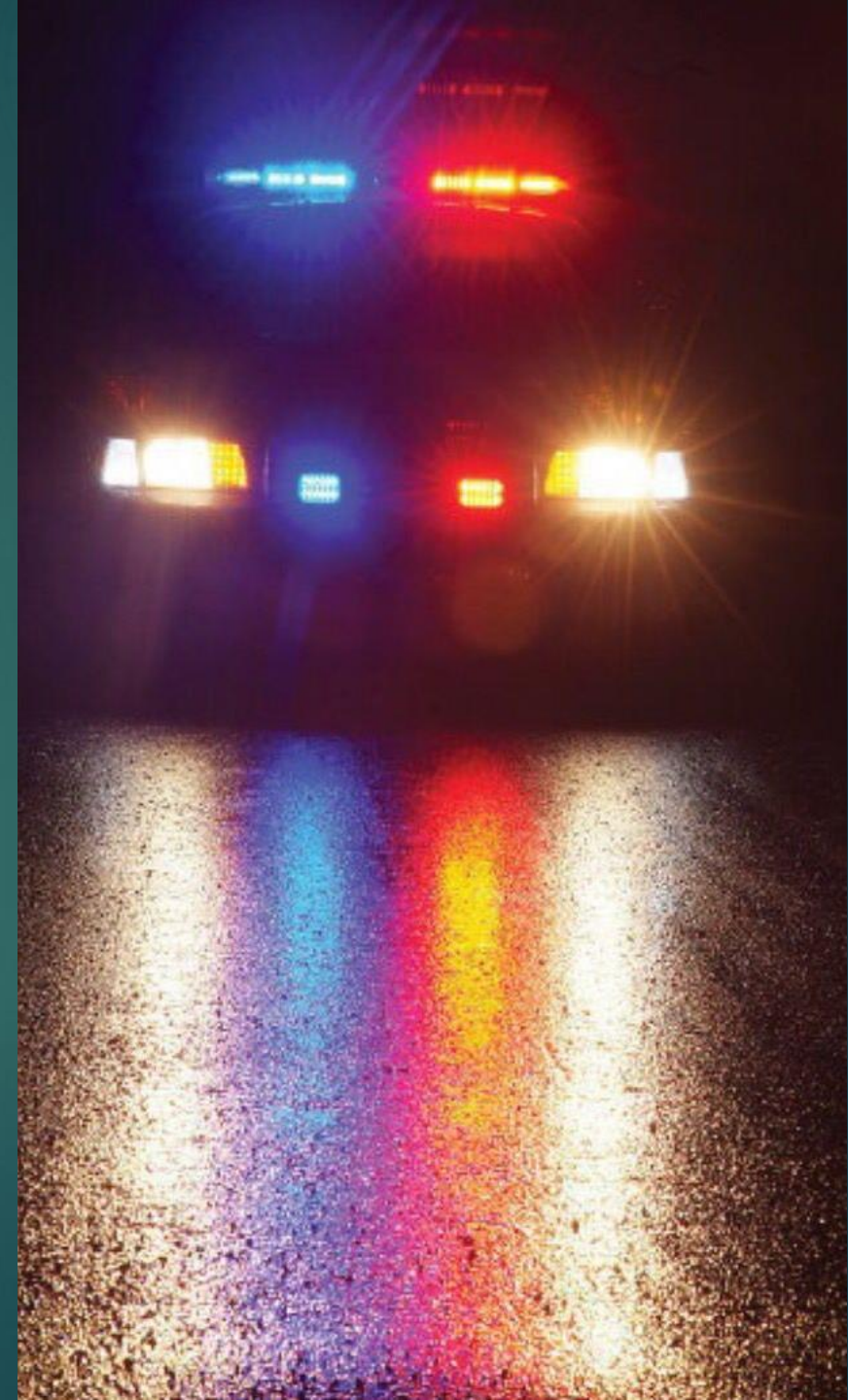


Proyecto 2: Análisis de las fatalidades tras intercambios violentos con la policía de EE.UU

INTRO CIENCIA DE DATOS

JUAN LORTHIOIS



Información del proyecto y de la base de datos

- ▶ Durante los años 2015 a 2017, el diario norte americano “The Washington Post” mantuvo un abase de datos en la que se recopiló información básica sobre cada fatalidad en la población civil que siguió un intercambio violento con la policía.
- ▶ Además se combinó con información socioeconómico y demográfica de la ciudad y el estado en el que ocurrió la fatalidad.
- ▶ Nota: el objetivo es hacer observaciones directamente sobre este dataset, basándose en los hechos presentados por el dataset.

Objetivos:

- ▶ 1. Obtener más información sobre las circunstancias en las que ocurren estas fatalidades.
- ▶ 2. Determinar si existen patrones particulares dentro de los datos estudiados.
- ▶ 3. Presentar la información mediante gráficas en la medida de lo posible.
- ▶ 4. Obtener “insights” sobre diversos aspectos sociales de la población norte americana.
- ▶ 5. Mantener la mayor neutralidad posible en la presentación de los resultados.

Supuestos:

- ▶ 1. La información reportada por la policía no fue alterada.
- ▶ 2. La información respectoa si la persona estaba atacando o huyendo no fue alterada.

Presentación de los datasets y de sus variables:

| | Geographic Area | | City | Median Income |
|---|-----------------|--|-----------------|---------------|
| 0 | AL | | Abanda CDP | 11207 |
| 1 | AL | | Abbeville city | 25615 |
| 2 | AL | | Adamsville city | 42575 |
| 3 | AL | | Addison town | 37083 |
| 4 | AL | | Akron town | 21667 |

| | Geographic Area | | City | poverty_rate |
|---|-----------------|--|-----------------|--------------|
| 0 | AL | | Abanda CDP | 78.8 |
| 1 | AL | | Abbeville city | 29.1 |
| 2 | AL | | Adamsville city | 25.5 |
| 3 | AL | | Addison town | 30.7 |
| 4 | AL | | Akron town | 42 |

| | Geographic Area | | City | percent_completed_hs |
|---|-----------------|--|-----------------|----------------------|
| 0 | AL | | Abanda CDP | 21.2 |
| 1 | AL | | Abbeville city | 69.1 |
| 2 | AL | | Adamsville city | 78.9 |
| 3 | AL | | Addison town | 81.4 |
| 4 | AL | | Akron town | 68.6 |

Presentación de los datasets y de sus variables:

| | Geographic area | | City | share_white | share_black | share_native_american | share_asian | share_hispanic |
|---|-----------------|-----------------|------|-------------|-------------|-----------------------|-------------|----------------|
| 0 | AL | Abanda CDP | | 67.2 | 30.2 | 0 | 0 | 1.6 |
| 1 | AL | Abbeville city | | 54.4 | 41.4 | 0.1 | 1 | 3.1 |
| 2 | AL | Adamsville city | | 52.3 | 44.9 | 0.5 | 0.3 | 2.3 |
| 3 | AL | Addison town | | 99.1 | 0.1 | 0 | 0.1 | 0.4 |
| 4 | AL | Akron town | | 13.2 | 86.5 | 0 | 0 | 0.3 |

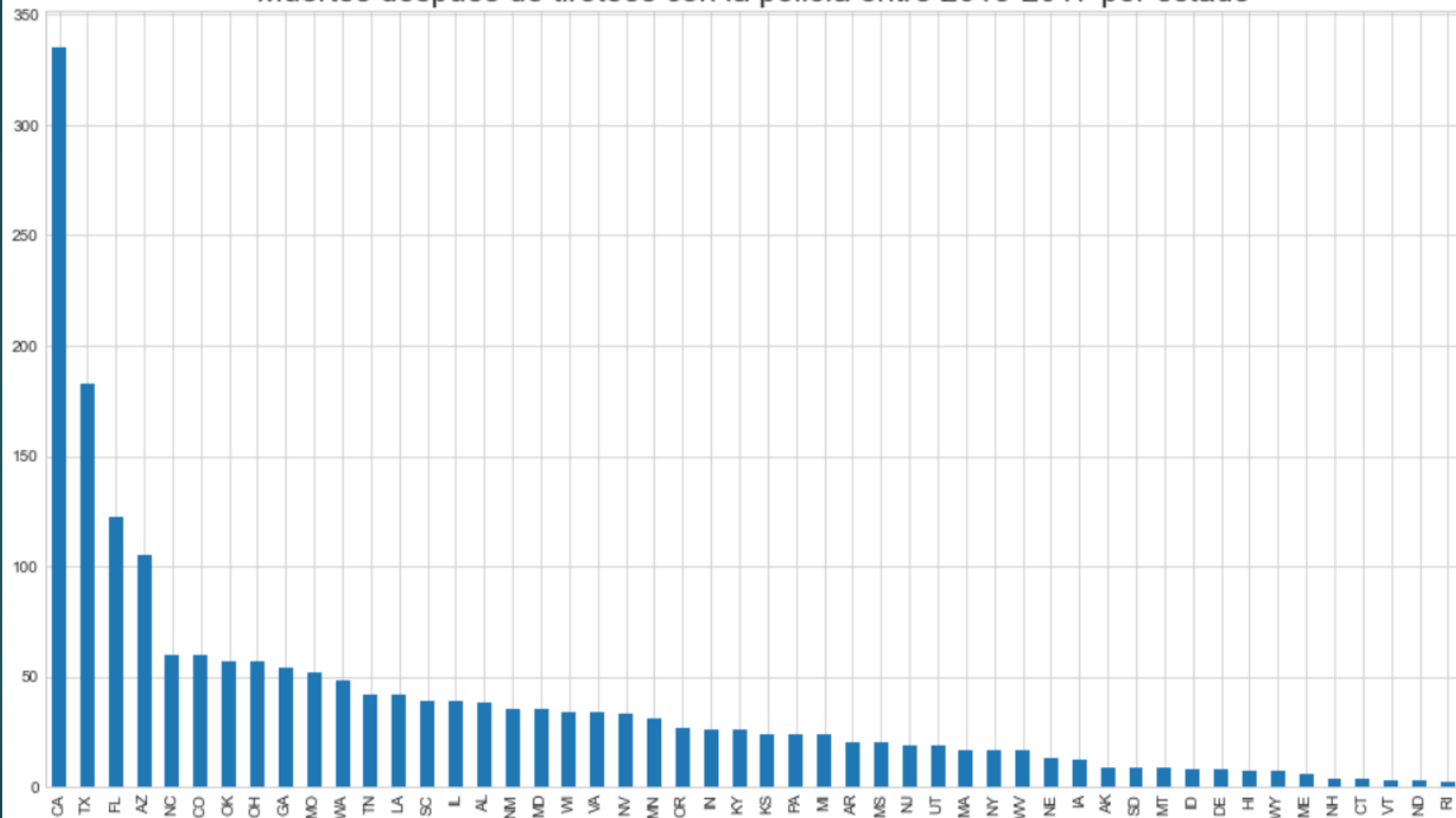
| | id | name | date | manner_of_death | armed | age | gender | race | city | state | signs_of_mental_illness | threat_level | flee | body_camera |
|---|----|--------------------|----------|------------------|------------|------|--------|------|---------------|-------|-------------------------|--------------|-------------|-------------|
| 0 | 3 | Tim Elliot | 02/01/15 | shot | gun | 53.0 | M | A | Shelton | WA | True | attack | Not fleeing | False |
| 1 | 4 | Lewis Lee Lembke | 02/01/15 | shot | gun | 47.0 | M | W | Aloha | OR | False | attack | Not fleeing | False |
| 2 | 5 | John Paul Quintero | 03/01/15 | shot and Tasered | unarmed | 23.0 | M | H | Wichita | KS | False | other | Not fleeing | False |
| 3 | 8 | Matthew Hoffman | 04/01/15 | shot | toy weapon | 32.0 | M | W | San Francisco | CA | True | attack | Not fleeing | False |
| 4 | 9 | Michael Rodriguez | 04/01/15 | shot | nail gun | 39.0 | M | H | Evans | CO | False | attack | Not fleeing | False |

Preguntas iniciales:

- ▶ ¿Cuál es el estado donde más fallecidos a manos de la polici se reportan?
- ▶ ¿Hay algún patrón inmediatamente aparente a partir del EDA?
- ▶ Dado que se cuenta con información socio económica para cada estado, es posible establecer una relación entre esto y los estados donde más se reportaron fallecidos?
- ▶ Para responder a las preguntas anteriores se juntaron los 5 datasets para tener una perspectiva global, por estado, asignando a cada reporte, las variables socioeconómicas correspondientes al lugar donde falleció la persona.

EDA

Muertes despues de tiroteos con la policia entre 2015-2017 por estado

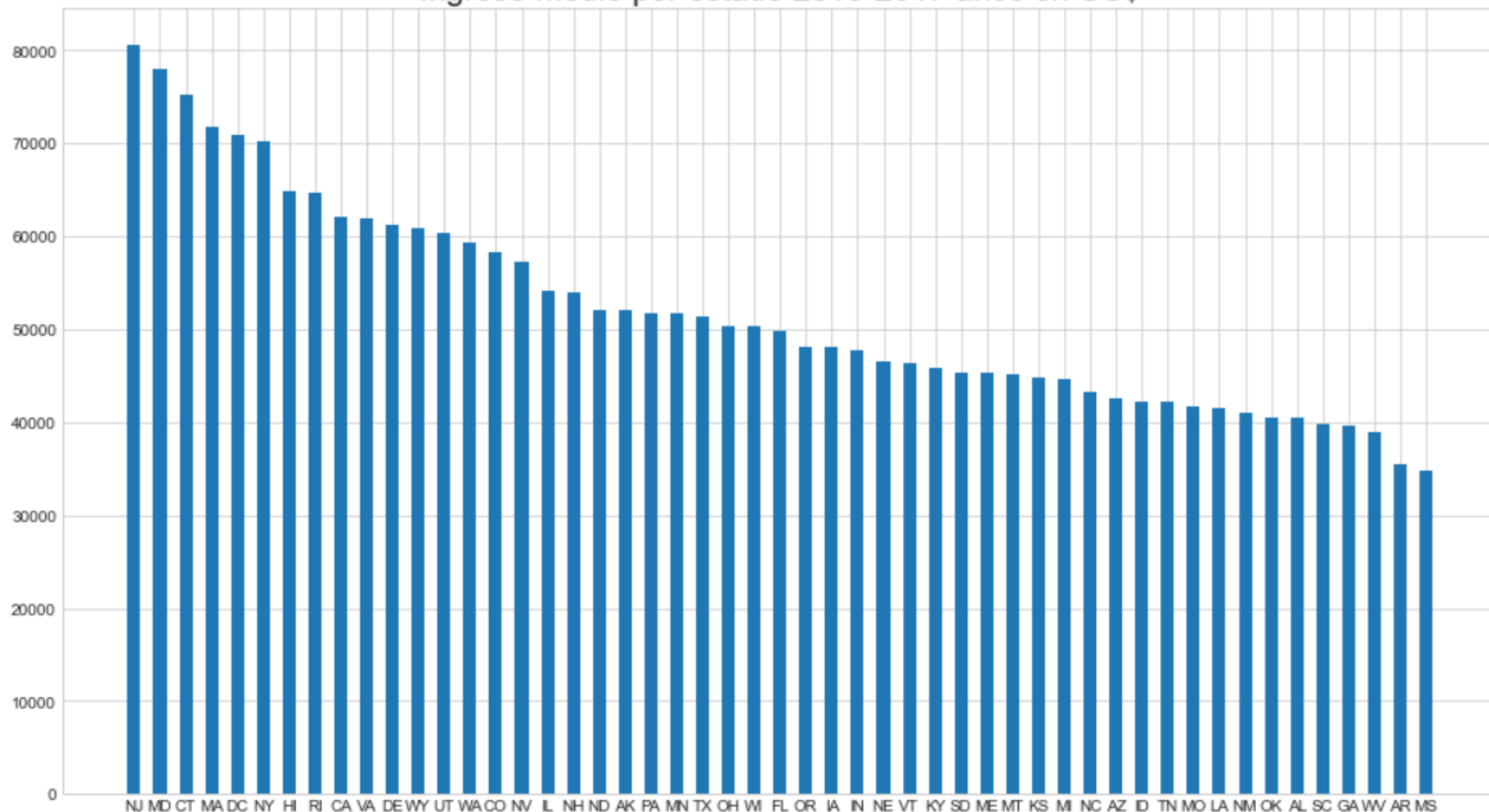


EDA

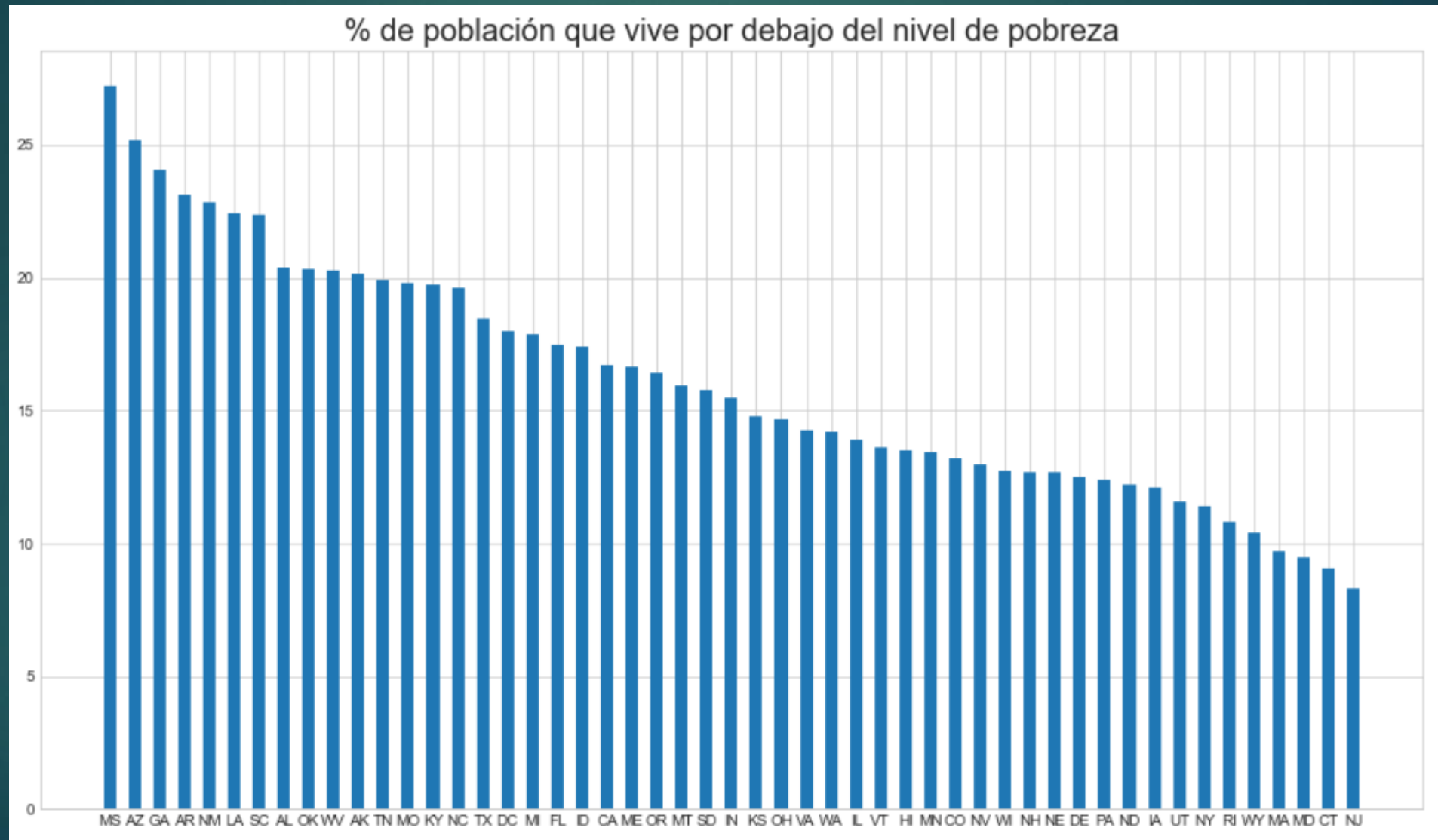
- ▶ De la gráfica anterior parece haber 3 niveles de violencia:
 - ▶ 1. Estados que reportan más de 100 fatalidades
 - ▶ 2. Los que reportan de 50 a 100
 - ▶ 3. Los que reportan menos de 50.
-
- ▶ ¿Qué tienen en común estos estados?
 - ▶ Continuamos con el EDA

EDA

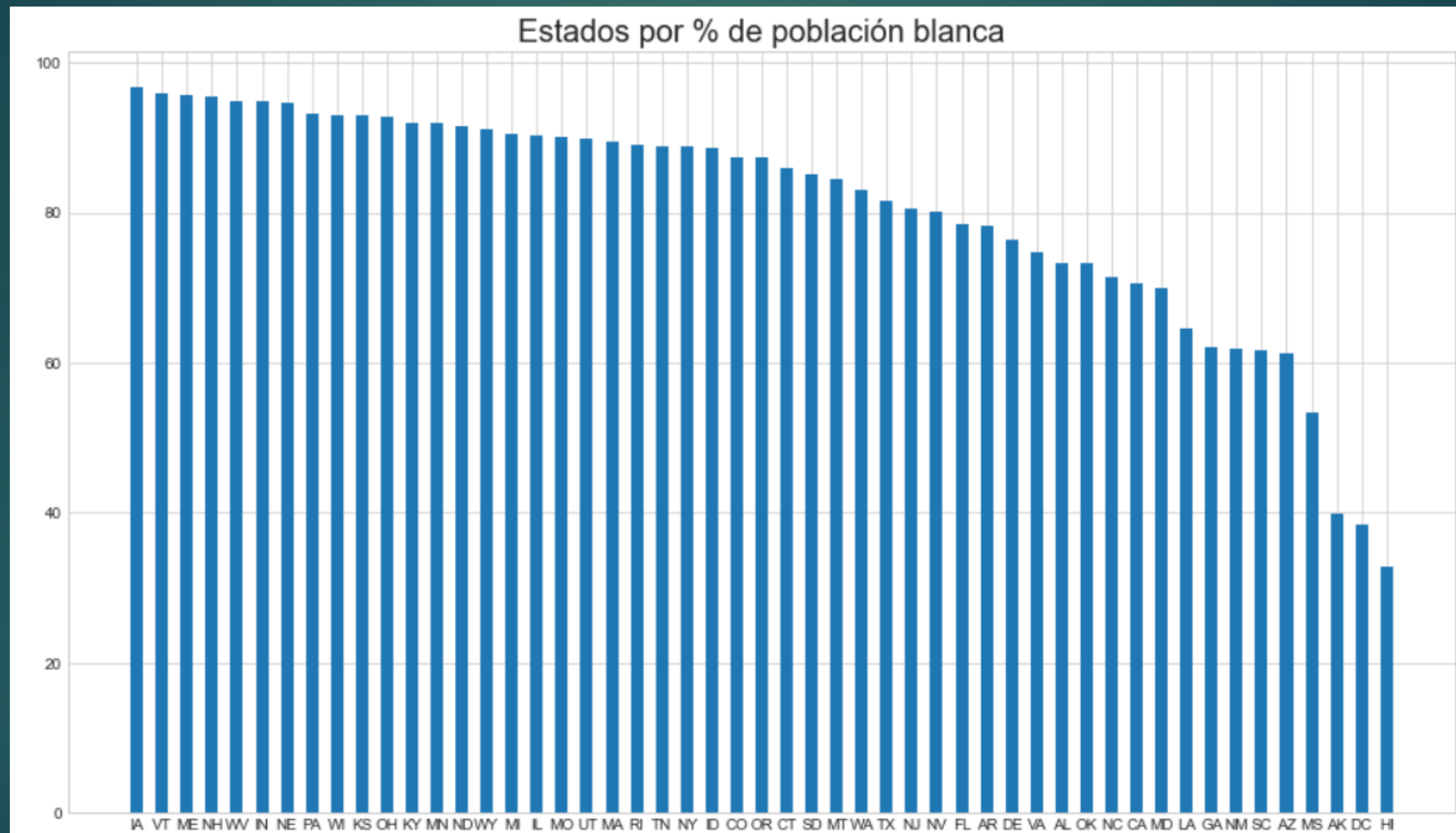
Ingreso medio por estado 2015-2017 años en US\$



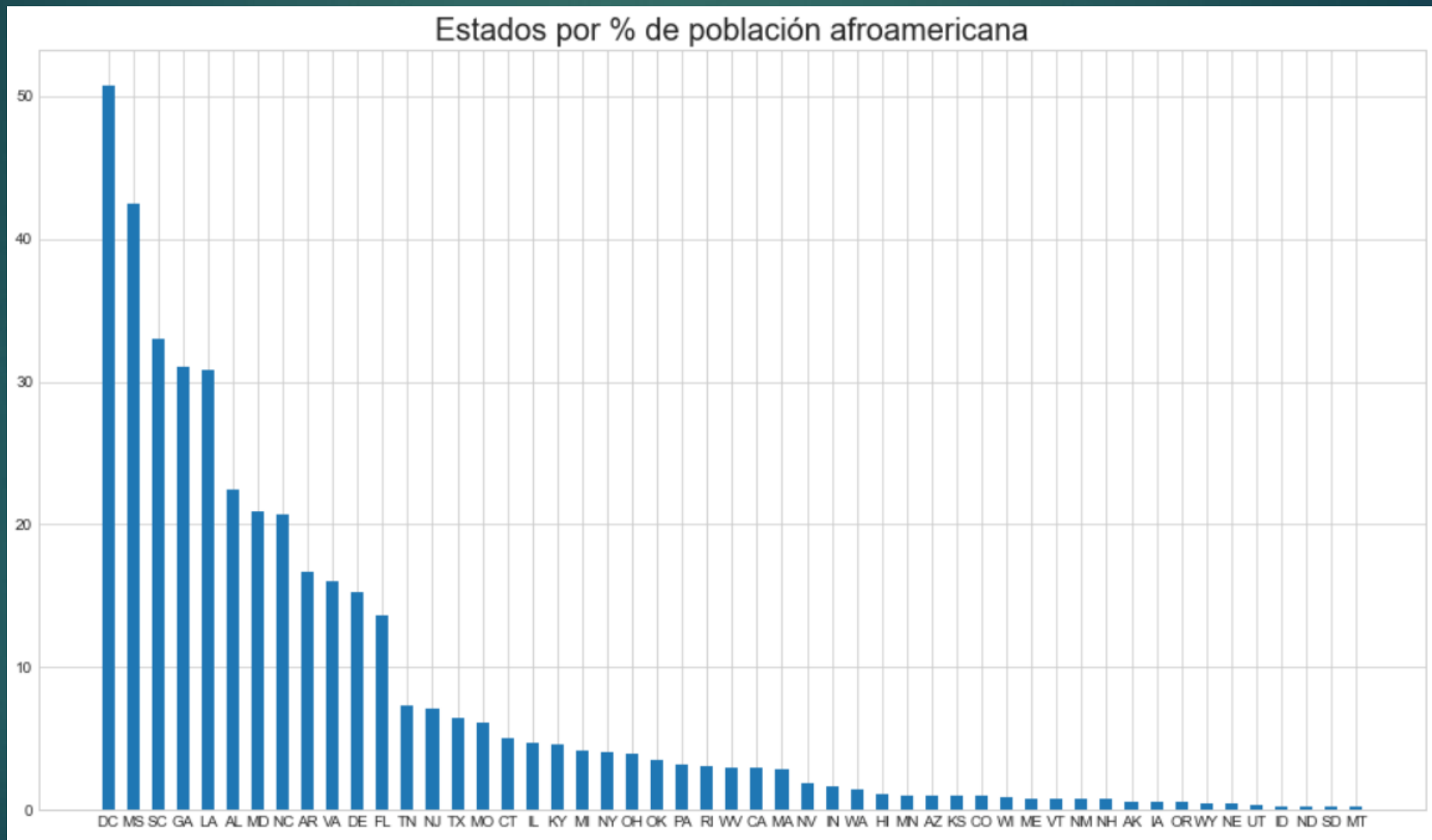
EDA



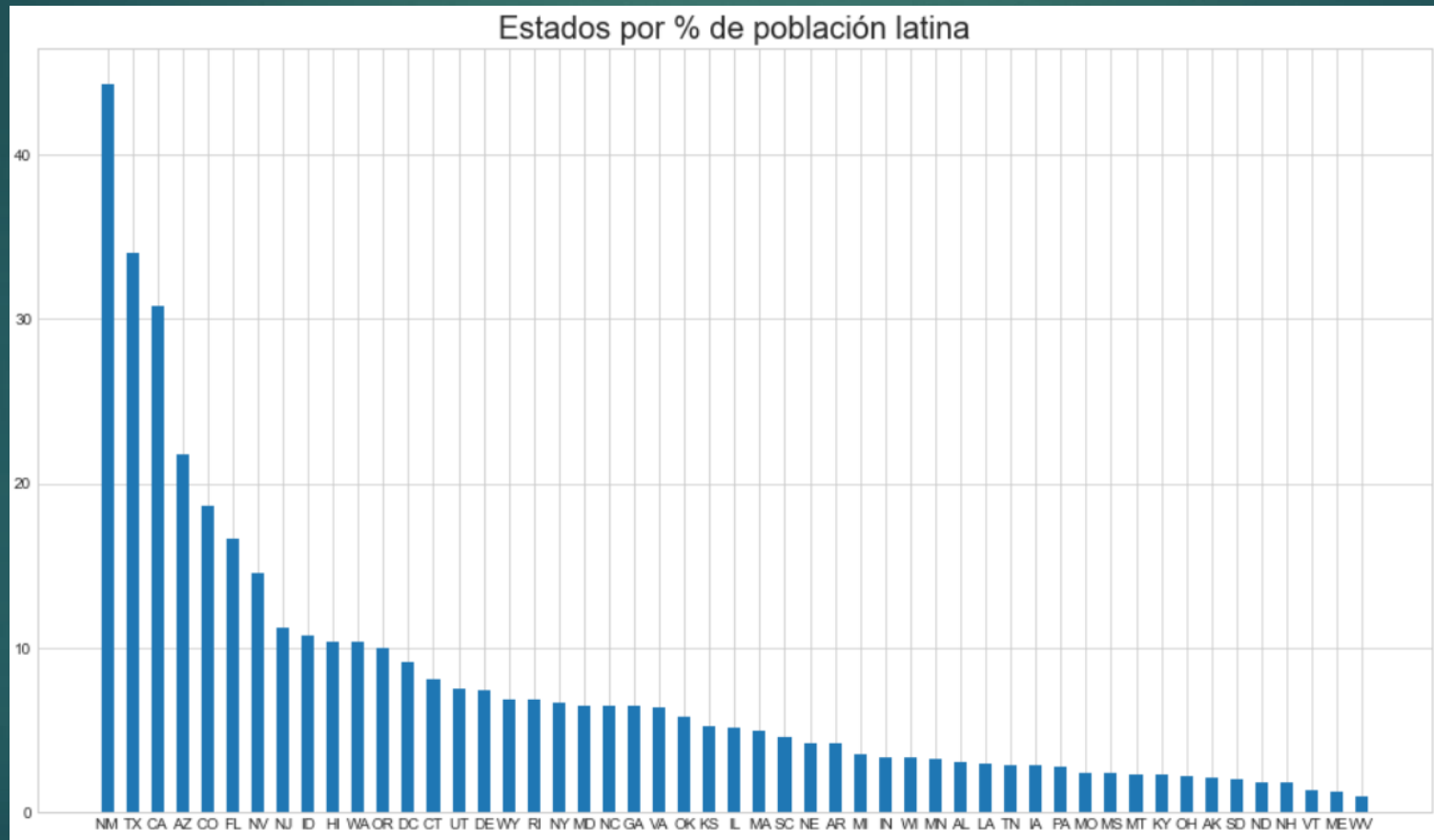
EDA



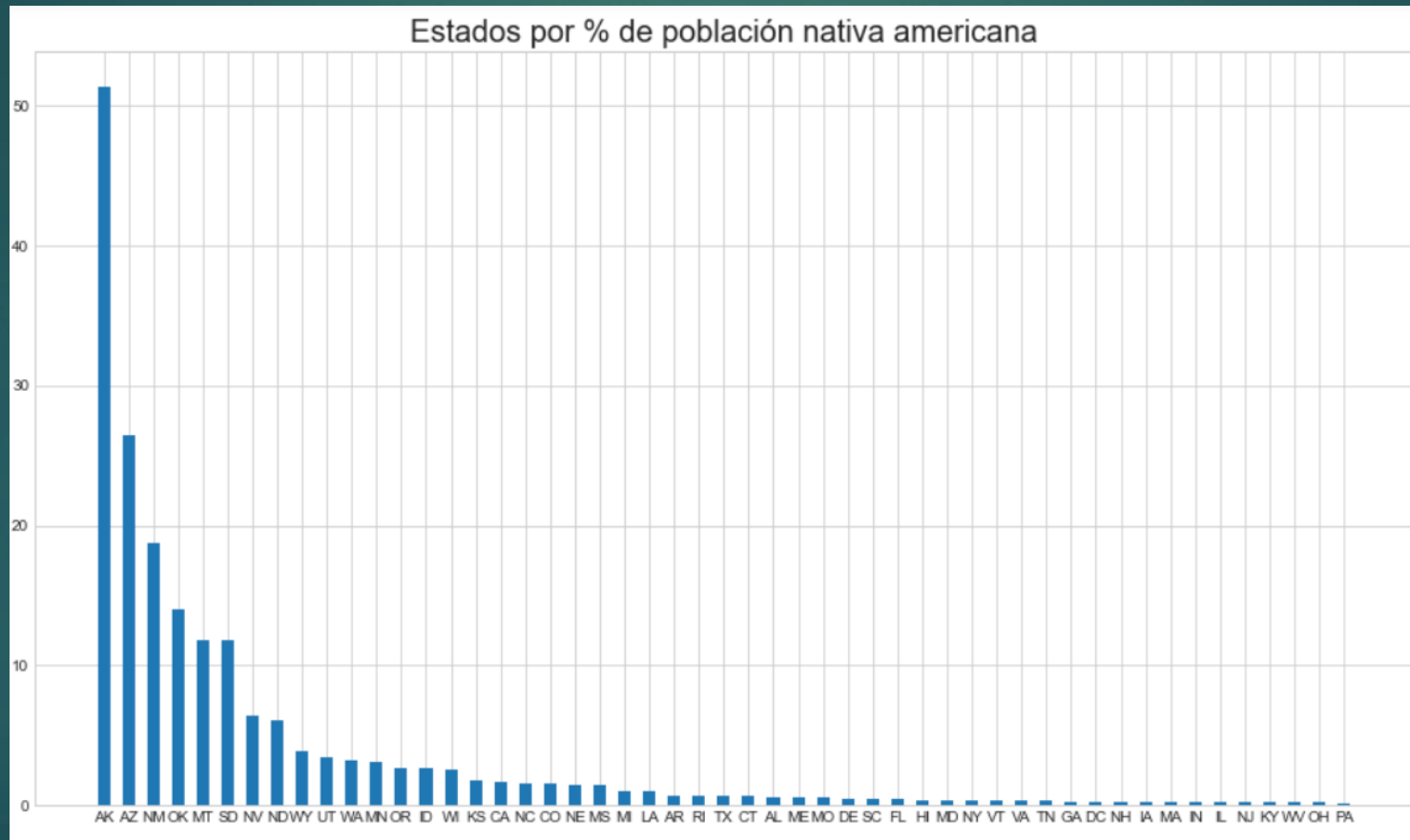
EDA



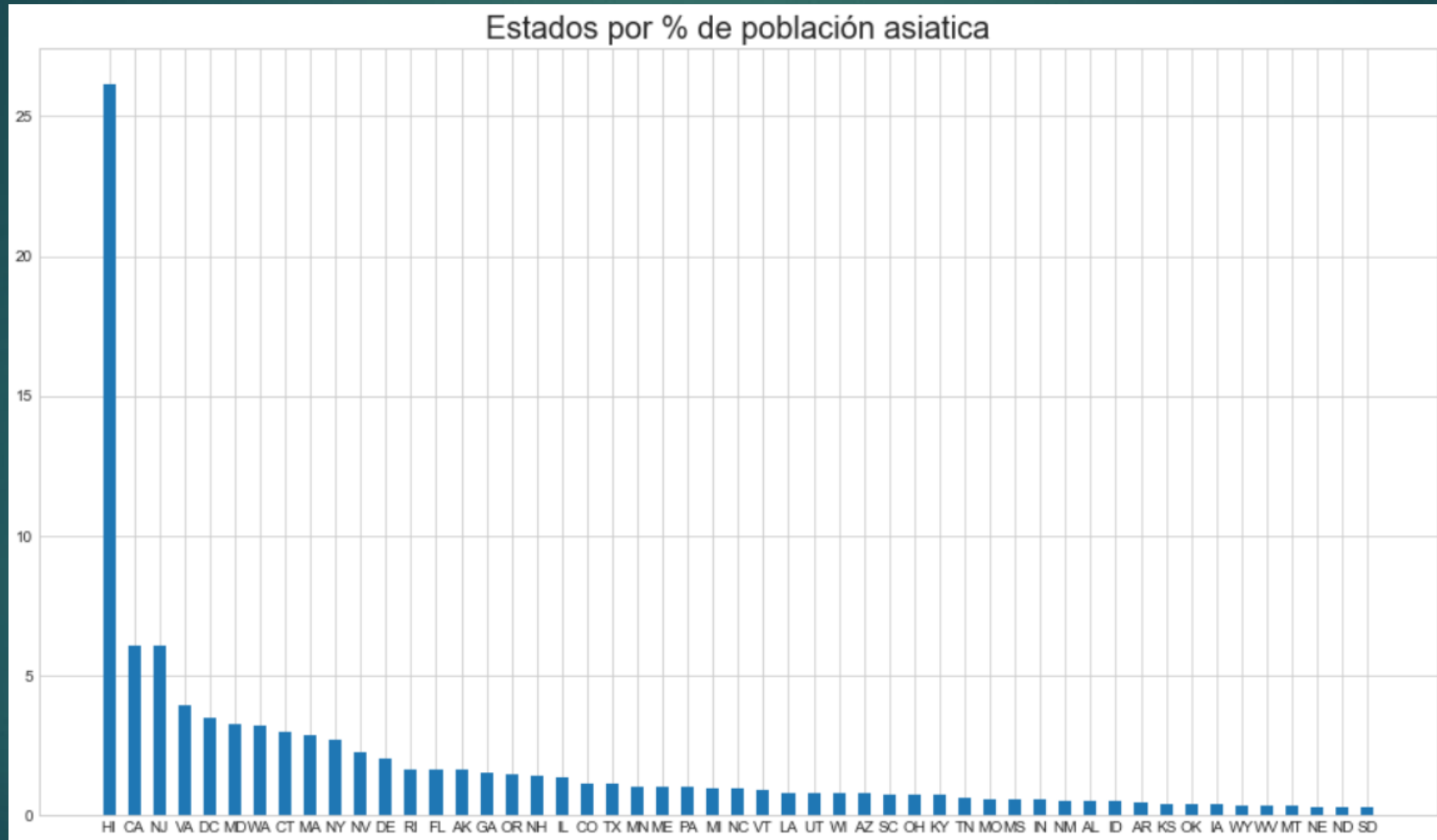
EDA



EDA



EDA



EDA: detalle de los 4 con mas fatalidades

```
Median Income      62165.578475
poverty_rate       16.725708
percent_completed_hs 81.802906
share_white        70.602235
share_black        2.919374
share_native_american 1.668480
share_asian        6.069449
share_hispanic     30.826975
Name: CA, dtype: float64
Median Income      51311.108214
poverty_rate       18.429036
percent_completed_hs 78.221159
share_white        81.597786
share_black        6.414388
share_native_american 0.676042
share_asian        1.126497
share_hispanic     33.971680
Name: TX, dtype: float64
```

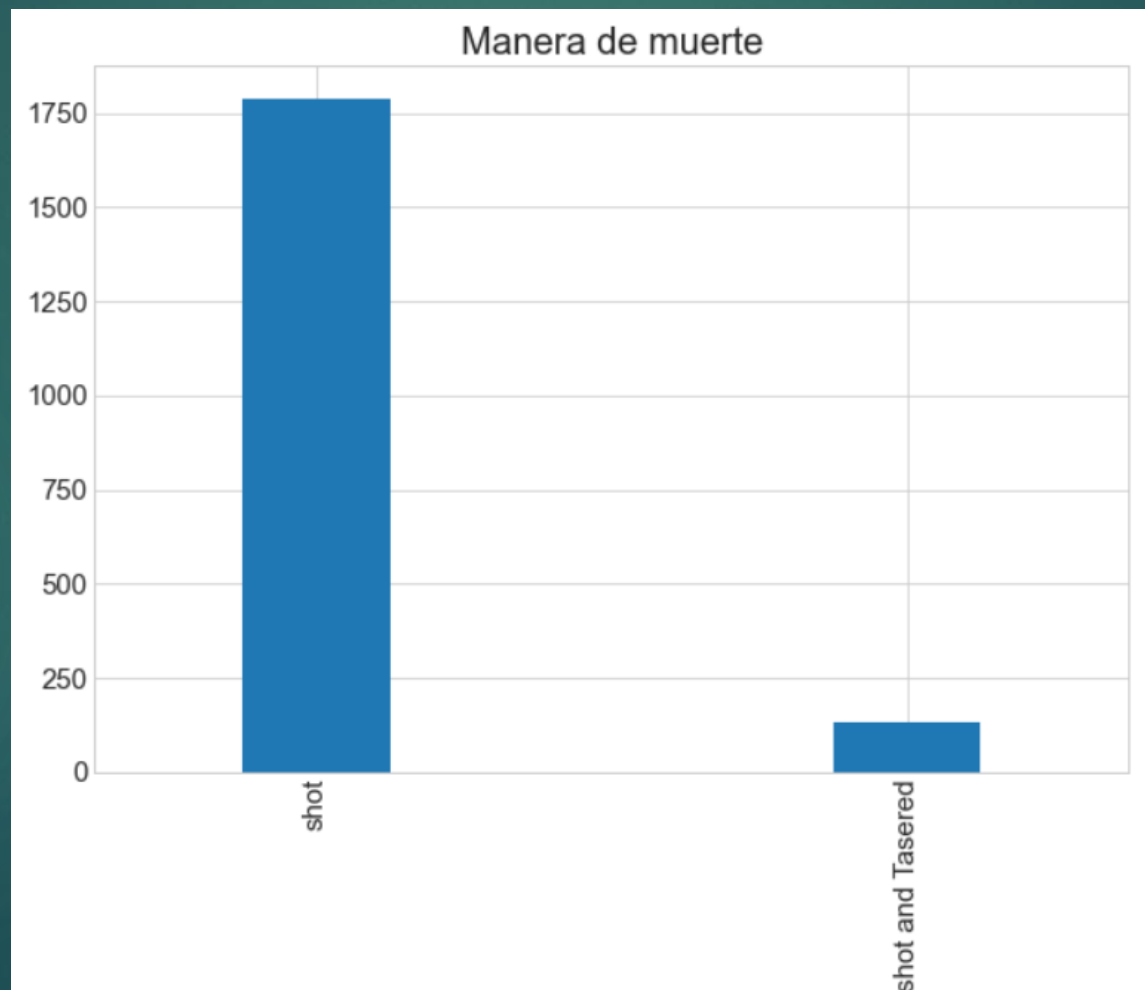
```
Median Income      49805.165533
poverty_rate       17.439480
percent_completed_hs 85.748869
share_white        78.487557
share_black        13.584502
share_native_american 0.450905
share_asian        1.652941
share_hispanic     16.596154
Name: FL, dtype: float64
Median Income      42512.921833
poverty_rate       25.143666
percent_completed_hs 80.553369
share_white        61.282480
share_black        1.011051
share_native_american 26.422102
share_asian        0.780863
share_hispanic     21.740162
Name: AZ, dtype: float64
```

EDA: detalle de los 4 con menos fatalidades

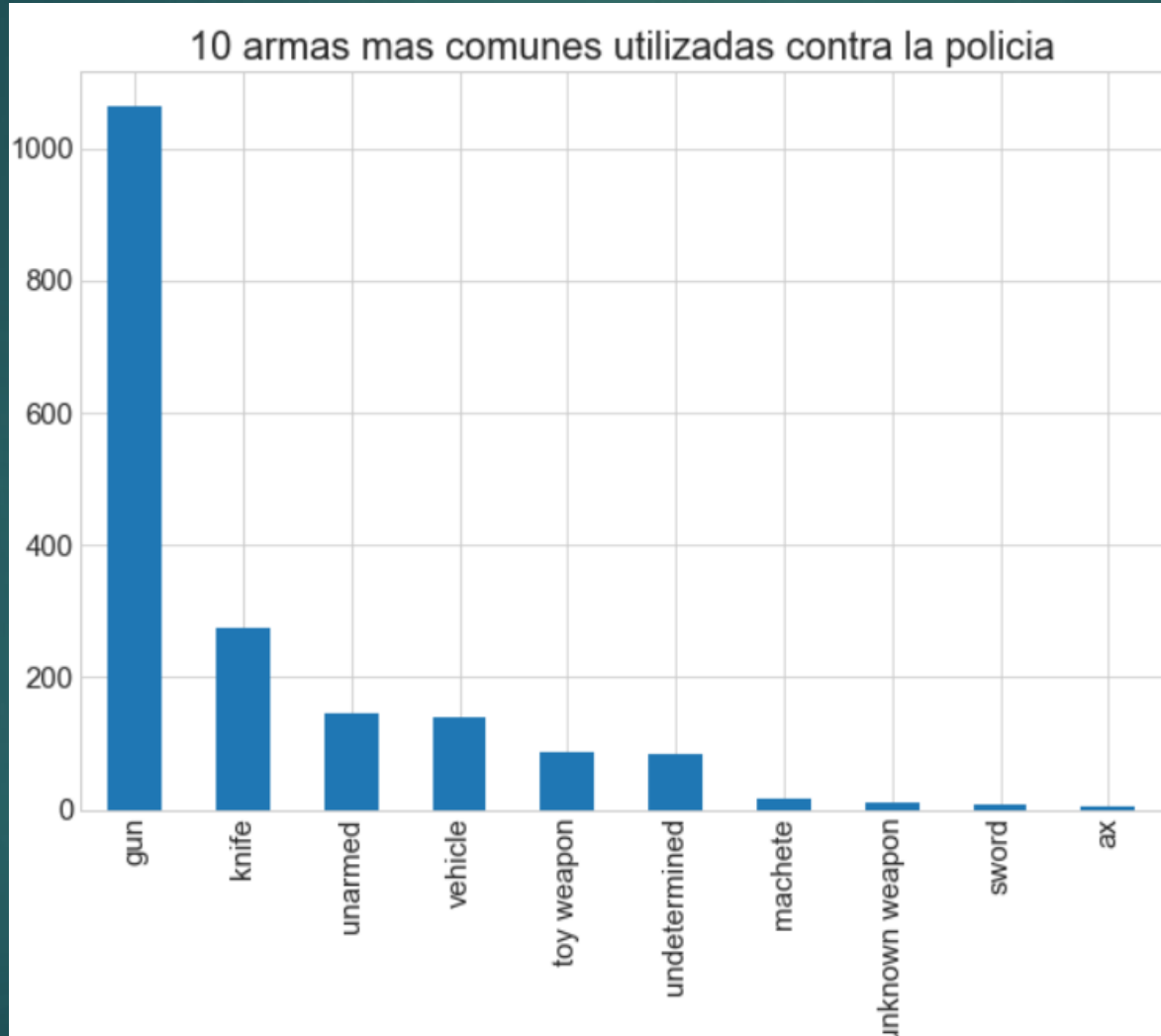
```
Median Income      53916.288889
poverty_rate       12.700000
percent_completed_hs  91.211111
share_white        95.531111
share_black        0.757778
share_native_american  0.280000
share_asian        1.415556
share_hispanic     1.806667
Name: NH, dtype: float64
Median Income      75200.321429
poverty_rate       9.055714
percent_completed_hs  91.520714
share_white        85.986429
share_black        5.040714
share_native_american  0.670000
share_asian        2.992143
share_hispanic     8.067857
Name: CT, dtype: float64
```

```
Median Income      52141.705882
poverty_rate       12.216176
percent_completed_hs  88.469706
share_white        91.496176
share_black        0.290294
share_native_american  6.110294
share_asian        0.312059
share_hispanic     1.828529
Name: ND, dtype: float64
Median Income      46373.982143
poverty_rate       13.629464
percent_completed_hs  90.112500
share_white        95.930357
share_black        0.785714
share_native_american  0.351786
share_asian        0.926786
share_hispanic     1.389286
Name: VT, dtype: float64
Median Income      64740.202020
```

EDA: Datset que detalla el enfrentamiento

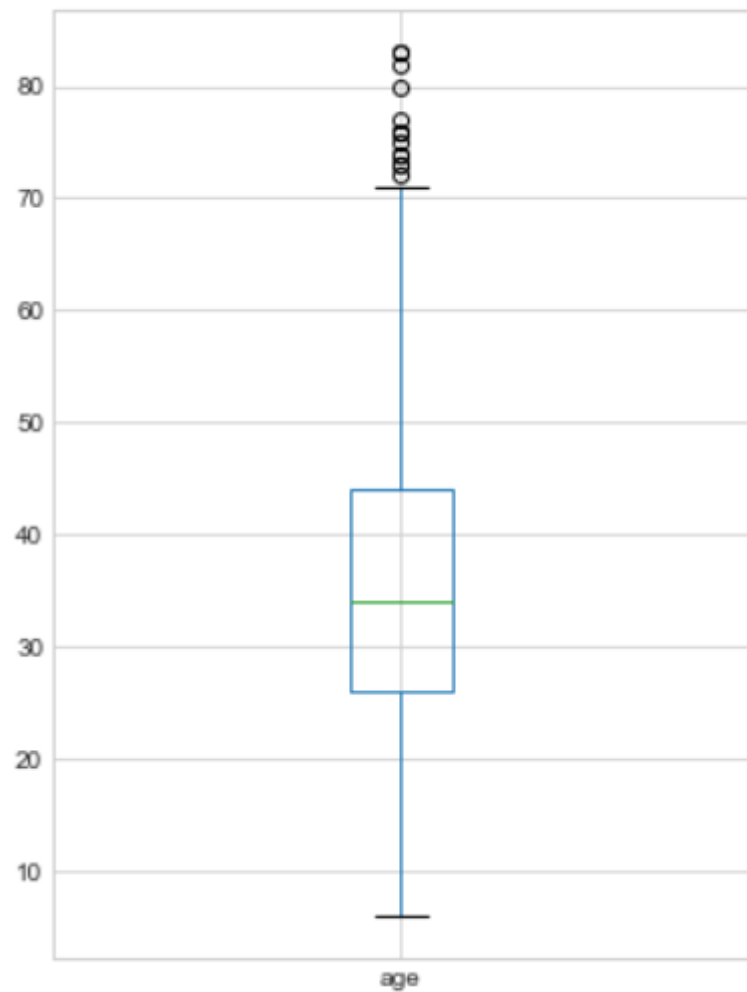


EDA

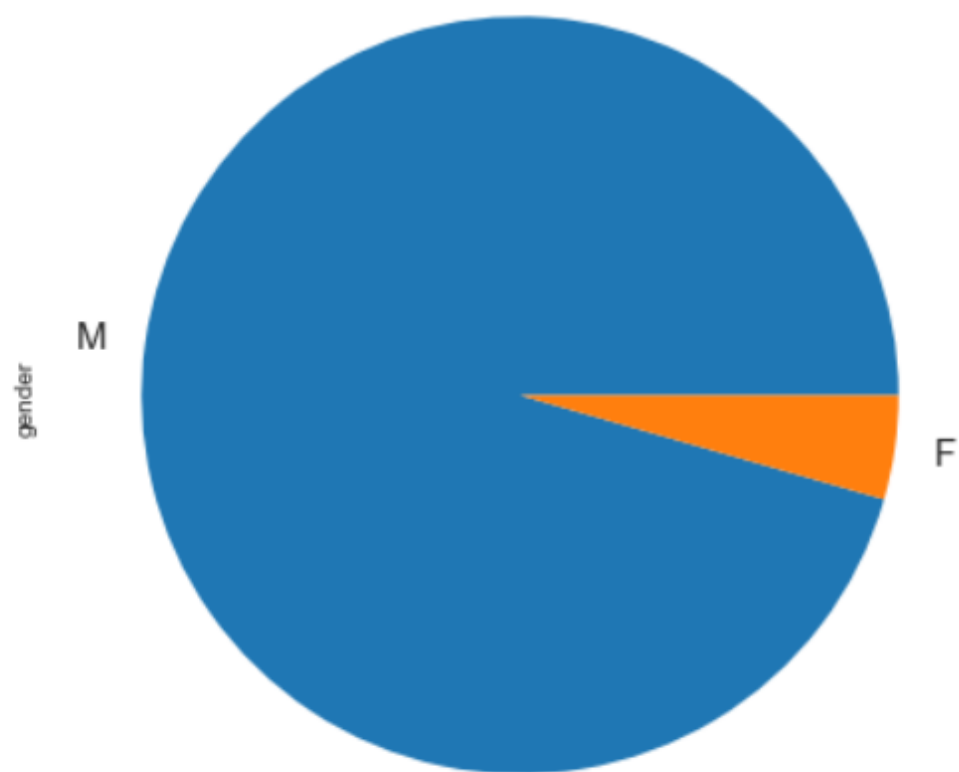


EDA

Distribucion edades fallecidos

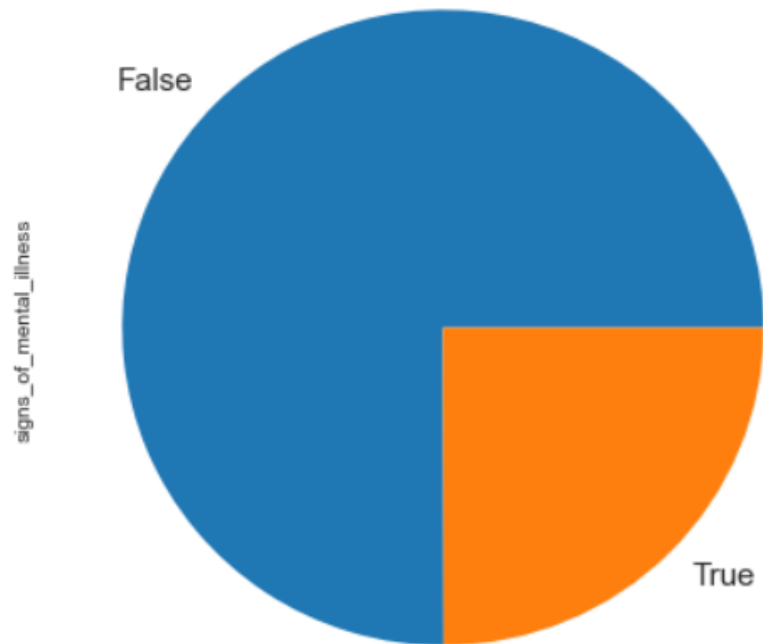


Fallecidos por genero

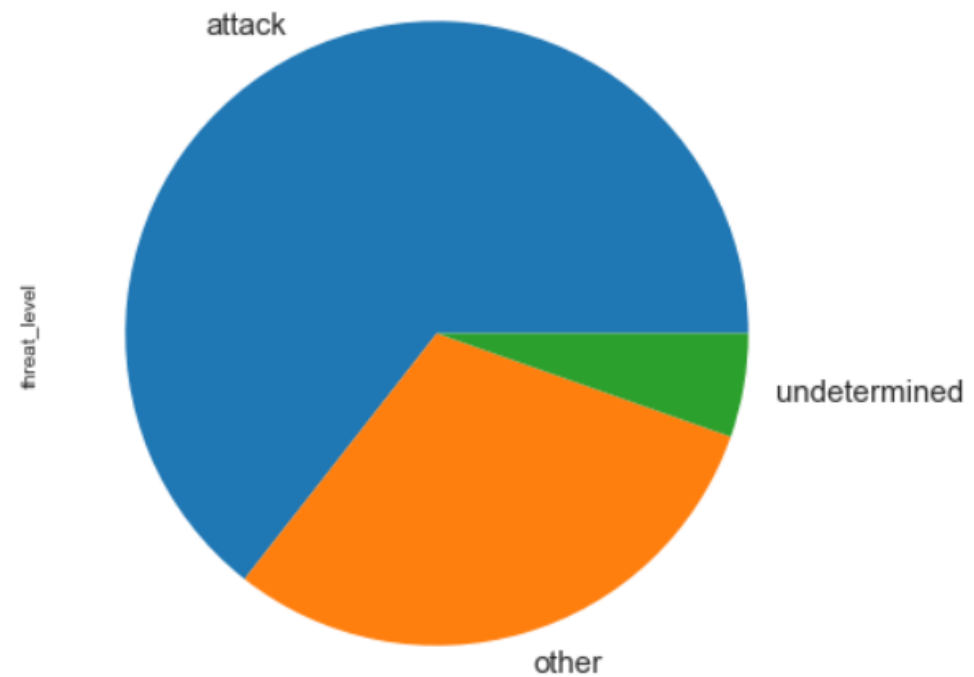


EDA

Trastornos psicológicos

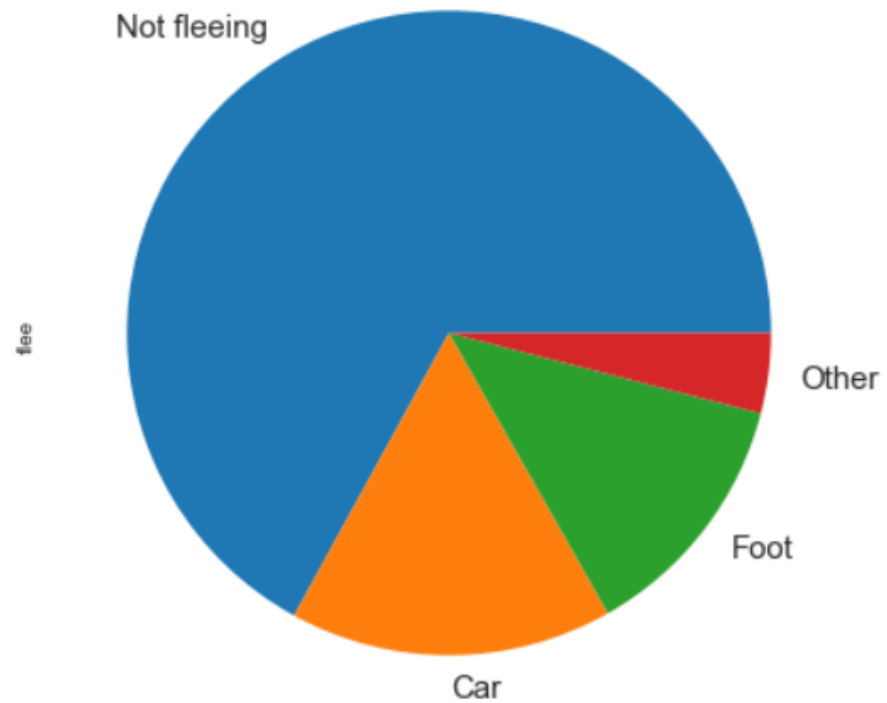


Nivel de amenaza reportado por la policía

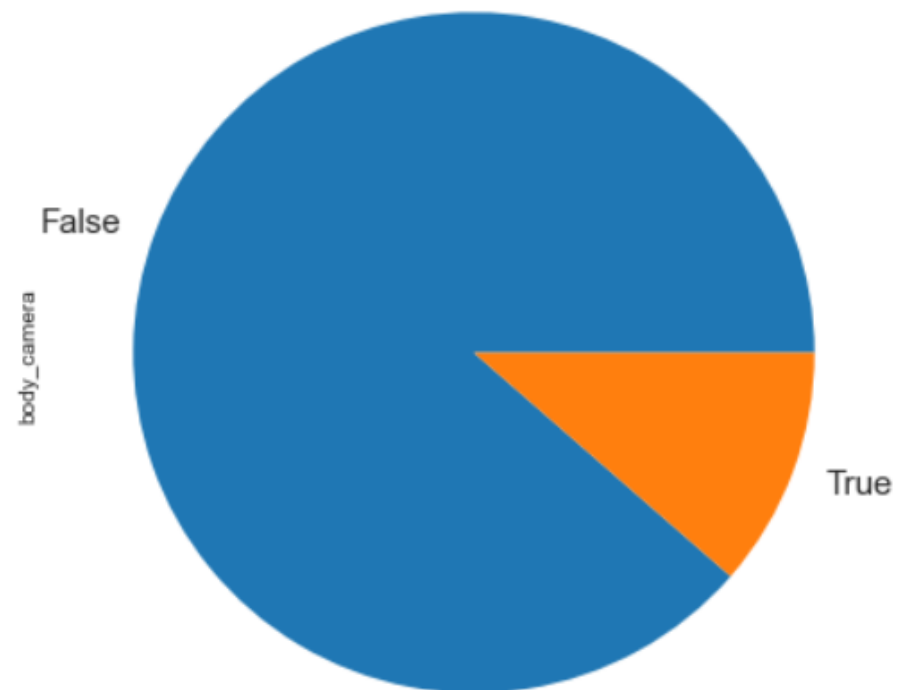


EDA

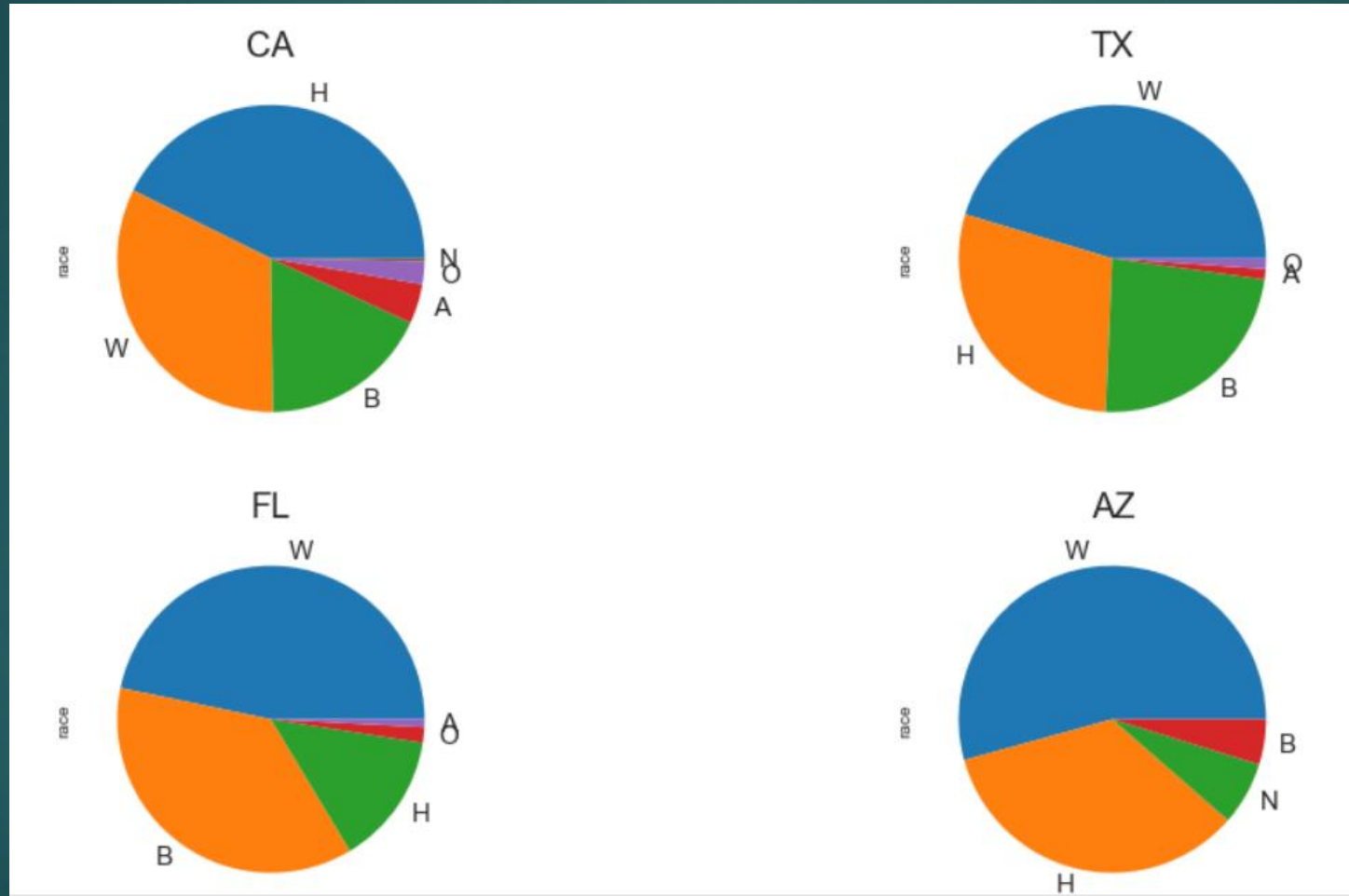
El sujeto estaba huyendo



El policia llevaba una camara



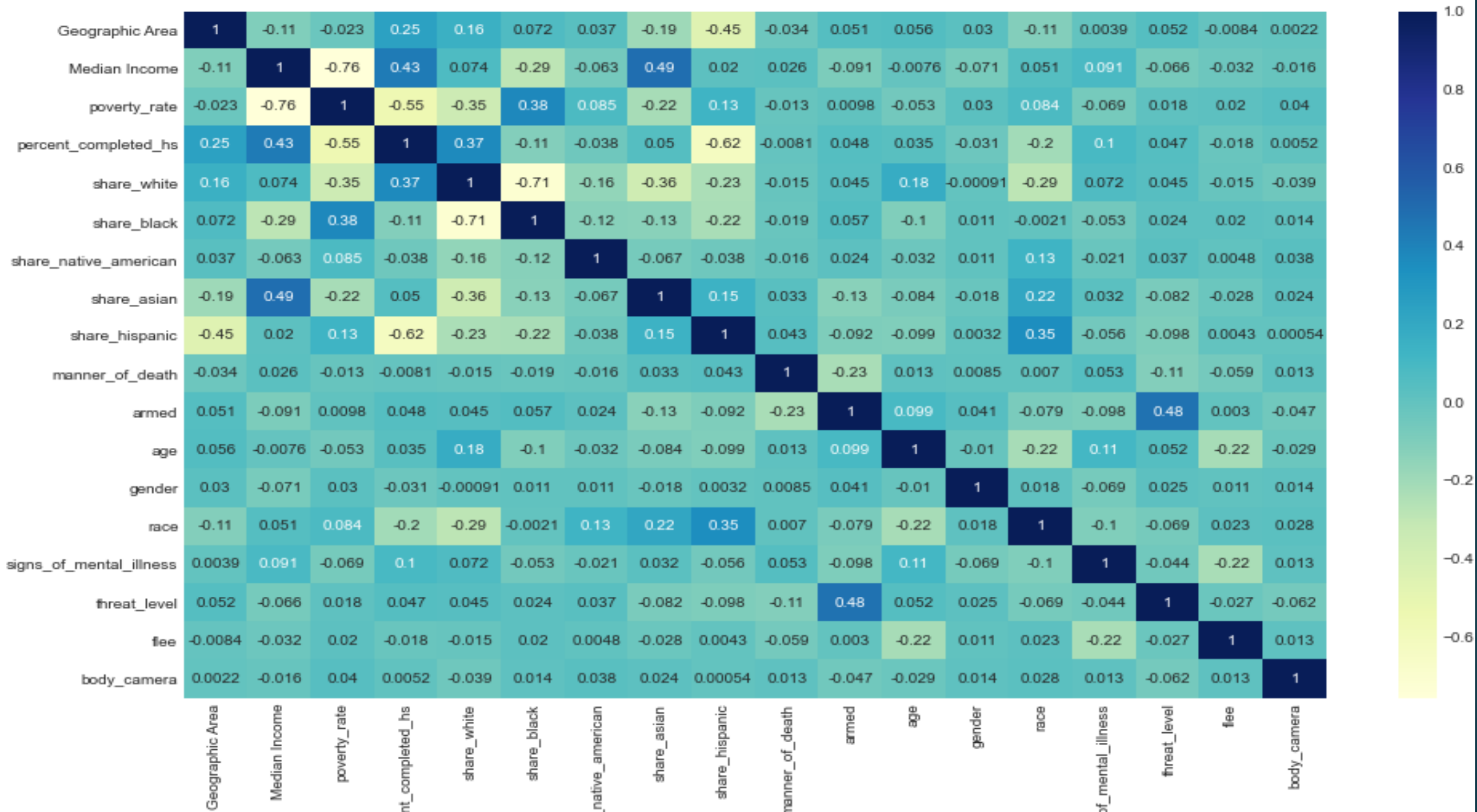
Distribucion etnias fallecidos por estado entre los 4 estados donde hubo mas de 100 fallecidos despues de un intercambio violento con la policia



Primeros resultados:

- ▶ La población que se identifica como “Blanca” parece ser, en general la que más fallecidos sufrió entre 2015 y 2017. Esto es consistente con el hecho que la comunidad mayoritaria en EE.UU. es “blanca”. En segundo lugar está la comunidad afroamericana, y en 3ero la comunidad latina.
- ▶ En California sin embargo, que también es el estado que más fallecidos reporta, la comunidad que más fallecidos sufrió fue la latina.
- ▶ Además, parece haber una correlación entre aquellos estados donde el ingreso medio es mayor, con niveles más bajos de fallecidos, y con niveles de educación superiores.

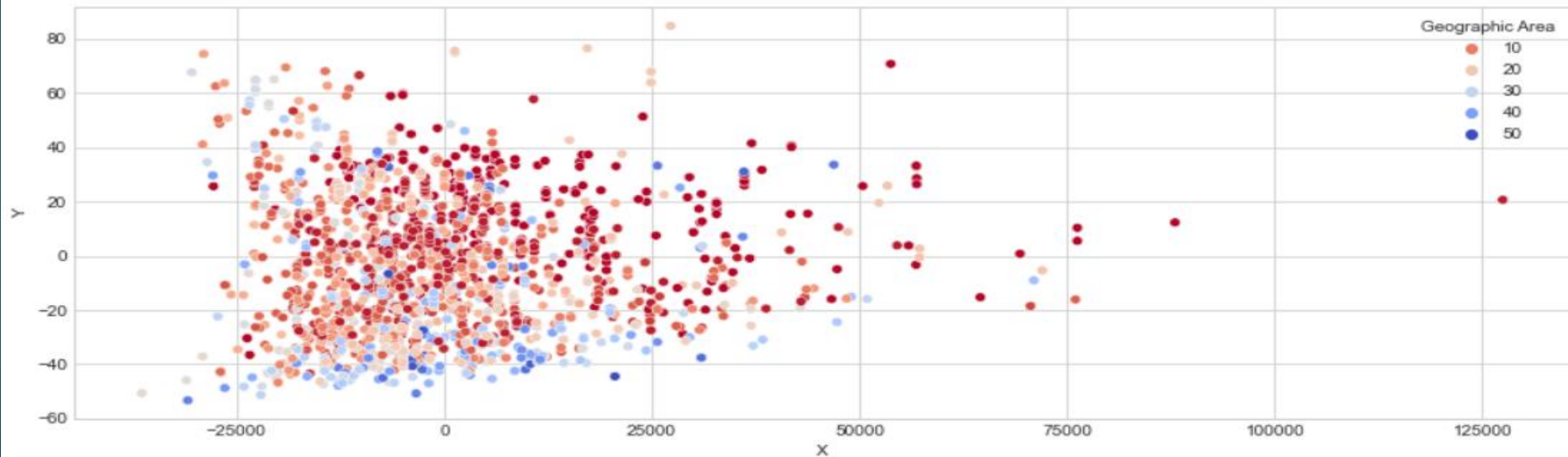
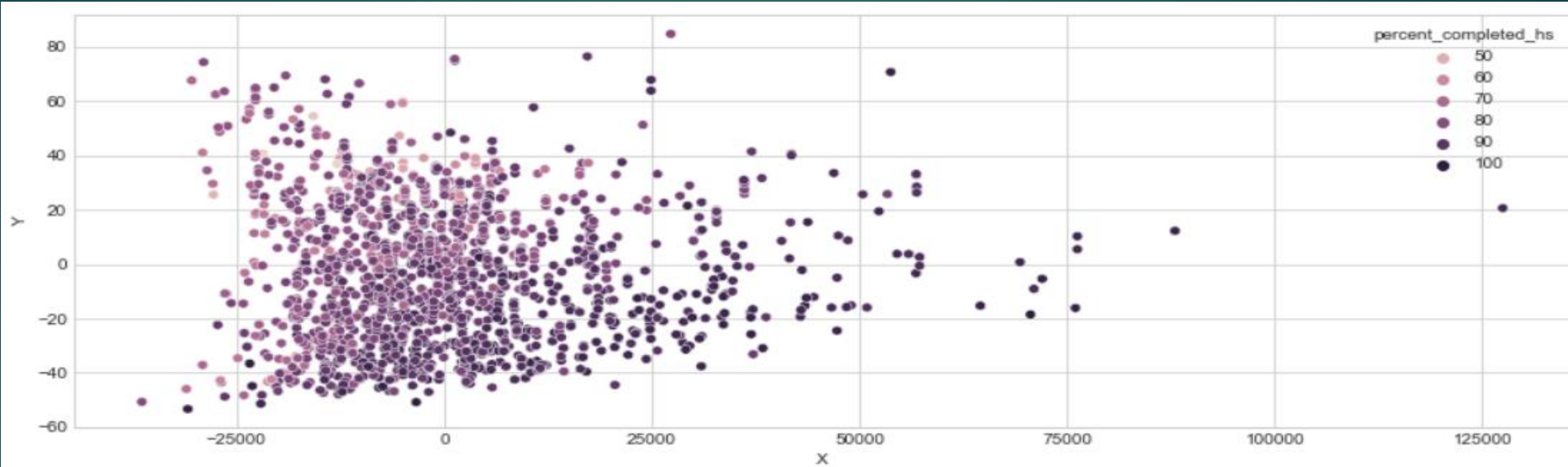
Matriz de correlacion



Agrupamiento con IsoMap

- ▶ Se aplicaron la mayoría de algoritmos estudiados en clase al dataset, como T-SNE, MDS, LLE, entre otros, sin embargo el que mejores resultados dio fue IsoMap.
- ▶ Es posible que esto sea gracias al hecho de preservar las distancias geodésicas.

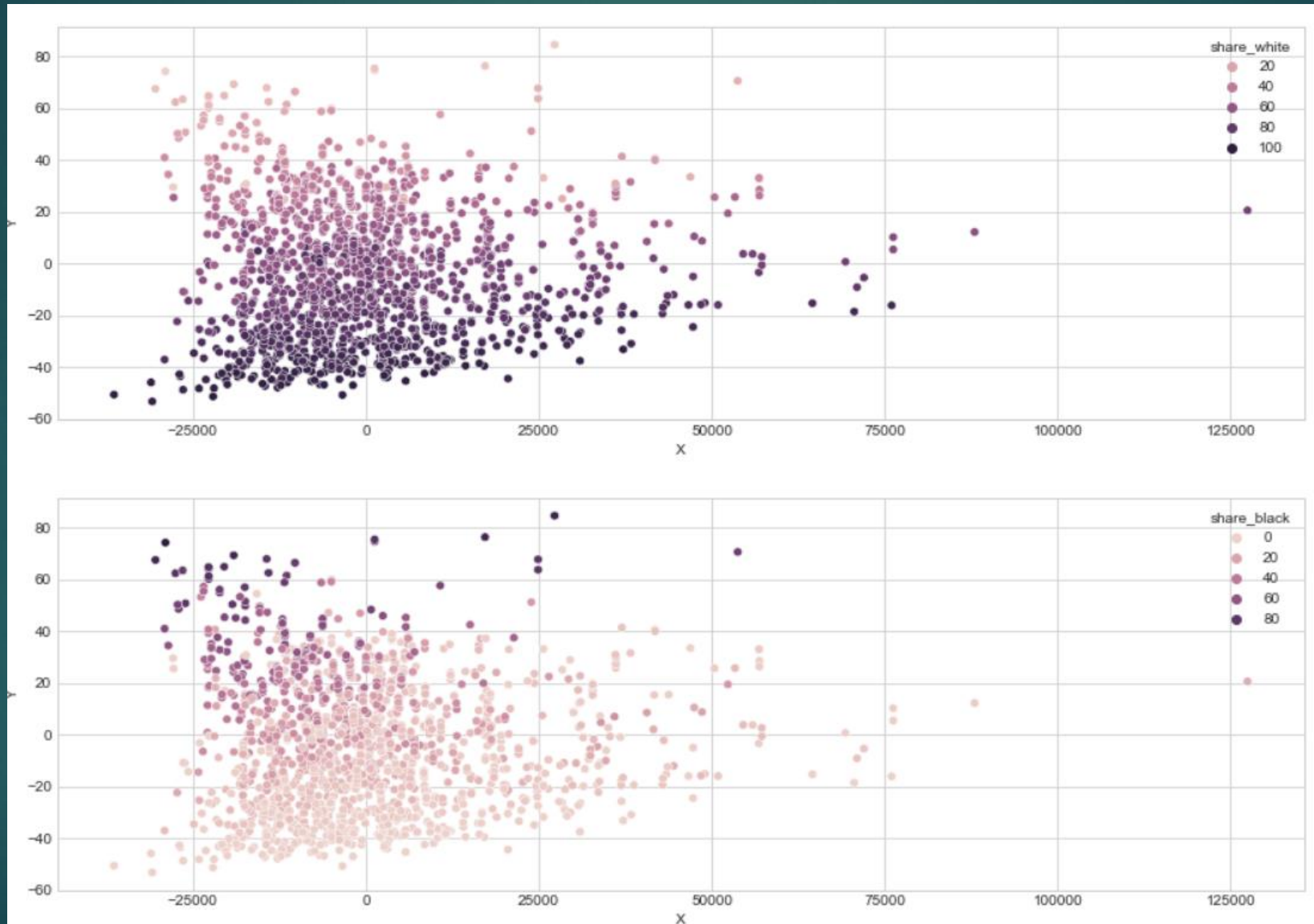
ISOMAP



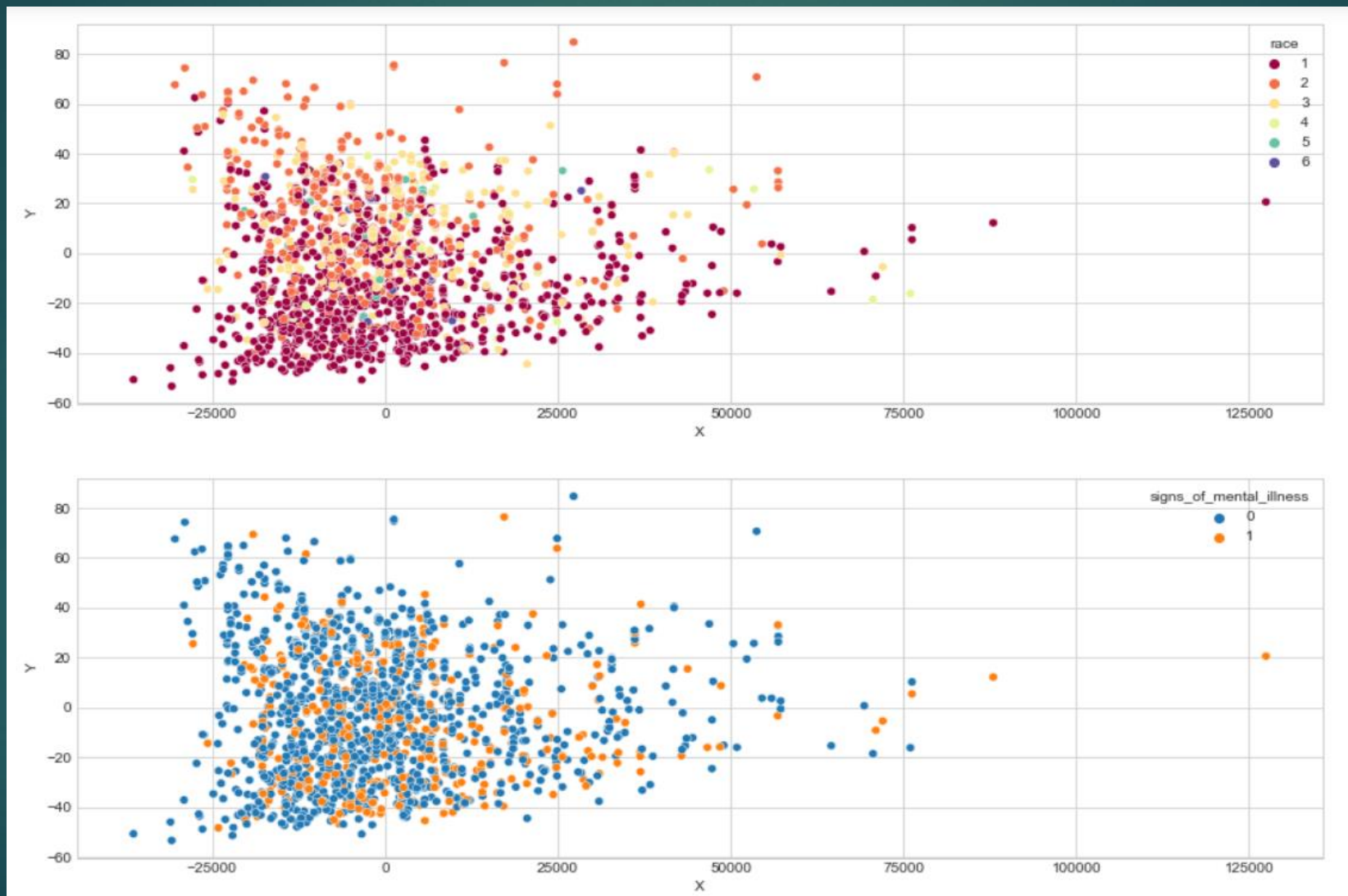
ISOMAP



ISOMAP



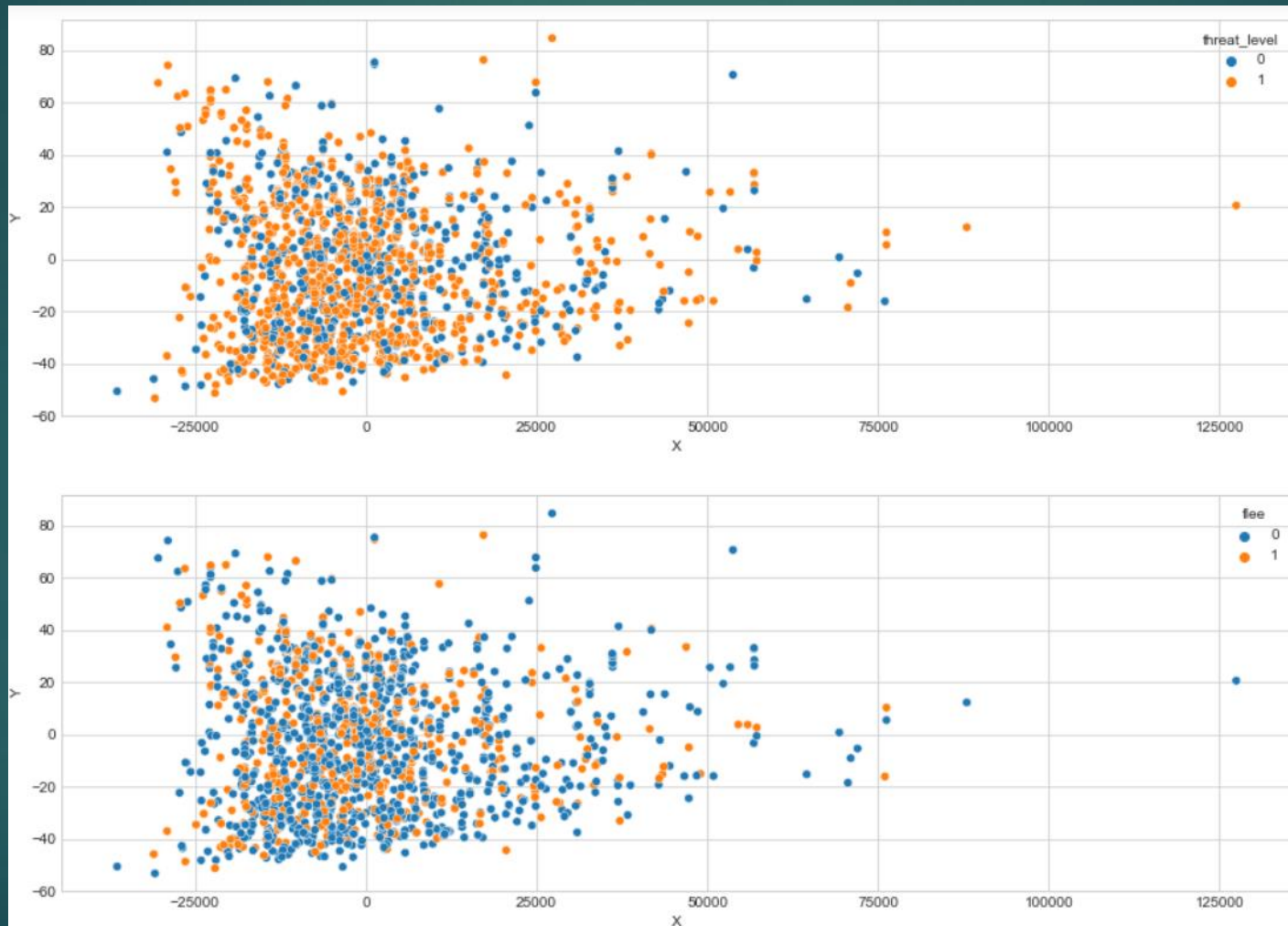
ISOMAP



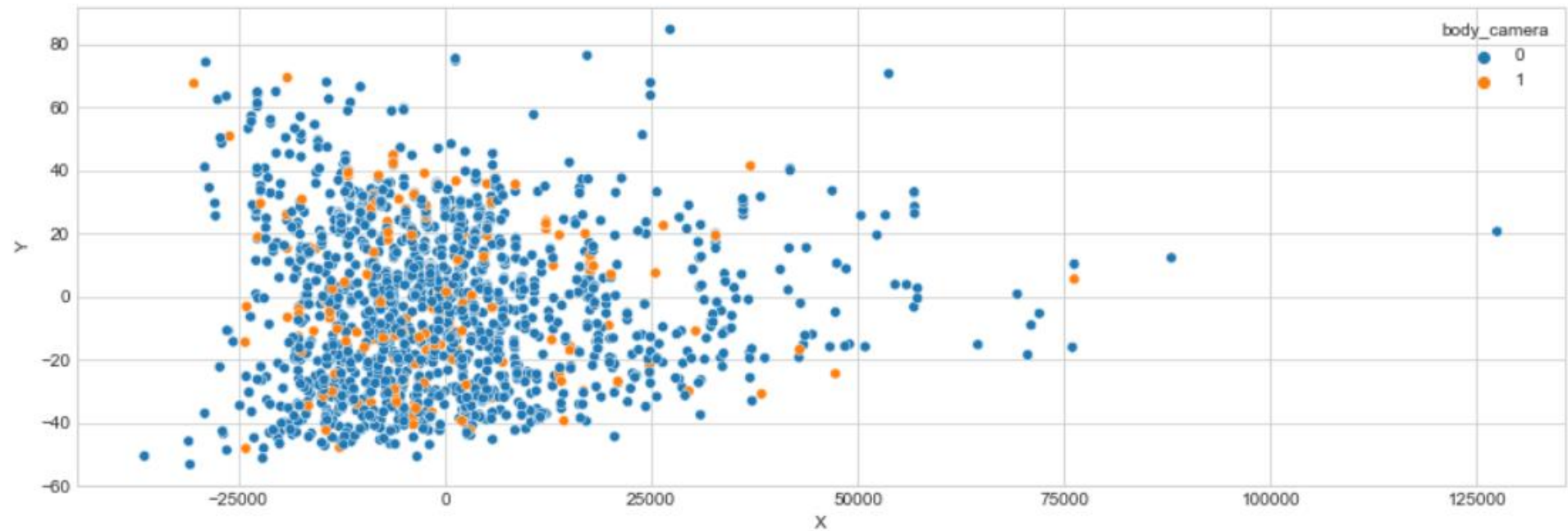
ISOMAP



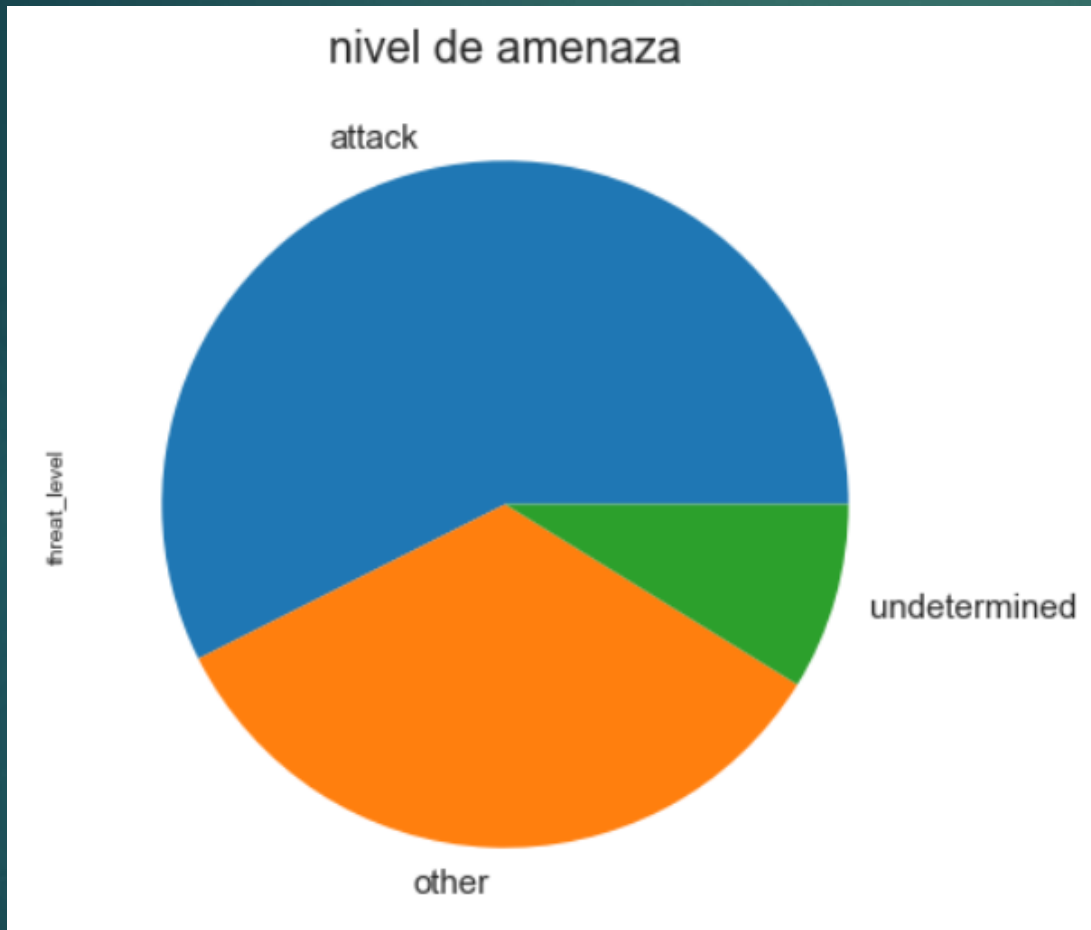
ISOMAP



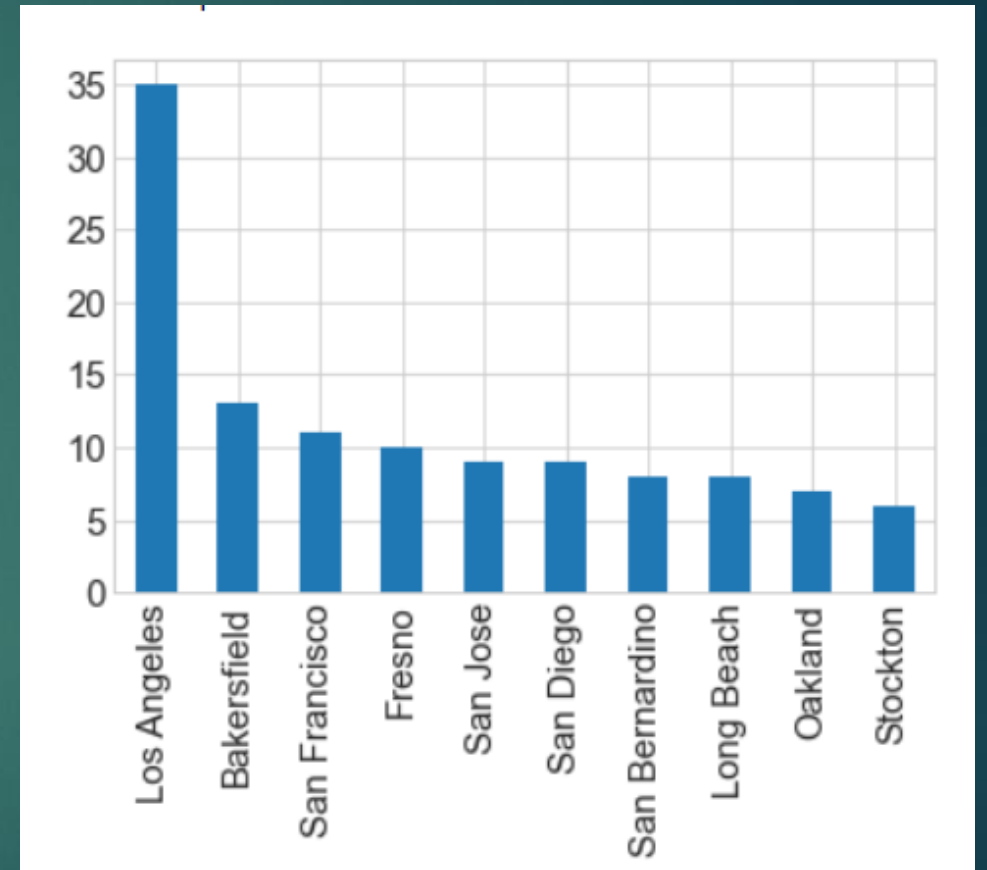
ISOMAP



Datos adicionales

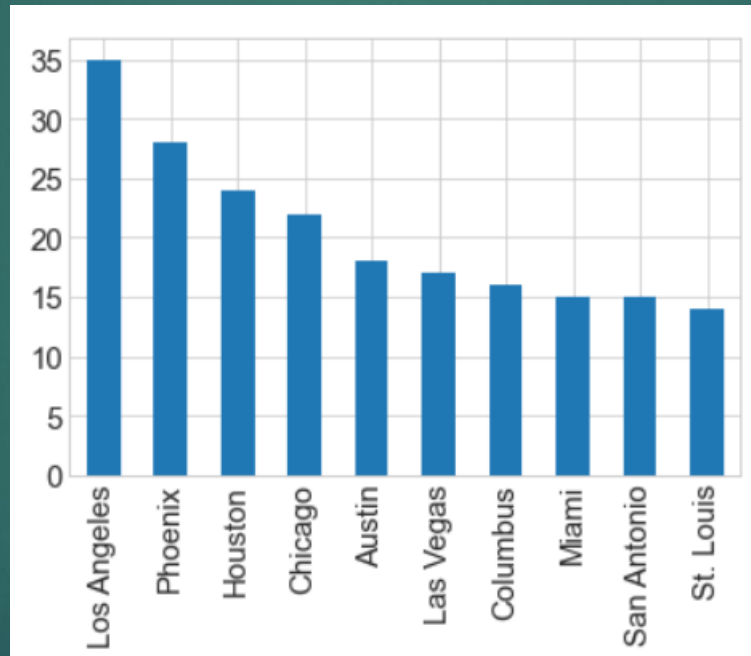


De los hispanos fallecidos, estaban atacando?



10 ciudades de CA que más fallecidos reportan

A nivel global



Conclusiones

- ▶ Hay una segregación notoria entre las comunidades blancas y las demás, en particular con la comunidad afroamericana.
- ▶ Esto parece indicar que la segregación entre la comunidad blanca y las demás es responsable de que la violencia se concentre en ciertos estados y ciudades más que en otros.
- ▶ Aunque hay casos particulares, como California, parece haber una correspondencia entre los estados que mayores niveles de educación e ingreso medio presentan, con los estados que menos fallecidos reportan.
- ▶ La correlación entre fuerte presencia de latinos y mayores números de fallecidos podría explicarse porque en varios de estos estados es notoria la presencia de pandillas latinas conocidas por sus enfrentamientos armados con la policía.