SELECT year_crime as Year, COUNT(*) as total
  FROM `map_usa`.`shr76_21`
 GROUP BY Year ;

 SELECT Year,  total,  ROUND(((total - LAG(total) OVER (ORDER BY Year)) / LAG(total) OVER (ORDER BY Year)) * 100, 2)
                        AS diferencia_porcentual
   FROM y_a
  ORDER BY Year;
```

Nos da un resultado como este:

| Year | total | diferencia_porcentual |
|---|---|---|
| 1976 | 17619 | NaN |
| 1977 | 18844 | 6,95 |
| 1978 | 19523 | 3,6 |
| 1979 | 21698 | 11,14 |
| 1980 | 23092 | 6,42 |
| 1981 | 21208 | -8,16 |
| 1982 | 20544 | -3,13 |
| 1983 | 19653 | -4,34 |
| 1984 | 18093 | -7,94 |
| 2019 | 16836 | -2,91 |
| 2020 | 21182 | 25,81 |

↑ Lo más llamativo es el aumento del **25%** de homicidios en 2020 con respecto a 2019, precisamente el año de la pandemia.
Pudo ser por la tensión política y las revueltas del país en el momento, pero también es probable es que ese año se hayan actualizado muchos casos anteriores o se hayan incluido en el sistema. También pudo haber más violencia doméstica por el confinamiento, o puede ser que se consideraran homicidios muertes derivadas por COVID.

## →PRESIDENTES

Para realizar la última consulta de esta parte, vamos a hacer una serie de cosas.

**Primero**, vamos a unir nuestros datos con una tabla que contiene nombres de presidentes, partido político, año de comienzo de mandato y año de fin. **Luego**, vamos a añadir una columna calculando el total de asesinatos por mandato de presidentes y, **finalmente**, dividiendo ese total por sus años de mandato, añadiremos una columna que indique los asesinatos medios por año de cada uno de los presidentes.

Añadiremos una cláusula **WHERE u.year_begin >=1976 AND u.year_end <=2021** para asegurarnos de que no pillamos a presidentes para los que solo contemos la mitad del mandato (ya que los datos del MAP empiezan en 1976 y acaban en 2021 en el archivo que tenemos).

```
SELECT u.pres_name, u.Party, u.year_begin, u.year_end, SUM(total) as total_asesinatos, ROUND(Sum(total)/(u.year_end - u.year_begin),0) as asesinatos_por_año
  FROM y_a y
  JOIN map_usa.us_presidents u
    ON y.Year BETWEEN u.year_begin AND u.year_end
 WHERE u.year_begin >=1976 AND u.year_end <=2021
 GROUP BY u.pres_name, u.Party, u.year_begin, u.year_end
 ORDER BY asesinatos_por_año DESC ;
```

| pres_name | Party | year_begin | year_end | total_asesinatos | asesinatos_por_año |
|---|---|---|---|---|---|
| George H.W. Bush | Republican | 1989 | 1993 | 111901 | 27975 |
| James (Jimmy) Carter | Democratic | 1977 | 1981 | 104365 | 26091 |
| Donald Trump | Republican | 2017 | 2021 | 95994 | 23999 |
| Ronald Reagan | Republican | 1981 | 1989 | 175439 | 21930 |
| William (Bill) Clinton | Democratic | 1993 | 2001 | 166849 | 20856 |
| George W. Bush | Republican | 2001 | 2009 | 148646 | 18581 |
| Barack Obama | Democratic | 2009 | 2017 | 144134 | 18017 |

Aunque sería fácil decir que Trump es el presidente con mayor media de asesinatos anuales desde 1993, hemos visto antes que en 2020 hay una anomalía en la cantidad de asesinatos, no deberíamos saltar a conclusiones precipitadas sin estudiarlo en profundidad.

# PARTE 2. EMPAREJANDO ASESINOS EN SERIE.

Siendo el schema **map_usa** y las tablas **shr76_21** (los datos de MAP), **us_presidents** (la tabla con los presidentes) y **sk_dt** (asesinos en serie).

## →¿QUÉ VAMOS A HACER?

Lo primero es tratar de unir nuestras dos tablas de forma que los años de actividad, estado en los que actuaban, modus operandi, etnia favorita, género de la víctima y edad de la víctima de nuestra tabla de asesinos seriales coincida con esas características en nuestra tabla de homicidios para casos que **no hayan sido resueltos**.

```
CREATE VIEW df_final AS
SELECT *
  FROM `map_usa`.`shr76_21` s
  JOIN  map_usa.sk_dt sk
    ON s.State = sk.state1 OR s.State = sk.state2
 WHERE s.year_crime BETWEEN sk.start_crimes AND sk.end_crimes
   AND (s.Weapon = sk.weapon1 OR s.Weapon = sk.weapon2 OR s.Weapon = sk.weapon3)
   AND s.VicRace = sk.victimrace
   AND s.VicSex = sk.victimsex
   AND s.VicAge BETWEEN sk.victimagemin AND sk.victimagemax
   AND s.Solved = "No";
```

```
SELECT COUNT(*) AS total_matches
  FROM df_final ;
```

| total_matches |
| --- |
| 4007 |

← ¡Hay 4007 coincidencias!

Vamos a guardar esta **view** como **df_final.**

## →¿MÁS ASESINATOS = MÁS ASESINOS EN SERIE?

Vamos ahora a comparar la tabla de estados con más asesinatos que ya teníamos con una tabla en la que consultaremos cuales son los estados con más asesinos en serie. Vamos a guardarla como **top_states**.

```
SELECT State, COUNT(DISTINCT full_name_sk) as total_asesinos
  FROM df_final
 GROUP BY State
 ORDER BY total_asesinos DESC
 LIMIT 5 ;
```

| State | total_asesinos |
|---|---|
| California | 8 |
| Illinois | 7 |
| New York | 5 |
| Texas | 4 |
| Missouri | 3 |

← La única sorpresa es Missouri, 4 asesinos emparejados sin encontrarse en el top 10 de estados con más asesinatos.

## →A LA CAZA DEL ASESINO

Ahora vamos a realizar un par de consultas: **primero**, vamos a usar nuestra lista **top_states** de los 5 estados con más asesinos en serie coincidentes en casos sin resolver. **Después** vamos a listar los 10 asesinos en serie con menos coincidencias como **lowest_sk_count.**

**Finalmente**, vamos a unir ambas listas para encontrar un asesino con pocos casos coincidentes en estado que no tenga demasiados asesinos emparejados, para conseguir no buscar en estados en los que pudiera haber muchos o buscar asesinos que, por su método o perfil de víctima, tengan muchas coincidencias.

```
CREATE VIEW lowest_sk_count AS
SELECT full_name_sk, COUNT(*) as total_coincidencias
  FROM df_final
 GROUP BY full_name_sk
 ORDER BY total_coincidencias ASC
 LIMIT 10 ;
```

*Asesinos con menos coincidencias*

```
SELECT full_name_sk, COUNT(*) AS total_matches
  FROM df_final
 WHERE State NOT IN (
                     SELECT State
                       FROM top_states
                     )
   AND full_name_sk IN (
                     SELECT full_name_sk
                       FROM lowest_sk_count
                     )
 GROUP BY full_name_sk
 ORDER BY total_matches ASC ;
```

*Asesinos con menos coincidencias en estados con menos asesinos*

| full_name_sk | total_matches |
|---|---|
| Connecticut River Valley Killer | 1 |
| Anthony Sowell | 1 |
| Frankford Slasher | 2 |
| Edwin Kaprat | 2 |
| Robert Hansen | 6 |
| William Sapp (serial killer) | 8 |

De los resultados podemos obviar al **Frankford Slasher** y al **Conneticut River Valley Killer**, ya que son asesinos no identificados y probablemente estén saltando los mismos casos que se le atribuyen y aún no están resueltos.

Sobre el resto, podemos delimitar su zona de acción y ver si encajan las fechas y métodos, cosa que haremos justo ahora.

## →CERRANDO LA SELECCIÓN

Sobre todo ahora tenemos que extraer el **condado** en que actuaba cada uno de los asesinos y cruzarlo con la columna **CNTYFIPS** del dataset de asesinatos en USA, que contiene el condado en el que se ha producido el incidente.

Los asesinos que han resultado, su método, radio de acción y tipo de víctima son los siguientes:

**-Anthony Sowell:** Actúa en Cleveland, Ohio, Cuyahoga County (código CNTYFIPS "**Cuyahoga, OH**") .

**Víctimas:** Mujeres negras muy obesas o muy delgadas, de entre 25 y 53 años. **Modus operandi:** Estrangulación

**-Edwin Kaprat:** Actúa en Hernando County (código CNTYFIPS "**Hernando, FL**") .

**Víctimas:** Mujeres de avanzada edad , de más de 70 años. **Modus operandi:** Paliza, y posterior incendio de la casa

**-Robert Hansen:** Actúa en Anchorage, Alaska (código CNTYFIPS "**Anchorage, AK**") .

**Víctimas:** Mujeres blancas , de entre 17 y 41 años. **Modus operandi:** Disparaba en el bosque con un fusil a modo de cacería.

**-William Sapp:** Actúa en Springfield, Ohio, Clark County (código CNTYFIPS "**Clark, OH**").

**Víctimas:** Mujeres blancas , de entre 11 y 58 años. **Modus operandi:** Paliza con objeto contundente.

Ahora podemos cercar por condados, y comprobar las circunstancias de los casos resultantes

Nuestra **consulta final** será :

```
SELECT full_name_sk, CNTYFIPS, State, year_crime, Month , VicAge, VicSex, VicRace, Weapon, Circumstance, Subcircum, Solved, Situation
  FROM df_final
 WHERE full_name_sk IN ("Anthony Sowell" , "Robert Hansen", "Edwin Kaprat" , "william Sapp (serial killer)")
   AND CNTYFIPS IN ("Hernando, FL", "Anchorage, AK", "Clark, OH", "Cuyahoga, OH") ;
```

Y el **resultado**:

| full_name_sk | CNTYFIPS | State | year_crime | Month | VicAge | VicSex | VicRace | Weapon | Circumstance | Subcircum | Solv | Situation |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Robert Hansen | Anchorage, AK | Alaska | 1979 | February | 26 | Female | White | Handgun - pistol, revolver, etc | Robbery | | No | Single victim/unknown offender(s) |
| William Sapp (serial killer) | Clark, OH | Ohio | 1992 | August | 12 | Female | White | Blunt object - hammer, club, etc | Other | | No | Multiple victims/unknown offender(s) |
| Anthony Sowell | Cuyahoga, OH | Ohio | 2007 | November | 45 | Female | Black | Blunt object - hammer, club, etc | Circumstances undetermined | | No | Single victim/unknown offender(s) |

-Podemos descartar la autoría del **primero:** Robert hansen no usaba una pistola sino un rifle en la mayoría de ocasiones y, sobre todo, su motivo no era el robo.

-Aunque el segundo es sospechoso, el hecho de que en la columna **Situation** diga **Multiple victims** , que la fecha sea **Agosto de 1992** y la edad **12 años** hace sospechar que se trata de Phree Marrow, asesinada por Sapp junto a Martha Leach, de 11 años, en agosto de 1992. Probablemente el que en la categoría **Solved** diga **No** sea una errata policial.

-El tercer caso **sí puede ser interesante**: Sowell no tiene registro de haber matado a ninguna mujer de 45 años en noviembre de 2007, y sus asesinatos comienzan en mayo de ese mismo año. Se encuentra en el rango de edad adecuado, con el modus operandi adecuado.

Podríamos investigar más a fondo ese caso de trabajar junto a la policía.

Quiero concluir recordando que este trabajo solo ha sido estimado para comprobar las posibilidades de cruzar ambas bases de datos y elaborar sobre los resultados.  Ninguna de las conclusiones a las que haya llegado pueden considerarse como 100% fiables y con más tiempo podría conseguirse un resultado mucho más fino, pero ha sido divertido como toma de contacto.

Un saludo,

Alonso.