

# Hate Crime Analysis

**Professor:** Pedro Furtado

**Disciplina:** Sistemas de Gestão de Dados

**Projeto:** Open-subject Project

**Projeto elaborado por:**

- Bruno Sequeira nº 2020235721 email - brunosequeira@student.dei.uc.pt
- Rui Santos nº 2020225542 email - rpsantos@student.dei.uc.pt

## Tema: Hate Crimes

### O que são os Hate Crimes?

Os crimes de ódio são uma manifestação extrema da intolerância e da hostilidade, que têm o poder de infligir danos profundos não apenas às vítimas diretas, mas também às suas comunidades. Estes atos, motivados por preconceitos relacionados a características como raça, etnia, religião, orientação sexual, identidade de género ou qualquer outra característica percebida como diferente, representam uma grave violação dos direitos humanos e dos princípios de igualdade e dignidade.

Em todo o mundo, indivíduos e grupos são alvo de ataques físicos, verbais e emocionais simplesmente por serem quem são. Esses ataques não apenas causam ferimentos físicos, mas também deixam cicatrizes emocionais profundas que podem durar uma vida inteira. O medo, a ansiedade e a sensação de insegurança que resultam desses atos afetam não apenas as vítimas, mas também as suas famílias e comunidades, minando a coesão social e alimentando o ciclo de ódio e violência.

É crucial reconhecer que os crimes de ódio não são apenas crimes contra indivíduos, mas também ataques contra os valores fundamentais de uma sociedade democrática e inclusiva. Quando permitimos que o ódio e a intolerância prosperem sem consequências adequadas, comprometemos os princípios de justiça e igualdade que formam a base de uma sociedade civilizada.

Combater os crimes de ódio requer uma abordagem multifacetada que envolva não apenas a aplicação rigorosa da lei, mas também a educação, a sensibilização e a promoção da tolerância e da diversidade. Isso significa criar leis e políticas que criminalizem explicitamente os crimes motivados pelo preconceito, mas também investir em programas de educação que promovam o respeito mútuo e a compreensão intercultural.

Além disso, é fundamental que as comunidades se unam em solidariedade contra o ódio, mostrando apoio às vítimas, denunciando o discurso de ódio e promovendo uma cultura de respeito e aceitação. Quando nos recusamos a aceitar o ódio como inevitável e nos comprometemos a construir uma sociedade baseada na igualdade e na dignidade de todos

os seres humanos, damos um passo importante na direção de um mundo mais justo e compassivo.

Iremos abordar agora umas questões, para avaliar se este tema é bom para analisar, tirando algumas conclusões, e contribuições que iremos dar a este tema.

- **Importância do Estudo dos Crimes de Ódio para a Sociedade:**

- Estudar os crimes de ódio é crucial para a sociedade por diversas razões. Primeiro, compreender a natureza e as causas desses crimes ajuda a desenvolver estratégias eficazes para preveni-los e combatê-los. Além disso, a análise desses crimes oferece insights valiosos sobre os padrões de discriminação e intolerância presentes numa sociedade, o que pode levar a mudanças legislativas e políticas para promover a igualdade e a justiça. Estudar o ódio também é importante para fornecer apoio adequado às vítimas e promover a cura nas comunidades afetadas.

- **O Estado Atual do Conhecimento sobre Crimes de Ódio:**

- Existem numerosos estudos e análises sobre crimes de ódio em várias partes do mundo. Algumas análises destacam a prevalência de crimes de ódio motivados por raça, religião, orientação sexual e outras características protegidas, enquanto outras exploram as tendências de aumento ou diminuição desses crimes ao longo do tempo.

- **Descobertas Relevantes e Novas Contribuições:**

- Para contribuir de forma significativa ao estudo dos crimes de ódio, iremos analisar o impacto destes ao longo dos anos nas sociedades, relacionando com outros aspetos que não encontramos noutras análises realizadas por terceiros. Esses aspetos são a densidade populacional, se este fator tem algum impacto no número de crimes de ódio. Outro aspeto bastante importante, é a riqueza dos países, se os países mais ricos são mais conservadores e têm uma percentagem de crimes menor em relação aos países menos desenvolvidos, a idade média é outro fator. Por último, a taxa de imigrantes, se esta tem impacto visto que quanto maior for a taxa de imigração implica que existem mais raças e pode ocorrer uma evolução dos crimes de ódio. Sendo esse um dado importante para os imigrantes analisarem antes de irem para um país com uma taxa de crimes de ódio relativamente alta.

- **Análise Comparativa e Relevância em Diferentes Sociedades:**

- Uma análise comparativa de crimes de ódio em diferentes sociedades pode revelar padrões globais e tendências regionais, bem como diferenças significativas entre contextos culturais. Portanto, iremos fazer uma comparação entre a Europa, (os países), com os Estados Unidos da América, (os estados). Analisar os países da europa, verificar qual deles têm uma taxa de crimes de ódio alta, analisar o tipo de crime, e o tipo de ofensa e analisar outros tipos de dados, sendo que esta análise poderá ajudar os responsáveis em cada sociedade a evoluir ajudando-os a ter uma sociedade mais respeitadora.

# Questão Principal: Como fatores demográficos e socioeconômicos influenciam as taxas de crimes de ódio nos Estados Unidos e na Europa?

A ocorrência de crimes de ódio é um fenômeno complexo e multifacetado, influenciado por uma variedade de fatores demográficos e socioeconômicos. Nos Estados Unidos e na Europa, as taxas de crimes de ódio variam significativamente, refletindo as diferenças nas estruturas sociais, políticas e econômicas de cada região. Compreender como fatores como população, densidade populacional, Produto Interno Bruto (GDP), idade média da população e taxas de imigração afetam essas taxas é crucial para desenvolver políticas eficazes de prevenção e mitigação. Este estudo busca explorar a influência desses fatores nos crimes de ódio, comparando estados dos EUA com países europeus, a fim de identificar padrões e fornecer insights sobre as causas subjacentes dessas ocorrências. Ao analisar dados demográficos e socioeconômicos detalhados, pretendemos delinear as dinâmicas que contribuem para a prevalência de crimes de ódio em diferentes contextos geográficos e culturais.

## Coleta de Dados

Para alcançar uma análise abrangente, planejamos coletar conjuntos de dados diversificados e representativos de cada região. Primeiramente, focaremos na obtenção de conjuntos de dados relacionados à Europa, dados entre 2016 e 2022 acerca dos crimes de ódio, sendo que queremos os tipos de crimes e os tipos de ofensas. Com base nos países desse dataset iremos procurar indicadores de riqueza nacional, como o PIB per capita, além de dados sobre idade média, população, área total e taxas de imigração. Em seguida, direcionaremos nossa atenção para os Estados Unidos, buscando conjuntos de dados comparáveis que nos permitam realizar uma análise paralela.

## Imports

Nesta seção iremos apresentar todas as packages que necessitamos:

```
In [ ]: import csv
import pandas as pd
import matplotlib.pyplot as plt
import matplotlib.patches as mpatches
import numpy as np
```

## DataSet encontrados e Limpeza

Tal como referido anteriormente, a nossa procura começará pela a europa.

Sendo que o dataset encontrado corresponde a esta organização: **hatecrime.osce.org**, sendo que este website apresenta dados e informações sobre crimes de ódio apresentados pelos 57 Estados participantes da Organização para a Segurança e Cooperação na Europa (OSCE), juntamente com incidentes e dados apresentados pela sociedade civil e organizações internacionais.

Apresentaremos a leitura do dataset, que apresenta os dados entre 2016 até 2022 dos crimes de ódio.

In [ ]:

```
dados = pd.read_csv("DataSet/hate_crime_europe.csv")
print(dados[:1])
```

	Date	Country	Bias motivations	Type of incident
0	2016-11	Belgium	Anti-Semitic hate crime	Violent attacks against people

	Source	Description
0	Antisemitisme.be	A man visibly identifiable as Jewish was pushe...

Com base no print, existem **6** colunas, a **data**, o **país**, o **motivo**, o **tipo de incidente**, a **fonte**, e a **descrição**.

Iremos verificar quantos países são, quantos motivos e quantos tipos de incidentes existem neste dataset.

In [ ]:

```
paises = dados[dados.columns[1]].unique()
print(paises, len(paises))
motivos = dados[dados.columns[2]].unique()
print(motivos, len(motivos))
incident = dados[dados.columns[3]].unique()
print(incident, len(incident))
```

```
['Belgium' 'Austria' 'Belarus' 'Bulgaria' 'Denmark' 'France' 'Greece'
 'United Kingdom' 'Hungary' 'Armenia' 'Italy' 'Turkey' 'Spain' 'Slovakia'
 'Russian Federation' 'Serbia' 'Slovenia' 'Georgia' 'Poland' 'Croatia'
 'North Macedonia' 'Kyrgyzstan' 'Tajikistan' 'Uzbekistan' 'Ukraine'
 'Moldova' 'Kazakhstan' 'Norway' 'Sweden' 'Switzerland' 'Lithuania'
 'Latvia' 'Germany' 'Canada' 'Albania' 'Bosnia and Herzegovina'
 'Czech Republic' 'Ireland' 'Netherlands' 'Romania' 'Turkmenistan'
 'Luxembourg' 'Cyprus' 'Montenegro' 'Estonia' 'Finland' 'Iceland'
 'Portugal' 'Azerbaijan' 'Monaco' 'Malta' 'Liechtenstein'] 52
['Anti-Semitic hate crime'
 'Anti-Semitic hate crime, Anti-LGBTI hate crime'
 'Racist and xenophobic hate crime' 'Anti-Muslim hate crime'
 'Racist and xenophobic hate crime, Anti-Muslim hate crime'
 'Anti-LGBTI hate crime' 'Anti-Christian hate crime'
 'Racist and xenophobic hate crime, Anti-LGBTI hate crime'
 'Racist and xenophobic hate crime, Anti-Roma hate crime'
 'Anti-Roma hate crime'
 'Racist and xenophobic hate crime, Anti-Semitic hate crime'
 'Disability hate crime'
 'Racist and xenophobic hate crime, Anti-Christian hate crime'
 'Gender-based hate crime'
 'Gender-based hate crime, Anti-LGBTI hate crime'
 'Anti-Semitic hate crime, Anti-Muslim hate crime'
 'Anti-Muslim hate crime, Anti-LGBTI hate crime'
 'Anti-Roma hate crime, Anti-Christian hate crime'
 'Racist and xenophobic hate crime, Disability hate crime'
 'Other hate crime based on religion or belief'
 'Anti-Roma hate crime, Anti-Semitic hate crime'
 'Racist and xenophobic hate crime, Anti-Muslim hate crime, Gender-based hate
 crime'
 'Racist and xenophobic hate crime, Anti-Semitic hate crime, Anti-Muslim hate
 crime'
 'Anti-Muslim hate crime, Gender-based hate crime'
 'Racist and xenophobic hate crime, Gender-based hate crime'
 'Disability hate crime, Anti-LGBTI hate crime'
 'Racist and xenophobic hate crime, Gender-based hate crime, Anti-LGBTI hate
 crime'
 'Anti-Christian hate crime, Gender-based hate crime']
```

'Racist and xenophobic hate crime, Other hate crime based on religion or belief'

'Racist and xenophobic hate crime, Anti-Roma hate crime, Anti-Semitic hate crime'

'Racist and xenophobic hate crime, Anti-Muslim hate crime, Anti-LGBTI hate crime'

'Anti-Roma hate crime, Gender-based hate crime'

'Anti-Christian hate crime, Anti-LGBTI hate crime'

'Disability hate crime, Gender-based hate crime'

'Anti-Semitic hate crime, Anti-Christian hate crime'

'Racist and xenophobic hate crime, Anti-Semitic hate crime, Anti-LGBTI hate crime'

'Racist and xenophobic hate crime, Anti-Semitic hate crime, Gender-based hate crime'

'Anti-Roma hate crime, Anti-Muslim hate crime'

'Racist and xenophobic hate crime, Anti-Muslim hate crime, Disability hate crime'

'Racist and xenophobic hate crime, Anti-Roma hate crime, Gender-based hate crime'

'Anti-Muslim hate crime, Disability hate crime'

'Anti-Roma hate crime, Anti-LGBTI hate crime'

'Anti-Semitic hate crime, Gender-based hate crime'

'Anti-Semitic hate crime, Anti-Muslim hate crime, Anti-Christian hate crime, Other hate crime based on religion or belief'

'Anti-Roma hate crime, Anti-Semitic hate crime, Anti-LGBTI hate crime'

'Racist and xenophobic hate crime, Anti-Roma hate crime, Anti-Semitic hate crime, Anti-LGBTI hate crime'

'Anti-Semitic hate crime, Gender-based hate crime, Anti-LGBTI hate crime'

'Racist and xenophobic hate crime, Anti-Semitic hate crime, Anti-Muslim hate crime, Anti-Christian hate crime'

'Other hate crime based on religion or belief, Gender-based hate crime, Anti-LGBTI hate crime'

'Racist and xenophobic hate crime, Anti-Roma hate crime, Anti-Muslim hate crime'

'Anti-Muslim hate crime, Anti-Christian hate crime'

'Anti-Roma hate crime, Disability hate crime'

'Anti-Semitic hate crime, Anti-Christian hate crime, Gender-based hate crime, Anti-LGBTI hate crime'

'Anti-Semitic hate crime, Disability hate crime'

'Racist and xenophobic hate crime, Anti-Semitic hate crime, Other hate crime based on religion or belief'

'Anti-Roma hate crime, Gender-based hate crime, Anti-LGBTI hate crime'

'Racist and xenophobic hate crime, Disability hate crime, Anti-LGBTI hate crime'

'Racist and xenophobic hate crime, Anti-Muslim hate crime, Anti-Christian hate crime'

'Racist and xenophobic hate crime, Anti-Semitic hate crime, Gender-based hate crime, Anti-LGBTI hate crime'

'Anti-Muslim hate crime, Other hate crime based on religion or belief'

'Anti-Semitic hate crime, Anti-Christian hate crime, Anti-LGBTI hate crime'

'Racist and xenophobic hate crime, Other hate crime based on religion or belief, Gender-based hate crime'

'Anti-Semitic hate crime, Anti-Muslim hate crime, Anti-Christian hate crime'

'Disability hate crime, Gender-based hate crime, Anti-LGBTI hate crime'

'Racist and xenophobic hate crime, Disability hate crime, Gender-based hate crime, Anti-LGBTI hate crime'] 65

['Violent attacks against people' 'Threats' 'Attacks against property'

'Threats, Violent attacks against people'

'Attacks against property, Violent attacks against people'] 5

Verificamos que existem 52 países.

Os motivos verificamos que estes não estão bem estruturados, sendo que existem alguns crimes de ódio com alguns motivos. Inicialmente são 63, mas se analisarmos melhor verificamos que são menos. É necessário limpar os dados da coluna **Bias motivation**.

Os tipos de incidentes, é exatamente o mesmo que o que referimos acima. Portanto é necessário limpar os dados desta coluna também.

## Limpeza do DataSet da Europa

In [ ]:

```
def Europe():

    FileNameClean = "OriginalDataSet/hate_crime_europeClean.csv"
    FileName = "DataSet/hate_crime_europe.csv"
    count = 0
    with open(FileName, 'r') as f, open(FileNameClean,"w") as FW:
        for linha in f:
            if count > 0:
                reader = csv.reader([linha], delimiter=',', quotechar='"')
                dados_completos = next(reader)
                date = dados_completos[0]
                divide = date.split("-")
                year = divide[0]
                if(len(divide) > 1 ):
                    month = divide[1]
                else:
                    month = ""
                country = dados_completos[1]
                continent = "Europe"
                bias_desc = dados_completos[2]
                bias = bias_desc.split(",")
                offense = dados_completos[3]
                typeoffense = offense.split(",")
                for b in bias:
                    for o in typeoffense:
                        helper = str(count) + "," + continent + "," + c
                        FW.write(helper)

            else:
                info = "incident_id,country,state_name,year,month,bias_de
                FW.write(info)
                print("Working...\n")

            count+=1
        print("Clean file Europe")
    Europe()
```

Working...

Clean file Europe

Com base no código acima, limpamos o dataset, colocamos o continente, dividimos a data pelo seu mês e ano, e dividimos os motivos, ou seja, se um caso tem cerca de 2 motivos, então foram adicionadas 2 linhas, uma para cada motivo.

Iremos analisar agora quantos motivos existem e os tipos de incidente.

In [ ]:

```
dadosEurope = pd.read_csv("OriginalDataSet/hate_crime_europeClean.csv")

países = dadosEurope[dadosEurope.columns[2]].unique()
print(países,len(países))
motivos = dadosEurope[dadosEurope.columns[5]].unique()
print(motivos,len(motivos))
incident = dadosEurope[dadosEurope.columns[6]].unique()
print(incident,len(incident))
```

```
[ 'Belgium' 'Austria' 'Belarus' 'Bulgaria' 'Denmark' 'France' 'Greece'
'United Kingdom' 'Hungary' 'Armenia' 'Italy' 'Turkey' 'Spain' 'Slovakia'
'Russian Federation' 'Serbia' 'Slovenia' 'Georgia' 'Poland' 'Croatia'
'North Macedonia' 'Kyrgyzstan' 'Tajikistan' 'Uzbekistan' 'Ukraine'
'Moldova' 'Kazakhstan' 'Norway' 'Sweden' 'Switzerland' 'Lithuania'
'Latvia' 'Germany' 'Canada' 'Albania' 'Bosnia and Herzegovina'
'Czech Republic' 'Ireland' 'Netherlands' 'Romania' 'Turkmenistan'
'Luxembourg' 'Cyprus' 'Montenegro' 'Estonia' 'Finland' 'Iceland'
'Portugal' 'Azerbaijan' 'Monaco' 'Malta' 'Liechtenstein'] 52
['Anti-Semitic hate crime' 'Anti-LGBTI hate crime'
'Racist and xenophobic hate crime' 'Anti-Muslim hate crime'
'Anti-Christian hate crime' 'Anti-Roma hate crime'
'Disability hate crime' 'Gender-based hate crime'
'Other hate crime based on religion or belief'] 9
['Violent attacks against people' 'Threats' 'Attacks against property'] 3
```

Analisando o resultado, é possível visualizar cerca de 52 países tal como anteriormente.

Temos cerca de 9 tipos de motivação, iremos apresentar os motivos, apresentando exemplos de vítimas.

### Motivos

- Anti-Semitic hate
  - membros da comunidade judaica.
- Anti-LGBTI hate
  - Agressão verbal ou física contra pessoas LGBTQ+
- Racist and xenophobic hate
  - assédio racial contra indivíduos de minorias étnicas.
- Anti-Muslim hate
  - religião islâmica
- Anti-Christian hate
  - Profanação de igrejas ou cemitérios cristãos.
- Anti-Roma hate
  - Discriminação contra a comunidade Roma
- Disability hate
  - bullying de pessoas com deficiência
- Gender-based hate
  - pessoas transgénero ou não conformes com o gênero.
- Other hate crime based on religion or belief
  - locais de culto de religiões minoritárias.

E 3 tipos de incidentes que são os seguintes:

### Tipos de incidentes

- Violent attacks against people
  - São todos os casos que envolvam a vítima diretamente, por exemplo, assalto, (simples e violentos), abusos, homicídios, raptos...
- Threats
  - Ameaças, intimidações.
- Attacks against property
  - Corresponde aos roubos, destruição de propriedade, estrago e vandalismo...

## DataSet Estados Unidos da América

Após fazermos uma limpeza do dataset da Europa, iremos abordar os estados unidos da américa.

In [ ]:

```
dados = pd.read_csv("DataSet/hate_crime.csv")
print(dados[:1])
```

```

incident_id  data_year  ori  pug_agency_name  pub_agency_unit  \
0      186845      2016  AK0010100      Anchorage      NaN

agency_type_name  state_abbr  state_name  division_name  region_name  ...  \
0           City           AK      Alaska      Pacific      West      ...

offender_race  offender_ethnicity  victim_count  offense_name  \
0           White           Unknown           1  Simple Assault

total_individual_victims  location_name  \
0              1.0  School-Elementary/Secondary

bias_desc  victim_types  multiple_offense  multiple_bia
S
0  Anti-Black or African American  Individual  S
S"
```

[1 rows x 28 columns]

Com base neste dataset, visualizamos que existem cerca de 28 colunas para cada crime cometido.

Logo, temos de comparar colunas que contenham relações com o dataset da europa. Sendo que neste caso são as colunas:

- incident\_id - nº do incidente
- Country - que é os estados unidos
- state\_name - estado
- incident\_date - dia do incidente
- bias desc - motivo/os
- offense\_name - ofensa/as

In [ ]:

```
estados = dados[dados.columns[7]].unique()
print('Estados', estados, len(estados))
motivo = dados[dados.columns[-4]].unique()
print('Motivos', len(motivo))
ofensa = dados[dados.columns[-7]].unique()
print('Ofensa', len(ofensa))
```

```
Estados ['Alaska' 'Alabama' 'Arkansas' 'Arizona' 'California' 'Colorado'
'Connecticut' 'District of Columbia' 'Delaware' 'Florida' 'Georgia'
'Iowa' 'Idaho' 'Illinois' 'Indiana' 'Kansas' 'Kentucky' 'Louisiana'
'Massachusetts' 'Maryland' 'Maine' 'Michigan' 'Minnesota' 'Missouri'
'Mississippi' 'Montana' 'Nebraska' 'North Carolina' 'North Dakota'
'New Hampshire' 'New Jersey' 'New Mexico' 'Nevada' 'New York' 'Ohio'
'Oklahoma' 'Oregon' 'Pennsylvania' 'Rhode Island' 'South Carolina'
'South Dakota' 'Tennessee' 'Texas' 'Utah' 'Virginia' 'Vermont'
'Washington' 'Wisconsin' 'West Virginia' 'Wyoming' 'Federal' 'Hawaii'
'Guam'] 53
Motivos 364
Ofensa 281
```



Existem bastantes motivos, mas tal como na europa, isto é a consequência das combinações, visto que um crime pode ter vários motivos e várias ofensas.

Portanto, iremos analisar as ofensas e estabelecer uma ligação com o dataset da europa, identificar cada tipo de motivo deste dataset com os 9 motivos do dataset anterior. E iremos fazer o mesmo para as ofensas.

Iremos fazer uma limpeza nos dados, com base no dataset anterior, para isso precisamos dos seguintes dicionários.

In [ ]:

```
# Dicionário dos tipos de motivos, relacionando os 2 datasets.
lib = {"Anti-Black or African American":"Racist and xenophobic hate crime",
      "Anti-Asian":"Racist and xenophobic hate crime",
      "Anti-American Indian or Alaska Native":"Racist and xenophobic hate cr
      "Anti-White":"Racist and xenophobic hate crime",
      "Anti-Hispanic or Latino":"Racist and xenophobic hate crime",
      "Anti-Lesbian (Female)":"Anti-LGBTI hate crime",
      "Anti-Transgender":"Gender-based hate crime",
      "Anti-Islamic (Muslim)":"Anti-Muslim hate crime",
      "Anti-Other Race/Ethnicity/Ancestry":"Racist and xenophobic hate crime
      "Anti-Gay (Male)":"Anti-LGBTI hate crime",
      "Anti-Jewish":"Racist and xenophobic hate crime",
      "Anti-Bisexual":"Racist and xenophobic hate crime",
      "Anti-Multiple Races":"Racist and xenophobic hate crime",
      "Anti-Other Religion":"Other hate crime based on religion or belief",
      "Anti-Other Christian":"Anti-Christian hate crime",
      "Anti-Catholic":"Anti-Christian hate crime",
      "Anti-Physical Disability":"Disability hate crime",
      "Anti-Sikh":"Racist and xenophobic hate crime",
      "Anti-Bisexual":"Anti-LGBTI hate crime",
      "Anti-Protestant":"Racist and xenophobic hate crime",
      "Anti-Native Hawaiian or Other Pacific Islander":"Racist and xenophob
      "Anti-Gender Non-Conforming":"Gender-based hate crime",
      "Anti-Multiple Religions":"Other hate crime based on religion or beli
      "Anti-Female":"Gender-based hate crime",
      "Anti-Mental Disability":"Disability hate crime",
      "Anti-Native Hawaiian or Other Pacific Islander":"Racist and xenophob
      "Anti-Eastern Orthodox (Russian)":"Racist and xenophobic hate crime",
      "Anti-Church of Jesus Christ":"Anti-Christian hate crime",
      "Anti-Arab":"Racist and xenophobic hate crime",
      "Anti-Lesbian":"Anti-LGBTI hate crime",
      "Anti-Heterosexual":"Gender-based hate crime",
      "Anti-Male":"Gender-based hate crime",
      "Anti-Hindu":"Other hate crime based on religion or belief",
      "Anti-Atheism/Agnosticism":"Other hate crime based on religion or beli
      "Anti-Jehovah's Witness":"Other hate crime based on religion or belief
      "Anti-Buddhist":"Other hate crime based on religion or belief",
      }

#Dicionário dos tipos de ofensas, relacionando os 2 datasets.
liboffenses = {
    "Intimidation":"Threats",
    "Destruction/Damage/Vandalism of Property":"Attacks against property"
    "Simple Assault":"Violent attacks against people",
    "Aggravated Assault":"Violent attacks against people",
    "All Other Larceny":"Attacks against property",
    "Shoplifting":"Attacks against property",
    "Rape":"Violent attacks against people",
    "Kidnapping/Abduction":"Violent attacks against people",
    "Sexual Assault With An Object":"Violent attacks against people",
    "Murder and Nonnegligent Manslaughter":"Violent attacks against peopl
```

```

"Motor Vehicle Theft":"Attacks against property",
"Theft From Motor Vehicle":"Attacks against property",
"Fondling":"Violent attacks against people",
"Purse-snatching":"Attacks against property",
"Stolen Property Offenses":"Attacks against property",
"Drug/Narcotic Violations":"Violent attacks against people",
"Burglary/Breaking & Entering":"Attacks against property",
}

```

Uma nota que podemos tirar, é que existem alguns casos em que iremos ignorar crimes em que as ofensas sejam insignificantes, isto porque a europa não contém esses dados. Por exemplo, os estados unidos contém alguns casos em que a motivação é desconhecida, o que é um dado desinteressante.

In [ ]:

```

def EUA():

    FileNameClean = "OriginalDataSet/hate_crime_EUA.csv"
    FileName = "DataSet/hate_crime.csv"
    count = 0
    with open(FileName, 'r') as f, open(FileNameClean,"w") as FW:
        for linha in f:
            if count > 0:
                reader = csv.reader([linha], delimiter=',', quotechar='"')
                dados_completos = next(reader)
                incident_id = dados_completos[0]
                country = "United States of America"
                state_name = dados_completos[7]
                incident_date = dados_completos[12]
                listdate = incident_date.split("-")
                year = listdate[0]
                month = listdate[1]
                bias_desc = dados_completos[-4]
                bias = bias_desc.split(",")
                multiple_bias = bias[0].split(";")
                offense_name = dados_completos[-7]
                multiple_offense = offense_name.split(";")
                offense_name = multiple_offense[0]
                newlinha = incident_id + "," + country + "," + state_name
                helper = newlinha

                if offense_name in liboffenses and state_name != 'Guam' and state_name != 'Virgin Islands':
                    for b in multiple_bias:
                        if b != 'Group' and b != "Unknown (offender's motivation)":
                            helper += lib[b] + "," + liboffenses[offense_name]
                            FW.write(helper)
                            helper = newlinha
                else:
                    info = "incident_id,country,state_name,year,month,bias_desc,offense_name"
                    FW.write(info)
                    print("Working..\n")
                    count+=1

            print("Clean file United States")

    EUA()

```

Working..

Clean file United States

In [ ]:

```

dadosEUA = pd.read_csv("OriginalDataSet/hate_crime_EUA.csv")

```

```

estados = dadosEUA[dadosEUA.columns[2]].unique()
print('Estados', len(estados))
motivo = dadosEUA[dadosEUA.columns[5]].unique()
print('Motivos', len(motivo))
ofensa = dadosEUA[dadosEUA.columns[6]].unique()
print('Ofensa', len(ofensa))

```

Estados 51  
 Motivos 7  
 Ofensa 3

Agora ambos os dataset apresentam 7 motivos e 3 tipos de ofensas, estas que já foram referidas anteriormente.

Após limpeza do dataset dos estados unidos, relacionando com os dados da europa, iremos juntar estes 2 num só.

In [ ]:

```

## Junção dos dados
def JoinEUAandEurope():
    FileNameClean = "OriginalDataSet/hate_crime_clean.csv"
    FileName = "OriginalDataSet/hate_crime_EUA.csv"
    count = 0
    with open(FileNameClean, "w") as FW:
        with open(FileName, 'r') as f:
            for linha in f:
                if count > 0:
                    FW.write(linha)

                else:
                    info = "incident_id, country, state_name, year, month, bia
                    FW.write(info)
                    print("Working..\n")
                    count+=1
    FileName = "OriginalDataSet/hate_crime_europeClean.csv"
    count = 0
    with open(FileName, 'r') as f:
        for linha in f:
            if count > 0:
                FW.write(linha)
            else:
                print("Working..\n")
                count+=1
    print("Join Done!\n")
JoinEUAandEurope()

```

Working..

Working..

Join Done!

In [ ]:

```

dadosclean = pd.read_csv("OriginalDataSet/hate_crime_clean.csv")
euavseurope = dadosclean[dadosclean.columns[1]].unique()
print('EUAvsEURO', len(euavseurope))
estados = dadosclean[dadosclean.columns[2]].unique()
print('Estados/Países', len(estados))
motivo = dadosclean[dadosclean.columns[5]].unique()
print('Motivos', motivo, len(motivo))
ofensa = dadosclean[dadosclean.columns[6]].unique()
print('Ofensa', len(ofensa))

```

```

EUAvsEURO 2
Estados/Países 102
Motivos ['Racist and xenophobic hate crime' 'Anti-LGBTI hate crime'
'Other hate crime based on religion or belief' 'Gender-based hate crime'
'Anti-Muslim hate crime' 'Disability hate crime'
'Anti-Christian hate crime' 'Anti-Semitic hate crime'
'Anti-Roma hate crime'] 9
Ofensa 3

```

Com estes resultados, conseguimos verificar que este dataset está pronto para análise, com cerca de 9 tipos de crimes, e cerca de 3 ofensas.

Faltando-nos apenas criar um dataset que contenha as informações acerca da taxa de imigração, nível de escolaridade e riqueza do país, entre 2016 e 2022.

É possível analisar que entre 2016 a 2022, os estados unidos têm cerca de 3 vezes mais casos de crimes de ódio do que a europa.

## Dados relevantes a população, área total, GDP per capita, idade média e taxa de imigrantes.

Para responder à pergunta inicialmente referida, é necessário encontrar dados acerca do nº de habitantes por país ao longo dos anos (Comparar a evolução deste com os crimes), e área total (esta irá nos ajudar a obter a densidade populacional), o GDP per capita para que possamos analisar se a riqueza do país tem algum impacto na realização de crimes. A idade média achamos um fator importante, visto que na teoria as pessoas com idades mais altas são mais 'racistas', do que as mais novas, visto que estas já foram criadas nestas sociedades. A taxa de imigrantes poderá ter um impacto, visto que é a junção de novas pessoas, com novas crenças, o que poderá ocorrer crimes de ódio.

### População

Para o nº de habitantes foi realizado uma busca por dados que contivessem esse número por cada ano. Após obtenção, tivemos de fazer uma limpeza.

Com base nos dataset que encontramos, tivemos de criar dicionários para nos ajudarem a organizar melhor os dados. Em que tivemos de criar o identificadores para cada estado. Visto que cada dataset tinha um identificador diferente.

```

In [ ]: libpopulation_eua = {
    "0": "Alaska",
    "2": "Alabama", "3": "Arkansas", "4": "Arizona", "7": "California", "10": "Colorado", "
    "12": "District of Columbia", "13": "Delaware", "14": "Florida", "15": "Georgia", "16
    "22": "Kentucky", "23": "Louisiana", "24": "Massachusetts", "25": "Maryland", "26": "M
    "31": "Montana", "34": "Nebraska", "35": "New Hampshire", "36": "New Jersey", "37": "N
    "42": "Oregon", "43": "Pennsylvania", "44": "Rhode Island", "45": "South Carolina",
    "51": "Vermont", "52": "Washington", "53": "Wisconsin", "54": "West Virginia",
    "55": "Wyoming"
}

libgdp = {
    "AKPCPI": "Alaska", "ALPCPI": "Alabama", "ARPCPI": "Arkansas", "AZPCPI": "Arizo
}

FileNameAreaEurope = "DataSet/PopulationDataSets/totalareaEurope.csv"
FileNameClean = "OriginalDataSet/population.csv"

```

```

FileName = "DataSet/PopulationDataSets/populationEUA.csv"
FileNameEurope = "DataSet/PopulationDataSets/populationEurope.csv"
FileNameEuropeGDP = "DataSet/gdpEurope.csv"
FileNameEUAGDP = "DataSet/gdpEua.csv"
count = 0
helper = {}
helperEurope = {}
areaEUA = {}
gdpEurope = {}
helperGdp = {}

```

O próximo código irá realizar a limpeza do dataset da população dos estados unidos.

Este dataset com caminho "DataSet/PopulationDataSets/populationEUA.csv", foi obtido apartir do link: <https://fredaccount.stlouisfed.org/datalists/322164>

```

In [ ]:
with open(FileName, 'r') as f:
    for linha in f:
        if count > 0:
            reader = csv.reader([linha], delimiter=',', quotechar='\"')
            dados_completos = next(reader)
            for num in range(len(dados_completos)+1):
                value = num
                if str(value) in libpopulation_eua.keys():
                    if libpopulation_eua[str(value)] in helper.keys():
                        helper[libpopulation_eua[str(value)]].append(dados_co
                    else:
                        helper[libpopulation_eua[str(value)]] = [dados_co

            else:
                reader = csv.reader([linha], delimiter=',', quotechar='\"')
                dados_completos = next(reader)

                count+=1

```

Iremos realizar o mesmo procedimento para a europa. Os dados da Europa estão a ser guardados num dicionário "HelperEurope". O dataset escolhido foi do seguinte site, em que colocamos os países da europa e os anos entre 2016 e 2022.

<https://databank.worldbank.org/reports.aspx?source=2&series=SP.POP.TOTL&country=ECS>

```

In [ ]:
count = 0
with open(FileNameEurope, 'r') as f:
    for linha in f:
        if count > 0:
            reader = csv.reader([linha], delimiter=',', quotechar='\"')
            dados_completos = next(reader)
            if dados_completos[2] != 'World' and dados_completos[2] != 'G
                helperEurope[dados_completos[2]] = [dados_completos[8], da
            else:
                count+=1

```

### Total Área

A área total é a soma de todas as terras delimitados por fronteiras internacionais ou litorais. A medida usada para medir a área dos países é o quilômetro quadrado (km²), que corresponde a 100 hectares (ha) ou 1 milhão de metros quadrados (m²).

Este dados é importante para o cálculo da densidade populacional.

Os dados foram obtidos pela wikipedia em que apresenta a área para cada estado dos Estados Unidos da América.

[https://simple.wikipedia.org/wiki/List\\_of\\_U.S.\\_states\\_and\\_territories\\_by\\_area](https://simple.wikipedia.org/wiki/List_of_U.S._states_and_territories_by_area)

In [ ]:

```
FileNameAreaUSA = "DataSet/PopulationDataSets/totalareaEUA.csv"
count = 0
with open(FileNameAreaUSA,"r") as f:
    for linha in f:
        reader = csv.reader([linha], delimiter=',', quotechar='"')
        dados_completos = next(reader)
        if count >= 1:
            area = dados_completos[3].replace(",","")

            local = dados_completos[0].strip()

            areaEUA[local] = area

        count+=1
```

Fizemos exatamente o mesmo para a europa.

O dataset foi do site da população de cada país,

<https://databank.worldbank.org/reports.aspx?source=2&series=SP.POP.TOTL&country=ECS>.

Em que tivemos de escolher cada país, em termos de anos, é indiferente visto que a área é sempre a mesma independentemente do ano que se esteja.

In [ ]:

```
areaEurope = {}
with open(FileNameAreaEurope,"r") as f:
    for linha in f:
        reader = csv.reader([linha], delimiter=',', quotechar='"')
        dados_completos = next(reader)

        areaEurope[dados_completos[0].strip()] = dados_completos[2].strip()
```

### GDP per capita

O Produto Interno Bruto (PIB) per capita é uma medida que representa a média de produção económica por pessoa num determinado país ou região durante um determinado período de tempo, geralmente um ano. É calculado dividindo-se o Produto Interno Bruto (PIB) de uma região pelo número de pessoas que vivem nessa região. Quanto maior o PIB per capita, maior é a produção económica média por pessoa e, geralmente, maior é o padrão de vida e o nível de desenvolvimento económico da região.

Com base no dataset <https://databank.worldbank.org/reports.aspx?source=2&series=NY.GDP.PCAP.CD&country=#>,

tivemos de realizar a limpeza visto que este continha dados que não eram necessários para a nossa análise, por exemplo dados de 2000, e que não são relevantes para a nossa pesquisa.

In [ ]:

```
with open(FileNameEuropeGDP,"r") as f:
    for linha in f:
        reader = csv.reader([linha], delimiter=',', quotechar='"')
```

```
dados_completos = next(reader)

gdpEurope[dados_completos[0]] = dados_completos[-9:-2]
```

Realizamos o mesmo procedimento, recolhendo apenas os dados que necessitávamos.

O site foi <https://fred.stlouisfed.org/> em que juntamos todos os estados, sendo que colocamos também para os anos de 2016 a 2022.

```
In [ ]:
gpdhelper = {}
count = 0
gpdfinal = {}
with open(FileNameEUAGDP, 'r') as f:
    for linha in f:
        if count > 0:
            reader = csv.reader([linha], delimiter=',', quotechar='"')
            dados_completos = next(reader)

            for chaves, values in gpdhelper.items():
                if values not in gpdfinal:
                    gpdfinal[values] = [str(dados_completos[chaves])]
                else:
                    gpdfinal[values].append(str(dados_completos[chaves]))
            else:
                reader = csv.reader([linha], delimiter=',', quotechar='"')
                dados_completos = next(reader)

                for num in range(1, len(dados_completos)):
                    if libgdp.get(dados_completos[num]) != None:
                        gpdhelper[num] = libgdp.get(dados_completos[num])
                count+=1
```

## Ages

A idade é um fator crucial a ser considerado ao analisar crimes de ódio e diversos outros aspectos sociais. O modo como as pessoas percebem e interagem com o mundo ao seu redor é profundamente influenciado pela sua idade e estágio de desenvolvimento.

A idade pode moldar a forma como as pessoas percebem e interpretam o mundo ao seu redor. As experiências de vida variam significativamente com a idade, e essas experiências moldam as atitudes, crenças e preconceitos de uma pessoa.

A idade pode influenciar a exposição e a receptividade à mídia e à cultura dominantes. Jovens adultos podem ser mais influenciados por tendências e modismos culturais, enquanto pessoas mais velhas podem ser mais influenciadas por valores tradicionais e convencionais.

Inicialmente, iremos realizar uma limpeza dos dados da Europa.

O dataset é o seguinte:

[https://www.census.gov/data-tools/demo/idb/#/table?COUNTRY\\_YEAR=2024&COUNTRY\\_YR\\_ANIM=2024&CCODE=&popPages=BYAGE&menu=table](https://www.census.gov/data-tools/demo/idb/#/table?COUNTRY_YEAR=2024&COUNTRY_YR_ANIM=2024&CCODE=&popPages=BYAGE&menu=table)

Por motivos de indentação, tivemos de passar as vírgulas para pontos, visto que a média é um valor decimal.



```
In [ ]:
FileName = "DataSet/MeanAgesDataSets/europeages.csv"
agesEurope = {}
with open(FileName,"r") as f:
    for linha in f:
        reader = csv.reader([linha], delimiter=',', quotechar='"')
        dados_completos = next(reader)
        values = dados_completos[0]
        if values not in agesEurope:

            age = dados_completos[4].replace(",", ".")
            agesEurope[values] = [str(age)]
        else:
            age = dados_completos[4].replace(",", ".")
            agesEurope[values].append(str(age))
```

Para os Estados Unidos, ficou mais complexo, visto que não encontramos os dados todos juntos, portanto tivemos de ir buscar ano a ano, sendo que o link seguinte é o de 2016.

Tivemos de analisar anteriormente cada ficheiro, para realização de uma lógica que pudesse analisar todos os ficheiros e juntar num dicionário de listas os dados para cada estado.

<https://data.census.gov/table/ACSST1Y2016.S0101?>

[q=median%20age%20by%20state&g=010XX00US&moe=false&tp=false](https://data.census.gov/table/ACSST1Y2016.S0101?q=median%20age%20by%20state&g=010XX00US&moe=false&tp=false)

```
In [ ]:
positionStates = {}
NotDone = True
agesEUA = {}
lines = [33,35,35,35,35,35,35]
col = [3,6,6,6,6,6,6]
for i in range(7):
    file = 2016 + i
    FileName = "DataSet/MeanAgesDataSets/EUA/"+str(file)+".csv"

    with open(FileName,"r") as f:
        n_linha = 0
        for linha in f:
            reader = csv.reader([linha], delimiter=',', quotechar='"')
            dados_completos = next(reader)

            if NotDone:

                for val in range(1,len(dados_completos),col[i]):

                    state = dados_completos[val].split("!!!")
                    name = state[0]
                    positionStates[name] = val

                NotDone = False

            if n_linha == lines[i]:

                for chaves, values in positionStates.items():
                    if chaves not in agesEUA:
                        agesEUA[chaves] = [str(dados_completos[values])]
                    else:
                        agesEUA[chaves].append(str(dados_completos[values]))
                n_linha += 1
                NotDone = True
```

## Percentagem de Imigrantes por País



O número de habitantes imigrantes numa determinada região pode ter um impacto significativo na incidência e na natureza dos crimes de ódio.

A presença de imigrantes traz consigo uma diversidade de culturas, etnias e identidades para uma comunidade. Embora essa diversidade possa enriquecer uma sociedade, também pode desencadear tensões e conflitos entre grupos étnicos, especialmente em áreas onde as diferenças culturais são percebidas como uma ameaça à identidade ou ao status da comunidade majoritária.

A presença de imigrantes pode aumentar os níveis de preconceito e estereótipos em relação a grupos étnicos específicos. Isso pode ser resultado de percepções negativas alimentadas pela mídia, políticas anti-imigração ou simplesmente pelo desconhecimento das culturas e tradições dos imigrantes.

Para a europa, utilizamos este data set:

<https://ec.europa.eu/eurostat/databrowser/view/tps00178/default/table>.

In [ ]:

```
FileNameForeign = "DataSet/ForeignBornDataSets/europe.csv"

countryForeign = {}
with open(FileNameForeign,"r") as f:
    for linha in f:
        reader = csv.reader([linha], delimiter=',', quotechar=',')
        dados_completos = next(reader)
        countryForeign[dados_completos[0]] = dados_completos[1:]
```

Iremos realizar o mesmo procedimento. Para o Estados Unidos foram utilizados 7 ficheiros, visto que o dataset que encontramos era dividido por anos. O dataset é

<https://data.census.gov/>

In [ ]:

```
lines = [115,115,115,116,116,117,117]
col = [2,2,2,2,2,2,2]
positionstates = {}
stateforeign = {}
NotDone = True
for i in range(7):
    file = 2016 + i
    FileName = "DataSet/ForeignBornDataSets/"+str(file)+".csv"

    with open(FileName,"r") as f:
        n_linha = 0
        for linha in f:
            reader = csv.reader([linha], delimiter=',', quotechar='"')
            dados_completos = next(reader)
            if NotDone:
                count = 0
                for val in range(1,len(dados_completos),col[i]):

                    state = dados_completos[val].split("!!!")
                    name = state[0]
                    positionstates[name] = val
                    count += 1
                NotDone = False
            if n_linha == lines[i]:

                for chaves, values in positionstates.items():
```

```

        if chaves not in stateforeign:
            stateforeign[chaves] = [str(dados_completos[values])]
        else:
            stateforeign[chaves].append(str(dados_completos[values]))
    n_linha += 1
    NotDone = True

```

### Junção dos Dados realizados nesta secção

Após limpeza dos dados das idades, gdp per capita, número de habitantes, área da região, temos de juntar estes dados num só, para que cada país contenha as suas informações, podendo assim relacioná-los com os crimes de ódio, tirando boas conclusões.

In [ ]:

```

FileNameClean = "OriginalDataSet/population.csv"
with open(FileNameClean, 'w') as FW:
    out = "EUAvsEUROPE,state_name,pop2016,pop2017,pop2018,pop2019,pop2020"
    FW.write(out)
    for chave, valor in helper.items():
        out = "United States of America" + "," + chave
        for val in valor:
            out += "," + val

        out += "," + str(areaEUA.get(chave))

        lista = gpdfinal[chave]
        for num in lista:
            num = num.replace(",", ".")
            out += "," + num

        ageslist = agesEUA[chave]

        for ageyear in ageslist:
            ageyear = ageyear.replace(",", ".")
            out += "," + ageyear

        foreign = stateforeign[chave]

        pop_index = 0
        for forei in foreign:
            forei = forei.replace(",", "")
            out += "," + str(round(int(forei)/int(valor[pop_index]),3))
            pop_index += 1
        out += "\n"
        FW.write(out)

    for chave, valor in helperEurope.items():
        out = "Europe" + "," + chave
        for val in valor:
            out += "," + val
        out += "," + str(areaEurope.get(chave))

        lista = gdpEurope[chave]
        for num in lista:
            out += "," + num

        ageslist = agesEurope[chave]
        for ageyear in ageslist:
            out += "," + ageyear

```

```
country = countryForeign[chave]
```

```
for yea in country:
    out += "," + str(round((float(yea) / 100),3))
out += "\n"
FW.write(out)
```

In [ ]:

```
dadospopulation = pd.read_csv(FileNameClean)
print(dadospopulation)
```

	EUAvsEUROPE	state_name	pop2016	pop2017	pop2018
0	United States of America	Alaska	742575	740983	736624
1	United States of America	Alabama	4866824	4877989	4891628
2	United States of America	Arkansas	2991815	3003855	3012161
3	United States of America	Arizona	6944767	7048088	7164228
4	United States of America	California	39149186	39337785	39437463
...	...	...	...	...	...
105	Europe	Turkmenistan	5868561	5968383	6065066
106	Europe	Ukraine	45004673	44831135	44622518
107	Europe	United Kingdom	65611593	66058859	66460344
108	Europe	Uzbekistan	31847900	32388600	32956100
109	Europe	Malta	455356	467999	484630

	pop2019	pop2020	pop2021	pop2022	totalarea	...	age2020	age2021
0	733603	732964	734923	733276	1723337.0	...	35.0	35.6
1	4907965	5031864	5050380	5073903	135767.0	...	39.4	39.8
2	3020985	3014348	3028443	3046404	137732.0	...	38.8	38.5
3	7291843	7186683	7272487	7365684	295234.0	...	38.3	38.6
4	39437610	39503200	39145060	39040616	423972.0	...	37.0	37.6
...	...	...	...	...	...	...	...	...
105	6158420	6250438	6341855	6430770	469930.0	...	29.4	29.8
106	44386203	44132049	43822901	38000000	579400.0	...	41.3	41.6
107	66836327	67081234	67026292	66971395	241930.0	...	40.3	40.4
108	33580350	34232050	34915100	35648100	440653.0	...	28.1	28.3
109	504062	515332	518536	531113	316.0	...	42.3	42.6

	age2022	populationIMI2016	populationIMI2017	populationIMI2018	\
0	35.9	0.019	0.020	0.021	
1	39.6	0.008	0.008	0.009	
2	38.9	0.010	0.011	0.011	
3	38.8	0.022	0.025	0.028	
4	37.9	0.037	0.042	0.046	
...	...	...	...	...	
105	30.3	0.005	0.006	0.007	
106	43.4	0.041	0.042	0.043	
107	40.5	0.134	0.135	0.136	
108	28.5	0.008	0.009	0.010	
109	42.9	0.178	0.182	0.187	

	populationIMI2019	populationIMI2020	populationIMI2021	\
0	0.025	0.025	0.025	
1	0.011	0.011	0.012	
2	0.013	0.013	0.013	
3	0.030	0.030	0.029	
4	0.048	0.048	0.054	
...	...	...	...	
105	0.008	0.009	0.010	
106	0.044	0.045	0.046	
107	0.138	0.140	0.141	
108	0.011	0.012	0.013	
109	0.191	0.194	0.198	

	populationIMI2022
0	0.027

1	0.015
2	0.017
3	0.036
4	0.061
..	...
105	0.011
106	0.047
107	0.143
108	0.014
109	0.202

[110 rows x 31 columns]

Os dados já estão nos ficheiros, e temos a informação necessária para responder às perguntas. Mas antes disso, iremos criar uma base de dados para importação destes dados.

## Criação da Base de Dados para Importação dos Dados

Após limpeza dos dados, juntando tudo em 2 ficheiros, sendo que um deles contém os crimes com informações acerca do país, o tipo de crime a razão, o outro contém dados sobre o país tais como GDP per capita, idade média dos países, o número de habitantes, o número da área total entre outros.

Os nossos dados irão ser guardados numa base de dados Postgres, e irão ser apresentados os resultados com o uso do matplotlib.

```
In [ ]: import psycopg2

host = "localhost"
dbname = "hate_crime"
user = "postgres"
password = "postgres"

conn = psycopg2.connect(host=host, dbname=dbname, user=user, password=password)

cur = conn.cursor()
create_table_query = '''
CREATE TABLE IF NOT EXISTS datastates_country (
    EUAvsEurope VARCHAR(100),
    state_name VARCHAR(100),
    pop2016 INTEGER,
    pop2017 INTEGER,
    pop2018 INTEGER,
    pop2019 INTEGER,
    pop2020 INTEGER,
    pop2021 INTEGER,
    pop2022 INTEGER,
    totalarea NUMERIC,
    gdp2016 NUMERIC,
    gdp2017 NUMERIC,
    gdp2018 NUMERIC,
    gdp2019 NUMERIC,
    gdp2020 NUMERIC,
    gdp2021 NUMERIC,
    gdp2022 NUMERIC,
    age2016 NUMERIC(5, 2),
    age2017 NUMERIC(5, 2),
    age2018 NUMERIC(5, 2),
```

```

age2019 NUMERIC(5, 2),
age2020 NUMERIC(5, 2),
age2021 NUMERIC(5, 2),
age2022 NUMERIC(5, 2),
populationIMI2016 FLOAT,
populationIMI2017 FLOAT,
populationIMI2018 FLOAT,
populationIMI2019 FLOAT,
populationIMI2020 FLOAT,
populationIMI2021 FLOAT,
populationIMI2022 FLOAT
);
'''

cur.execute(create_table_query)

create_table_query = '''
CREATE TABLE IF NOT EXISTS crimes (
    incident_id INTEGER,
    country VARCHAR(100),
    state_name VARCHAR(100),
    year INTEGER,
    month INTEGER,
    bias_desc TEXT,
    offense_name TEXT
);
'''

cur.execute(create_table_query)

```

O próximo código irá colocar os dados que foram limpos anteriormente nas tabelas, se estas tabelas já contiverem dados então será ignorado.

```

In [ ]:
cur.execute("SELECT COUNT(*) FROM crimes")
count = cur.fetchone()[0]

if count == 0:

    with open('OriginalDataSet/hate_crime_clean.csv', 'r') as f:
        next(f)
        cur.copy_from(f, 'crimes', sep=',', null='', columns=('incident_id',

    print("Dados enviados com sucesso!")
else:
    print("Tabela com dados!")

conn.commit()

cur.execute("SELECT COUNT(*) FROM datastates_country")
count = cur.fetchone()[0]

if count == 0:

    with open('OriginalDataSet/population.csv', 'r') as f:

        next(f)
        cur.copy_from(f, 'datastates_country', sep=',', null='', columns=('eu

    print("Dados enviados com sucesso!")

```

```

else:
    print("Tabela com dados!")
    conn.commit()

```

Tabela com dados!  
Tabela com dados!

## Análise de Resultados

Com os dados que contemos, poderemos analisar para responder à pergunta inicial. Serão apresentadas 10 perguntas, e iremos analisar o resultado de modo a responder à pergunta inicial, se fatores demográficos e socioeconômicos influenciam as taxas de crimes de ódio nos Estados Unidos e na Europa, realizando a análise comparando estes 2.

### Qual a distribuição dos crimes de ódio entre os Estados Unidos da América e a Europa?

A distribuição dos crimes de ódio entre os Estados Unidos da América (EUA) e a Europa é uma questão crucial para a análise e compreensão desse fenómeno complexo e multifacetado. Examinar as diferenças e semelhanças na incidência e nas características dos crimes de ódio em diferentes regiões pode fornecer insights valiosos para a formulação de políticas eficazes e a implementação de medidas preventivas.

Iremos analisar com base na tabela crimes.

In [ ]:

```

cur.execute("select count(distinct(incident_id)),country from crimes group by
rows = cur.fetchall()

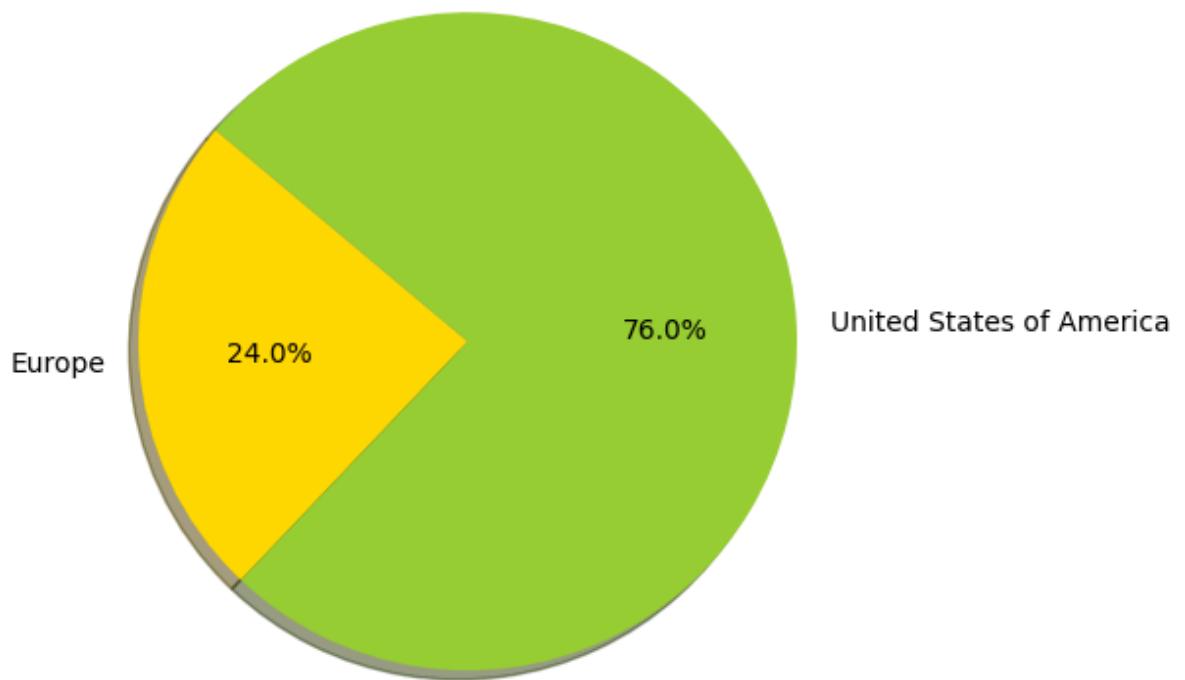
labels = [rows[0][1],rows[1][1]]
sizes = [rows[0][0], rows[1][0]]
colors = ['gold', 'yellowgreen']

plt.pie(sizes, labels=labels, colors=colors,
        autopct='%1.1f%%', shadow=True, startangle=140)

plt.axis('equal')
plt.title('Nº de casos de crimes de ódio')
plt.show()
print(rows)

```

## Nº de casos de crimes de ódio



```
[ (18112, 'Europe'), (57297, 'United States of America') ]
```

De acordo com os dados disponíveis, a Europa registrou 18.112 incidentes de crimes de ódio, enquanto os Estados Unidos da América reportaram um total de 57.297 incidentes.

Com 57.297 incidentes, os EUA apresentam um número substancialmente maior de crimes de ódio em comparação com a Europa. Esse volume pode ser atribuído a uma combinação de fatores, incluindo a diversidade étnica e racial, tensões sociais e políticas.

Na Europa, os 18.112 incidentes reportados também indicam um problema significativo, embora em menor escala numérica comparado aos EUA. A diversidade de culturas, línguas e políticas de imigração entre os países europeus influencia a dinâmica dos crimes de ódio na região.

## Qual a distribuição dos tipos de crimes (razões) entre EUA e Europa?

A análise dos tipos e razões dos crimes de ódio é fundamental para compreender a natureza e a dinâmica desses atos em diferentes regiões. Crimes de ódio, motivados por preconceitos contra raça, religião, orientação sexual, etnia entre outros, representam um grave problema social que afeta a coesão e a segurança das comunidades.

Portanto, iremos analisar para cada um (EUA/Europa) a percentagem de crimes para cada tipo de razões (poderá ser racismo, anti lgbti,etc)

In [ ]:

```
query = """
select country, bias_desc, count(*) from crimes
group by country, bias_desc
order by country, bias_desc;"""

cur = conn.cursor()
cur.execute(query)
```

```

rows = cur.fetchall()
europelabels = []
europe = rows[:9]
labels= []
values = []
for eu in europe:
    labels.append(eu[1])
    values.append(eu[2])

eua = rows[10:]

plt.figure(figsize=(12, 8))

plt.pie(values, autopct='%1.1f%%', shadow=True, startangle=160,pctdistance=1.
plt.axis('equal')
plt.title('Percentagem de casos de crimes de ódio - Europe')
plt.legend(labels,
            title='Tipo de Crime',
            fontsize=8,
            handletextpad=0.1,

            loc='upper left')

plt.show()
eurosoma = sum(values)
print("Percentagem de casos de crimes de ódio - Europa:\n")
for i in range(len(values)):
    print(labels[i],round((values[i]/eurosoma)*100,2), "% ")

labels= []
values = []
for eu in eua:
    labels.append(eu[1])
    values.append(eu[2])

plt.figure(figsize=(12, 8))

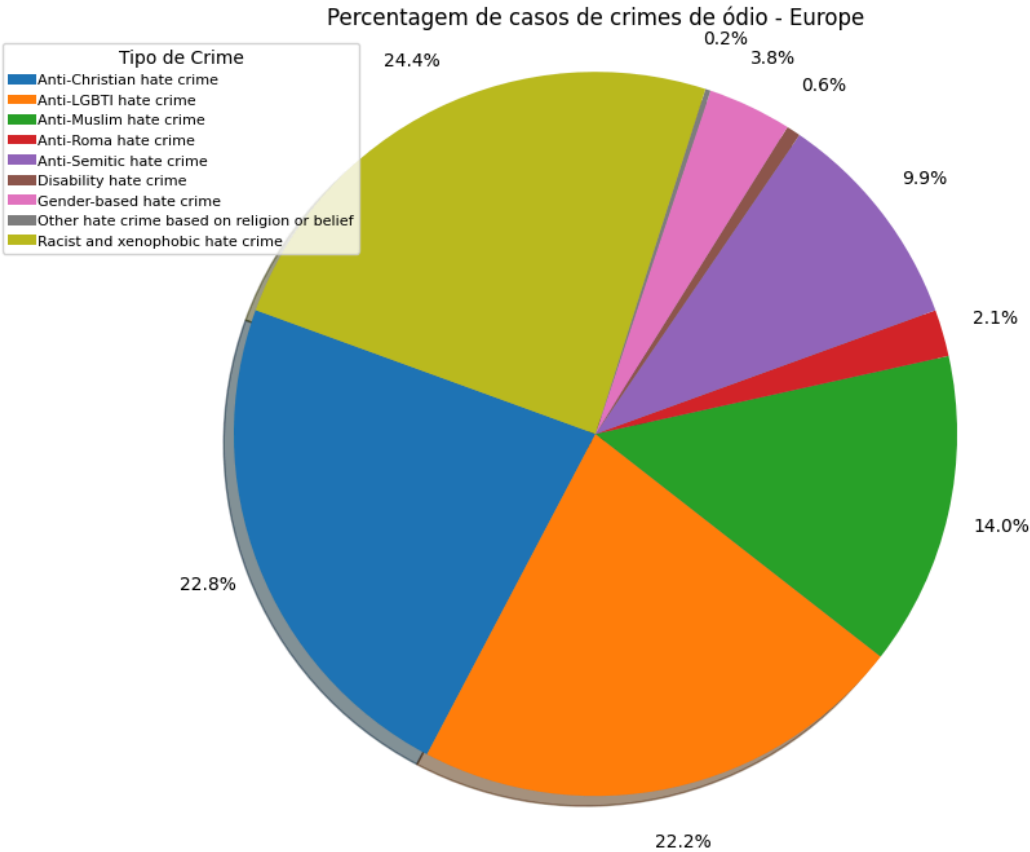
plt.pie(values, autopct='%1.1f%%', shadow=True, startangle=160,pctdistance=1.
plt.axis('equal')
plt.title('Percentagem de casos de crimes de ódio - United States of America')
plt.legend(labels,
            title='Tipo de Crime',
            fontsize=8,
            handletextpad=0.1,

            loc='upper left')

plt.show()
eurosoma = sum(values)
print("Percentagem de casos de crimes de ódio - United States of America:\n")
for i in range(len(values)):
    print(labels[i],round((values[i]/eurosoma)*100,2), "% ")

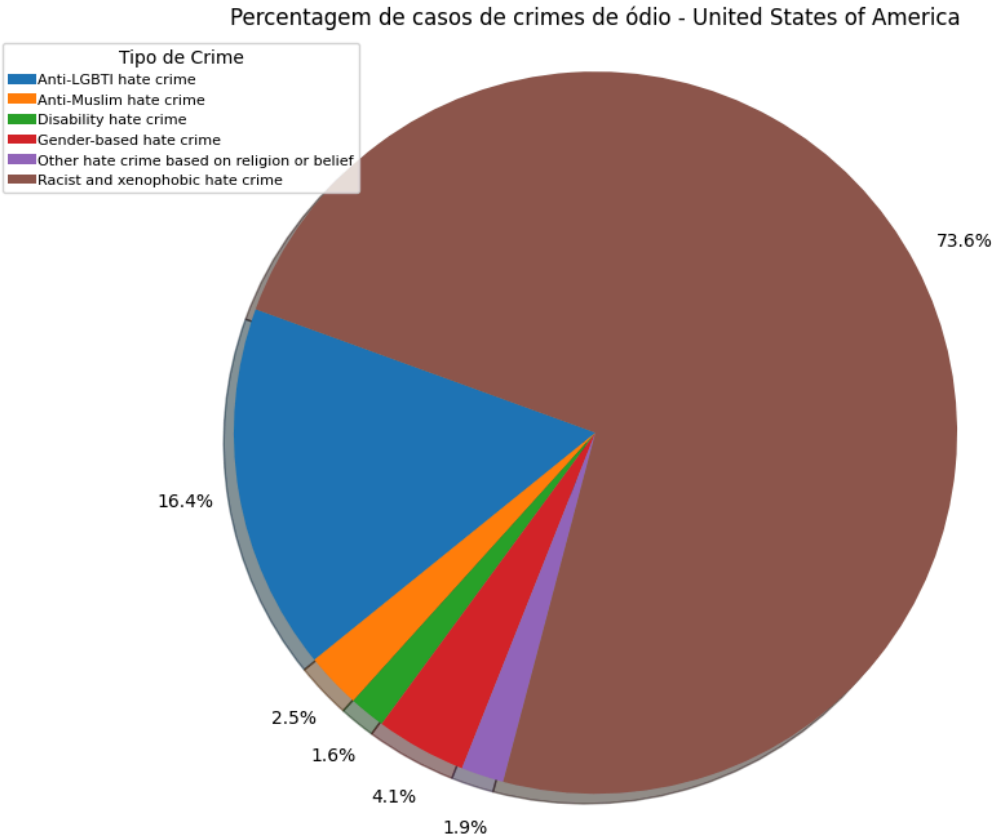
```





Percentagem de casos de crimes de ódio – Europa:

Anti-Christian hate crime 22.84 %  
Anti-LGBTI hate crime 22.17 %  
Anti-Muslim hate crime 14.0 %  
Anti-Roma hate crime 2.09 %  
Anti-Semitic hate crime 9.92 %  
Disability hate crime 0.63 %  
Gender-based hate crime 3.75 %  
Other hate crime based on religion or belief 0.24 %  
Racist and xenophobic hate crime 24.36 %



## Percentagem de casos de crimes de ódio – United States of America:

Anti-LGBTI hate crime 16.38 %  
Anti-Muslim hate crime 2.49 %  
Disability hate crime 1.61 %  
Gender-based hate crime 4.05 %  
Other hate crime based on religion or belief 1.91 %  
Racist and xenophobic hate crime 73.56 %

- **Percentagem de Casos de Crimes de Ódio na Europa**

- **Anti-Christian Hate Crime (22.84%):** A maior parte dos crimes de ódio na Europa é dirigida contra cristãos, representando quase um quarto do total. Isso pode refletir tensões religiosas internas em países com uma história de conflitos religiosos ou em regiões onde o cristianismo não é a religião majoritária.
- **Anti-LGBTI Hate Crime (22.17%):** Este tipo de crime de ódio também é significativo na Europa, indicando uma prevalência de atitudes homofóbicas e transfóbicas. A proximidade percentual com os crimes anti-cristãos sugere uma alta incidência de discriminação com base na orientação sexual e identidade de gênero.
- **Anti-Muslim Hate Crime (14.0%):** O anti islão é outra forma de ódio prevalente na Europa, refletindo possivelmente a tensão em sociedades que enfrentam desafios de integração de populações muçulmanas ou reações a eventos políticos e terroristas.
- **Anti-Roma Hate Crime (2.09%):** Embora menor em comparação com outros tipos de crimes de ódio, a discriminação contra os ciganos (Roma) destaca uma persistente marginalização desta comunidade na Europa.
- **Anti-Semitic Hate Crime (9.92%):** O anti-semitismo continua a ser um problema significativo, alcançando quase 10% do total de crimes de ódio, o que pode estar ligado a antigos preconceitos e a novos movimentos extremistas.
- **Disability Hate Crime (0.63%):** Relativamente baixo, sugerindo que a discriminação com base na deficiência não é tão comum quanto outras formas de ódio, embora qualquer incidência seja inaceitável.
- **Gender-based Hate Crime (3.75%):** Indica a presença de discriminação contra pessoas baseadas no seu gênero, refletindo a necessidade de abordar questões de igualdade de gênero.
- **Other Hate Crime Based on Religion or Belief (0.24%):** Muito baixo, o que pode indicar que outras religiões ou crenças específicas não são frequentemente alvo de crimes de ódio na Europa.
- **Racist and Xenophobic Hate Crime (24.36%):** Este é o tipo mais comum de crime de ódio na Europa, apontando para um grave problema de racismo e xenofobia que afeta diversas comunidades.
- **Percentagem de Casos de Crimes de Ódio nos Estados Unidos da América**
- **Anti-LGBTI Hate Crime (16.38%):** A discriminação contra pessoas LGBTI nos EUA também é significativa, embora menor do que na Europa, mas ainda representa uma parte importante dos crimes de ódio.

- **Anti-Muslim Hate Crime (2.49%)**: Relativamente baixo comparado à Europa, sugerindo que o anti islão pode ser menos prevalente ou reportada de forma diferente nos EUA.
- **Disability Hate Crime (1.61%)**: Similarmente baixo, mas um pouco mais alto do que na Europa, indicando algum nível de discriminação com base na deficiência.
- **Gender-based Hate Crime (4.05%)**: Ligeiramente mais alto do que na Europa, possivelmente refletindo tensões culturais específicas em torno de questões de género nos EUA.
- **Other Hate Crime Based on Religion or Belief (1.91%)**: Uma percentagem baixa, semelhante à da Europa, indicando que outras religiões ou crenças não são frequentemente alvo.
- **Racist and Xenophobic Hate Crime (73.56%)**: Extremamente alto em comparação com a Europa, apontando para um problema significativo de racismo e xenofobia nos EUA. Esta elevada percentagem sugere que a discriminação racial é a principal forma de crime de ódio nos EUA, superando todas as outras categorias combinadas.

A análise revela diferenças e semelhanças nas formas de crimes de ódio entre a Europa e os Estados Unidos:

#### **Europa:**

A distribuição dos crimes de ódio é mais diversificada, com várias formas de discriminação aparecendo em percentagens significativas. Os crimes racistas e xenofóbicos são os mais prevalentes, seguidos de perto pelos crimes anti-cristãos e anti-LGBTI.

#### **Estados Unidos:**

A grande maioria dos crimes de ódio é racial ou xenofóbica, sugerindo que o racismo é o problema mais grave e dominante. Outras formas de crimes de ódio, como anti-LGBTI, anti-musulmanos e de género, são significativamente menores, embora ainda presentes.

Essa distribuição reflete as diferentes realidades socioculturais e políticas das duas regiões. Na Europa, a diversidade de tipos de crimes de ódio pode estar ligada a uma maior variedade de grupos étnicos, religiosos e culturais. Nos EUA, a predominância esmagadora de crimes racistas aponta para tensões raciais profundamente enraizadas e históricas.

## **Qual a evolução dos tipos de crimes ao longo dos anos (EUA vs Europe)?**

A análise da evolução dos tipos de crimes de ódio ao longo dos anos é fundamental para compreender as dinâmicas sociais e culturais que moldam as atitudes e comportamentos em diferentes regiões. Nos Estados Unidos e na Europa, os crimes de ódio podem assumir várias formas, incluindo discriminação racial, religiosa, de género, contra pessoas LGBTI e contra indivíduos com deficiência. Cada região possui suas particularidades históricas, políticas e sociais que influenciam a prevalência e a natureza desses crimes.

Nos últimos anos, diversos eventos globais e locais, como crises migratórias, ataques terroristas, movimentos sociais e mudanças políticas, têm potencialmente alterado o panorama dos crimes de ódio. Analisar essas tendências ao longo do tempo nos Estados Unidos e na Europa pode revelar padrões importantes e destacar períodos de aumento ou diminuição desses crimes. Essa análise pode também indicar o impacto de políticas públicas, intervenções sociais e campanhas de conscientização sobre a intolerância e o preconceito.

```
In [ ]: query = """select bias_desc,count(*),year from crimes
group by country,bias_desc,year
order by country,bias_desc,year;
"""

cur = conn.cursor()
cur.execute(query)

rows = cur.fetchall()
europelabels = []
europe = rows[:62]
listafinal = []
labels = []
help = []
count = 0
eua = rows[63:]

for eu in europe:
    help.append(eu[1])
    count += 1
    if(count == 7):
        labels.append(eu[0])
        listafinal.append(help)
        help = []
        count = 0

print(europe)

plt.figure(figsize=(15, 8))
anos = [2016,2017,2018,2019,2020,2021,2022]
labels = ["Anti-Christian hate crime","Anti-LGBTI hate crime",
          "Anti-Muslim hate crime","Anti-Roma hate crime","Anti-Semitic hate
          "Disability hate crime","Gender-based hate crime","Other hate crime
count = 0
for tipo_crime in listafinal:

    plt.plot(anos, tipo_crime,label=labels[count])
    count+=1

plt.xlabel('Ano')
plt.ylabel('Quantidade de Crimes')
plt.title('Evolução Anual dos Crimes - Europe')
plt.legend()
plt.grid(True)
plt.show()

eurosoma = sum(values)
print("Evolução dos tipos de casos de crimes de ódio - Europe:\n")

count = 0
```

```

for tipo_crime in listafinal:
    print(labels[count],tipo_crime)
    count+=1
print("\n")

help = []
listafinal = []
eua = rows[63:]
print(eua)
count = 0
for eu in eua:
    help.append(eu[1])
    count += 1
    if(count == 7):
        labels.append(eu[0])
        listafinal.append(help)
        help = []
        count = 0

plt.figure(figsize=(15, 8))
anos = [2016,2017,2018,2019,2020,2021,2022]
labels = ["Anti-Christian hate crime","Anti-LGBTI hate crime","Anti-Muslim ha
count = 0
for tipo_crime in listafinal:

    plt.plot(anos, tipo_crime,label=labels[count])
    count+=1

plt.xlabel('Ano')
plt.ylabel('Quantidade de Crimes')
plt.title('Evolução Anual dos Crimes - United States of America')
plt.legend()
plt.grid(True)

plt.show()
print("Evolução dos tipos de casos de crimes de ódio - United States of Ameri

count = 0

for tipo_crime in listafinal:
    print(labels[count],tipo_crime)
    count+=1
print("\n")

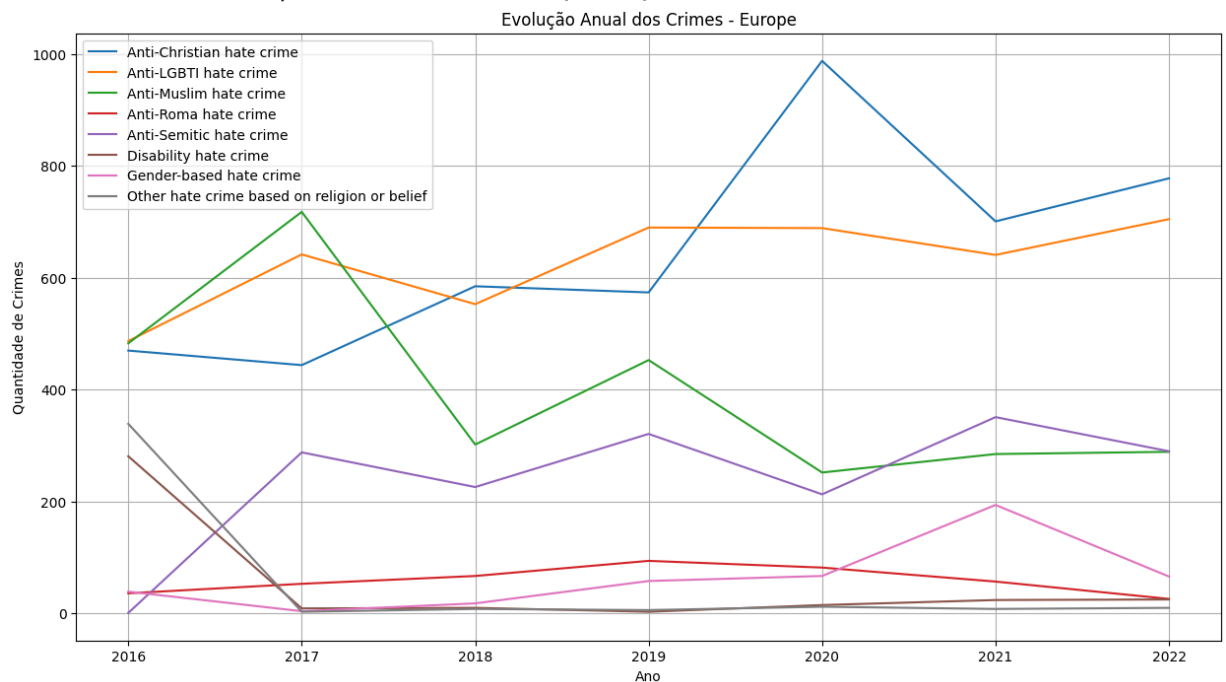
```

```

[('Anti-Christian hate crime', 470, 2016), ('Anti-Christian hate crime', 444,
2017), ('Anti-Christian hate crime', 585, 2018), ('Anti-Christian hate crim
e', 574, 2019), ('Anti-Christian hate crime', 988, 2020), ('Anti-Christian ha
te crime', 701, 2021), ('Anti-Christian hate crime', 778, 2022), ('Anti-LGBTI
hate crime', 487, 2016), ('Anti-LGBTI hate crime', 642, 2017), ('Anti-LGBTI h
ate crime', 553, 2018), ('Anti-LGBTI hate crime', 690, 2019), ('Anti-LGBTI ha
te crime', 689, 2020), ('Anti-LGBTI hate crime', 641, 2021), ('Anti-LGBTI hat
e crime', 705, 2022), ('Anti-Muslim hate crime', 483, 2016), ('Anti-Muslim ha
te crime', 718, 2017), ('Anti-Muslim hate crime', 302, 2018), ('Anti-Muslim h
ate crime', 453, 2019), ('Anti-Muslim hate crime', 252, 2020), ('Anti-Muslim
hate crime', 285, 2021), ('Anti-Muslim hate crime', 289, 2022), ('Anti-Roma h
ate crime', 36, 2016), ('Anti-Roma hate crime', 53, 2017), ('Anti-Roma hate c
rime', 67, 2018), ('Anti-Roma hate crime', 94, 2019), ('Anti-Roma hate crim
e', 82, 2020), ('Anti-Roma hate crime', 57, 2021), ('Anti-Roma hate crime', 2
6, 2022), ('Anti-Semitic hate crime', 1, 2010), ('Anti-Semitic hate crime', 2
88, 2016), ('Anti-Semitic hate crime', 226, 2017), ('Anti-Semitic hate crim
e', 321, 2018), ('Anti-Semitic hate crime', 213, 2019), ('Anti-Semitic hate c
rime', 351, 2020), ('Anti-Semitic hate crime', 290, 2021), ('Anti-Semitic hat
e crime', 281, 2022), ('Disability hate crime', 9, 2016), ('Disability hate c

```

rime', 10, 2017), ('Disability hate crime', 3, 2018), ('Disability hate crime', 15, 2019), ('Disability hate crime', 24, 2020), ('Disability hate crime', 25, 2021), ('Disability hate crime', 39, 2022), ('Gender-based hate crime', 4, 2016), ('Gender-based hate crime', 18, 2017), ('Gender-based hate crime', 58, 2018), ('Gender-based hate crime', 67, 2019), ('Gender-based hate crime', 194, 2020), ('Gender-based hate crime', 66, 2021), ('Gender-based hate crime', 339, 2022), ('Other hate crime based on religion or belief', 3, 2016), ('Other hate crime based on religion or belief', 8, 2018), ('Other hate crime based on religion or belief', 6, 2019), ('Other hate crime based on religion or belief', 12, 2020), ('Other hate crime based on religion or belief', 8, 2021), ('Other hate crime based on religion or belief', 10, 2022), ('Racist and xenophobic hate crime', 473, 2016), ('Racist and xenophobic hate crime', 607, 2017), ('Racist and xenophobic hate crime', 741, 2018), ('Racist and xenophobic hate crime', 929, 2019), ('Racist and xenophobic hate crime', 735, 2020), ('Racist and xenophobic hate crime', 569, 2021)]

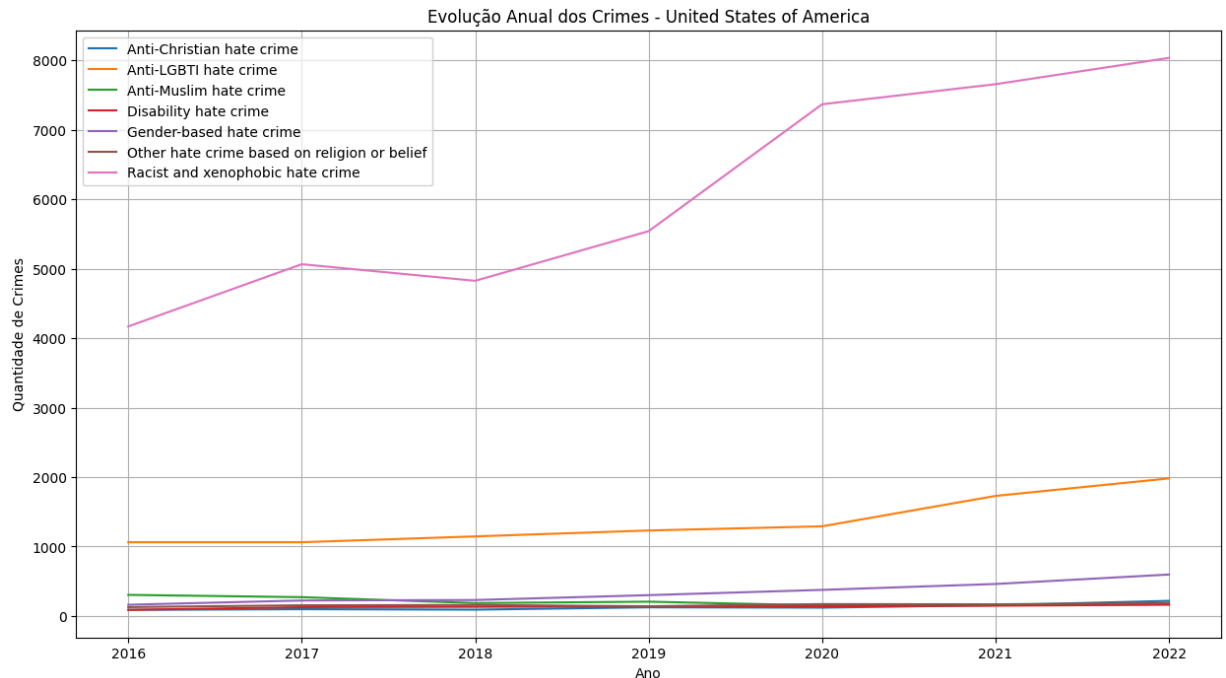


Evolução dos tipos de casos de crimes de ódio – Europe:

Anti-Christian hate crime [470, 444, 585, 574, 988, 701, 778]  
 Anti-LGBTI hate crime [487, 642, 553, 690, 689, 641, 705]  
 Anti-Muslim hate crime [483, 718, 302, 453, 252, 285, 289]  
 Anti-Roma hate crime [36, 53, 67, 94, 82, 57, 26]  
 Anti-Semitic hate crime [1, 288, 226, 321, 213, 351, 290]  
 Disability hate crime [281, 9, 10, 3, 15, 24, 25]  
 Gender-based hate crime [39, 4, 18, 58, 67, 194, 66]  
 Other hate crime based on religion or belief [339, 3, 8, 6, 12, 8, 10]

[('Anti-Christian hate crime', 88, 2016), ('Anti-Christian hate crime', 98, 2017), ('Anti-Christian hate crime', 91, 2018), ('Anti-Christian hate crime', 125, 2019), ('Anti-Christian hate crime', 120, 2020), ('Anti-Christian hate crime', 155, 2021), ('Anti-Christian hate crime', 218, 2022), ('Anti-LGBTI hate crime', 1062, 2016), ('Anti-LGBTI hate crime', 1062, 2017), ('Anti-LGBTI hate crime', 1146, 2018), ('Anti-LGBTI hate crime', 1231, 2019), ('Anti-LGBTI hate crime', 1291, 2020), ('Anti-LGBTI hate crime', 1729, 2021), ('Anti-LGBTI hate crime', 1981, 2022), ('Anti-Muslim hate crime', 304, 2016), ('Anti-Muslim hate crime', 271, 2017), ('Anti-Muslim hate crime', 186, 2018), ('Anti-Muslim hate crime', 206, 2019), ('Anti-Muslim hate crime', 152, 2020), ('Anti-Muslim hate crime', 156, 2021), ('Anti-Muslim hate crime', 171, 2022), ('Disability hate crime', 88, 2016), ('Disability hate crime', 127, 2017), ('Disability hate crime', 132, 2018), ('Disability hate crime', 138, 2019), ('Disability hate crime', 140, 2020), ('Disability hate crime', 148, 2021), ('Disability hate crime', 162, 2022), ('Gender-based hate crime', 162, 2016), ('Gender-based hate crime', 223, 2017), ('Gender-based hate crime', 230, 2018), ('Gender-based hate crime', 300, 2019), ('Gender-based hate crime', 376, 2020), ('Gender-based hate crime', 461, 2021), ('Gender-based hate crime', 597, 2022), ('Other hate crime based on religion or belief', 128, 2016), ('Other hate crime based on religion or belief', 8, 2017), ('Other hate crime based on religion or belief', 6, 2018), ('Other hate crime based on religion or belief', 12, 2019), ('Other hate crime based on religion or belief', 8, 2020), ('Other hate crime based on religion or belief', 10, 2021), ('Other hate crime based on religion or belief', 10, 2022)]

ased on religion or belief', 154, 2017), ('Other hate crime based on religion or belief', 158, 2018), ('Other hate crime based on religion or belief', 140, 2019), ('Other hate crime based on religion or belief', 171, 2020), ('Other hate crime based on religion or belief', 168, 2021), ('Other hate crime based on religion or belief', 187, 2022), ('Racist and xenophobic hate crime', 4169, 2016), ('Racist and xenophobic hate crime', 5065, 2017), ('Racist and xenophobic hate crime', 4826, 2018), ('Racist and xenophobic hate crime', 5541, 2019), ('Racist and xenophobic hate crime', 7368, 2020), ('Racist and xenophobic hate crime', 7657, 2021), ('Racist and xenophobic hate crime', 8036, 2022)]



Evolução dos tipos de casos de crimes de ódio – United States of America:

Anti-Christian hate crime [88, 98, 91, 125, 120, 155, 218]

Anti-LGBTI hate crime [1062, 1062, 1146, 1231, 1291, 1729, 1981]

Anti-Muslim hate crime [304, 271, 186, 206, 152, 156, 171]

Disability hate crime [88, 127, 132, 138, 140, 148, 162]

Gender-based hate crime [162, 223, 230, 300, 376, 461, 597]

Other hate crime based on religion or belief [128, 154, 158, 140, 171, 168, 187]

Racist and xenophobic hate crime [4169, 5065, 4826, 5541, 7368, 7657, 8036]

A análise da evolução dos tipos de casos de crimes de ódio na Europa e nos Estados Unidos ao longo dos últimos sete anos revela padrões distintos e dinâmicas específicas em cada região.

- **Europa**

- **Anti-Christian Hate Crime:** Houve uma variação significativa nos casos, começando com 470 incidentes em 2016 e atingindo um pico de 988 em 2020, seguido por uma queda para 778 em 2022. Esse aumento está relacionado a tensões religiosas específicas e eventos socio-políticos na Europa.

- **Anti-LGBTI Hate Crime:** Os casos de crimes de ódio contra pessoas LGBTI mostram uma tendência geral de crescimento, começando com 487 em 2016 e subindo para 705 em 2022. Isso pode refletir tanto um aumento na visibilidade e aceitação da comunidade LGBTI quanto uma maior incidência de discriminação e violência. O facto de que as pessoas não têm receio de mostrar quem elas são, faz com que ocorram mais crimes de ódio relacionados à comunidade LGBTI.

- **Anti-Muslim Hate Crime:** Os incidentes contra muçulmanos tiveram um pico em 2017 com 718 casos, mas depois diminuíram significativamente, estabilizando-se em torno de 289 em 2022. Isso pode ser reflexo de flutuações nas tensões religiosas e políticas, bem como na resposta governamental e social ao terrorismo e à imigração.
- **Anti-Roma Hate Crime:** Os crimes de ódio contra os Roma apresentam números relativamente baixos, com um aumento modesto de 36 em 2016 para um pico de 94 em 2019, seguido por uma queda para 26 em 2022. Isso pode indicar uma discriminação persistente, mas com uma variabilidade de ano para ano.
- **Anti-Semitic Hate Crime:** Os crimes anti-semitas tiveram uma trajetória ascendente, com um pico de 351 incidentes em 2021. Este aumento está associado a tensões geopolíticas e ao ressurgimento de movimentos extremistas na Europa.
- **Disability Hate Crime:** Embora os números sejam baixos, há uma tendência de crescimento, começando com 281 incidentes em 2016 e estabilizando-se em torno de 25 em 2022. Isso pode indicar uma maior conscientização e reporte de tais incidentes. Existem poucas ocorrências, porque a maior parte da população percebe que não foi a escolha da vítima e que esta nasceu ou ficou desabilitada ou fisicamente ou psicologicamente.
- **Gender-based Hate Crime:** Os crimes de ódio com base no gênero aumentaram significativamente, especialmente após 2019, atingindo 194 casos em 2021. Isso reflete uma crescente atenção às questões de gênero. O pico ocorreu no final da pandemia, em que a população saía à rua, e começaram a ver pessoas que se identificavam com outros gêneros o que antigamente não era muito presente na sociedade.
- **Other Hate Crime Based on Religion or Belief:** Os casos variam bastante, com uma diminuição significativa após 2016, indicando talvez uma categorização ou reporte diferente desses crimes ao longo dos anos.

## Estados Unidos

- **Anti-Christian Hate Crime:** Os casos aumentaram consistentemente de 88 em 2016 para 218 em 2022. Esse crescimento pode estar ligado a crescentes polarizações religiosas e políticas.
- **Anti-LGBTI Hate Crime:** Os incidentes aumentaram substancialmente, de 1062 em 2016 para 1981 em 2022. Este aumento pode refletir tanto maior visibilidade e aceitação da comunidade LGBTI quanto uma reação negativa, tal como referi na Europa.
- **Anti-Muslim Hate Crime:** Os crimes contra muçulmanos diminuíram de 304 em 2016 para 171 em 2022. Esta diminuição pode ser atribuída a uma variedade de fatores, incluindo mudanças nas políticas de imigração e nas relações internacionais.
- **Disability Hate Crime:** Houve um aumento nos casos, começando com 88 em 2016 e subindo para 162 em 2022. Isso pode refletir uma maior conscientização e reporte de tais crimes. Este valor é bastante parecido ao do Europa apesar de que nos Estados Unidos tem uma maior população.



- **Gender-based Hate Crime:** Os incidentes aumentaram drasticamente de 162 em 2016 para 597 em 2022, indicando uma crescente conscientização e reporte de crimes de gênero.
- **Other Hate Crime Based on Religion or Belief:** Houve flutuações nos incidentes, com um leve aumento de 128 em 2016 para 187 em 2022. Isso pode refletir uma variedade de tensões religiosas e sociais.
- **Racist and Xenophobic Hate Crime:** Este é o tipo de crime de ódio mais prevalente nos EUA, com um aumento significativo de 4169 em 2016 para 8036 em 2022. Este aumento pode ser atribuído a uma variedade de fatores, incluindo tensões raciais persistentes e políticas polarizadoras. Sendo que este corresponde a 73.6% dos tipos de crimes cometidos na América.

A evolução dos tipos de crimes de ódio na Europa e nos Estados Unidos revela padrões preocupantes e distintos. Na Europa, vemos flutuações significativas em crimes de ódio com base na religião e na orientação sexual, refletindo tensões específicas da região. Nos Estados Unidos, há um aumento pronunciado em crimes de ódio, especialmente aqueles baseados em raça e orientação sexual, indicando profundas divisões sociais e políticas. Ambas as regiões mostram uma tendência de aumento em crimes de ódio ao longo dos anos, destacando a necessidade de políticas robustas e intervenções sociais para combater a intolerância e promover a inclusão.

## Qual a distribuição dos tipos de ofensas entre EUA e Europa?

A análise da distribuição dos tipos de ofensas nos crimes de ódio é essencial para entender a natureza e a gravidade dos incidentes que ocorrem em diferentes regiões. Nos crimes de ódio, as ofensas podem variar amplamente, incluindo assaltos físicos, ameaças, vandalismo, assédio e outros tipos de violência. Compreender como esses tipos de ofensas são distribuídos entre os Estados Unidos e a Europa pode oferecer insights valiosos sobre as características específicas dos crimes de ódio em cada região e ajudar na formulação de políticas de prevenção e intervenção mais eficazes.

Neste caso, existem cerca de 3 tipos de ofensas, **Attacks againts property** que corresponde a vandalismo por exemplo, **Violent Attacks againts people** que corresponde a ataques físicos, assaltos (assaltos simples, e assaltos violentos), ataques psicológicos, assédio e por fim **Threats** que corresponde a ameaças e intimidações realizadas pelo agressor.

In [ ]:

```
query = """select offense_name,count(*),country from crimes where year >= 2016
group by offense_name,country
order by country,offense_name;"""

cur = conn.cursor()
cur.execute(query)

rows = cur.fetchall()

europe = rows[:3]
labels= []
```

```

values = []
for eu in europe:
    labels.append(eu[0])
    values.append(eu[1])

print(europe)

plt.figure(figsize=(12, 8))

plt.pie(values, autopct='%1.1f%%', shadow=True, startangle=160, pctdistance=0.
plt.axis('equal')
plt.title('Percentagem do tipo de ofensa dos crimes de ódio - Europe')
plt.legend(labels,
            title='Tipo de Ofensa',
            fontsize=8,
            handletextpad=0.1,

            loc='upper left')

plt.show()

eurosoma = sum(values)
print("Percentagem do tipo de ofensa dos crimes de ódio - Europa:\n")
for i in range(len(values)):
    print(labels[i], round((values[i]/eurosoma)*100,2), "% ")

eua = rows[3:]

labels= []
values = []
for eu in eua:
    labels.append(eu[0])
    values.append(eu[1])

plt.figure(figsize=(12, 8))

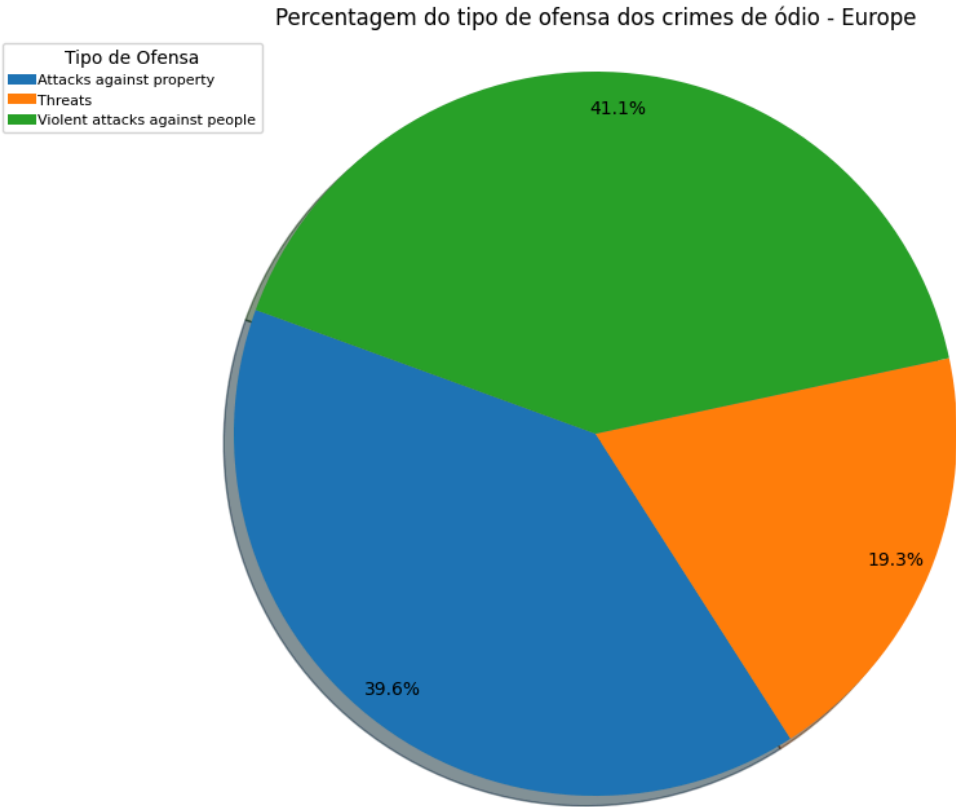
plt.pie(values, autopct='%1.1f%%', shadow=True, startangle=160, pctdistance=0.
plt.axis('equal')
plt.title('Percentagem do tipo de ofensa dos crimes de ódio - United States of Am
plt.legend(labels,
            title='Tipo de Ofensa',
            fontsize=8,
            handletextpad=0.1,

            loc='upper left')

plt.show()
eurosoma = sum(values)
print("Percentagem do tipo de ofensa dos crimes de ódio - United States of Am
for i in range(len(values)):
    print(labels[i], round((values[i]/eurosoma)*100,2), "% ")

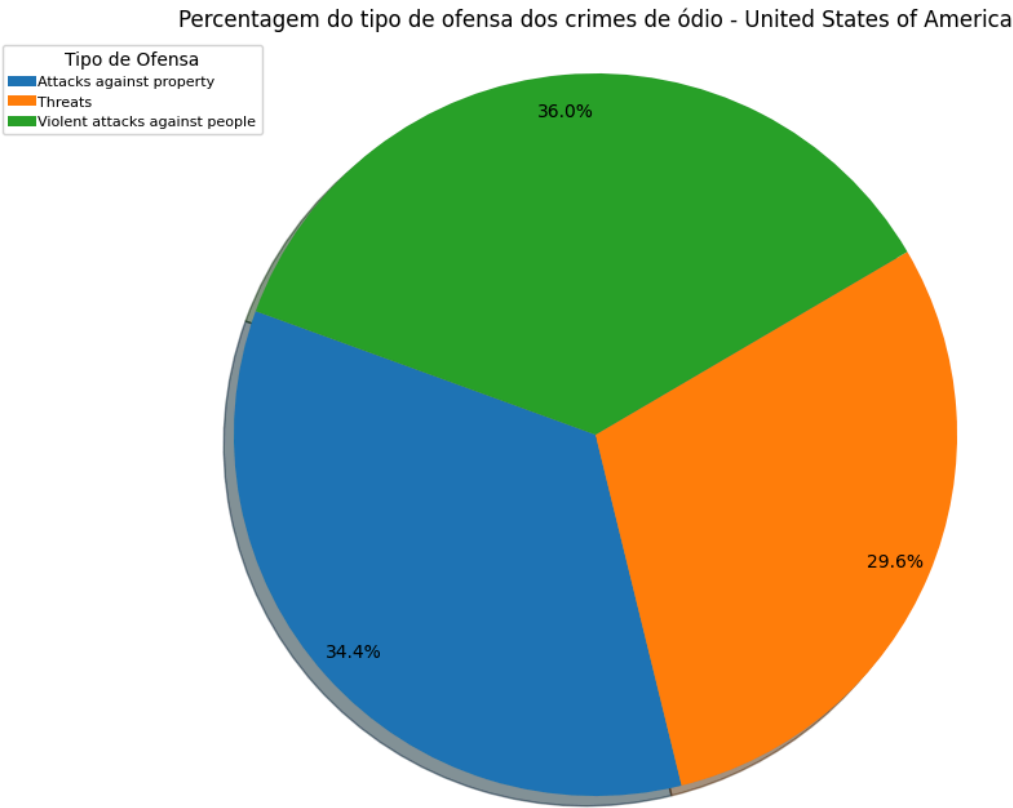
[('Attacks against property', 7874, 'Europe'), ('Threats', 3833, 'Europe'),
('Violent attacks against people', 8167, 'Europe')]

```



Percentagem do tipo de ofensa dos crimes de ódio – Europa:

Attacks against property 39.62 %  
Threats 19.29 %  
Violent attacks against people 41.09 %



Percentagem do tipo de ofensa dos crimes de ódio – United States of America:

Attacks against property 34.38 %  
Threats 29.62 %  
Violent attacks against people 36.0 %

Comparando os dados de ambas as regiões, observamos algumas diferenças notáveis:

**Attacks against property** : A Europa tem uma ligeiramente maior incidência de vandalismo (39.62%) em comparação com os Estados Unidos (34.38%). Isso pode refletir diferenças nas formas de manifestação de ódio e nos tipos de bens que são alvo de tais crimes.

**Threats**: Nos Estados Unidos, as ameaças representam uma parcela maior dos crimes de ódio (29.62%) em comparação com a Europa (19.29%). Este fato pode indicar que, nos EUA, existe uma maior tendência para a intimidação verbal e psicológica, que pode preceder ou substituir a violência física.

**Violent attacks against people**: A Europa apresenta uma maior incidência de ataques físicos (41.09%) em comparação com os Estados Unidos (36.0%). Este dado pode sugerir que, na Europa, os crimes de ódio têm maior probabilidade de resultar em violência física direta.

A análise revela que, embora haja semelhanças na distribuição dos tipos de ofensas nos crimes de ódio entre a Europa e os Estados Unidos, existem também diferenças significativas. A Europa tende a ter uma maior proporção de ataques violentos contra pessoas e ataques contra a propriedade, enquanto os Estados Unidos apresentam uma maior incidência de ameaças. Essas variações podem ser atribuídas a diferentes contextos culturais, sociais e legais que influenciam como o ódio se manifesta em cada região.

## Qual a evolução dos tipos de ofensas ao longo dos anos (EUA vs Europe)?

In [ ]:

```
query = """
select offense_name,count(*),year from crimes where year >= 2016
group by offense_name,country,year
order by country,offense_name,year;
"""

import numpy as np
cur = conn.cursor()
cur.execute(query)

rows = cur.fetchall()

europe = rows[:21]

help = []
listafinal = []

count = 0
for eu in europe:
    help.append(eu[1])
    count += 1
    if(count == 7):
        labels.append(eu[0])
        listafinal.append(help)
        help = []

        count = 0

print(europe)

plt.figure(figsize=(10,6))
```

```
plt.title('Evolução do tipo de ofensa dos crimes de ódio - Europe')
barWidth=0.25
r1 = np.arange(len(listafinal[0]))
r2 = [x + barWidth for x in r1]
r3 = [x + barWidth for x in r2]

plt.bar(r1,listafinal[0],width=barWidth,label='Attacks against property')
plt.bar(r2,listafinal[1],width=barWidth,label='Threats')
plt.bar(r3,listafinal[2],width=barWidth,label='Violent attacks against people')

plt.xticks([r + barWidth for r in range(len(listafinal[0]))],[2016,2017,2018,
plt.legend()
plt.show()
```

```
eua = rows[21:]
```

```
help = []
listafinal = []
print(eua)
```

```
count = 0
for eu in eua:
    help.append(eu[1])
    count += 1
    if(count == 7):
        labels.append(eu[0])
        listafinal.append(help)
        help = []

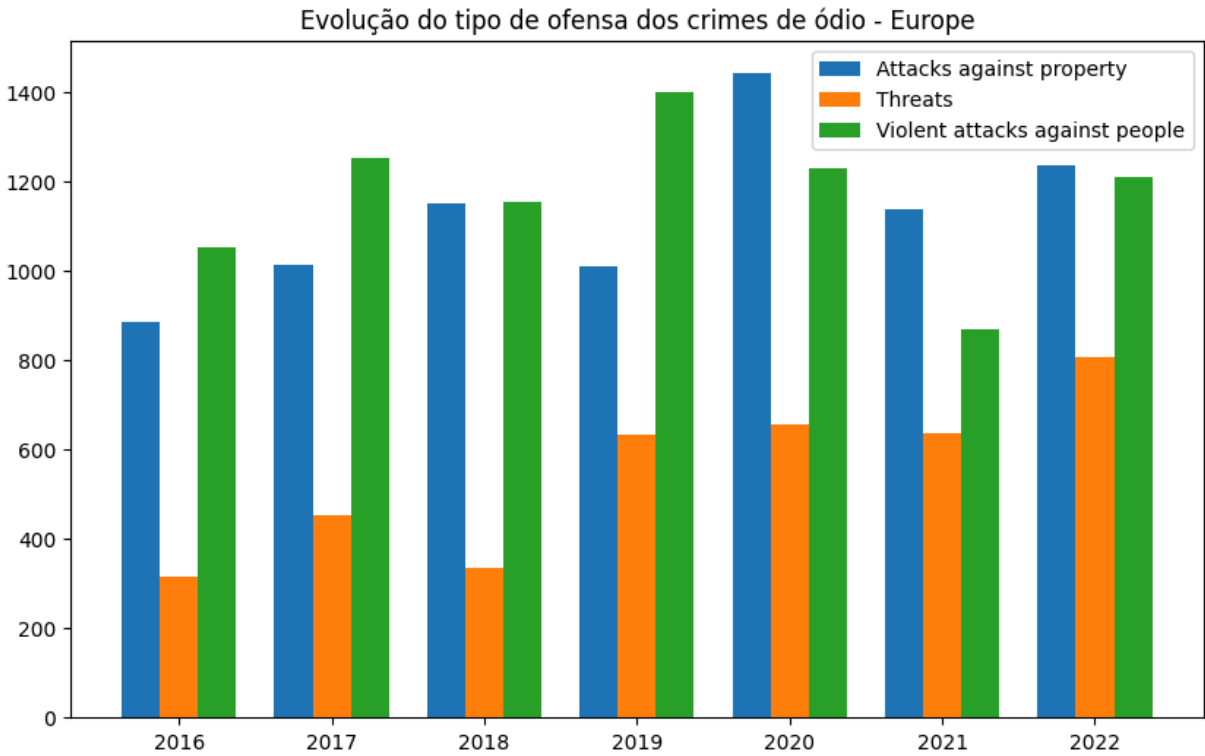
    count = 0
```

```
plt.figure(figsize=(10,6))
plt.title('Evolução do tipo de ofensa dos crimes de ódio - United States of A
barWidth=0.25
r1 = np.arange(len(listafinal[0]))
r2 = [x + barWidth for x in r1]
r3 = [x + barWidth for x in r2]

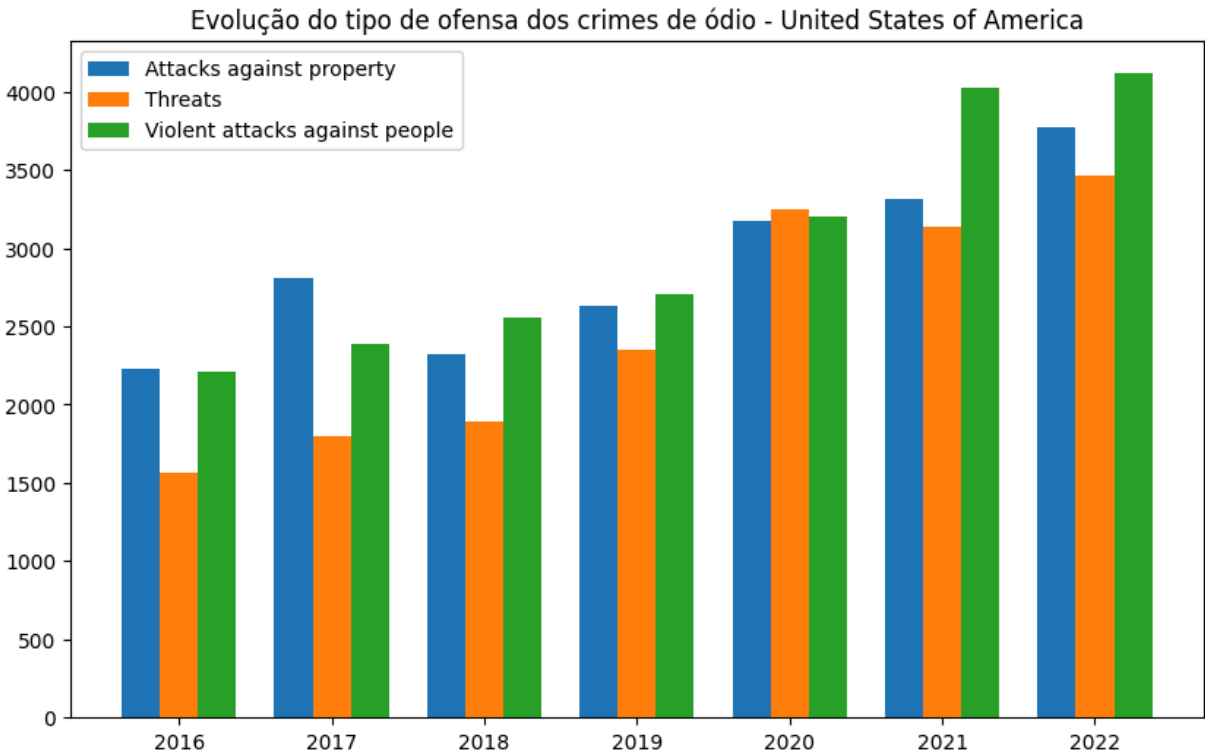
plt.bar(r1,listafinal[0],width=barWidth,label='Attacks against property')
plt.bar(r2,listafinal[1],width=barWidth,label='Threats')
plt.bar(r3,listafinal[2],width=barWidth,label='Violent attacks against people')

plt.xticks([r + barWidth for r in range(len(listafinal[0]))],[2016,2017,2018,
plt.legend()
plt.show()
```

```
[('Attacks against property', 885, 2016), ('Attacks against property', 1012,
2017), ('Attacks against property', 1150, 2018), ('Attacks against property',
1009, 2019), ('Attacks against property', 1442, 2020), ('Attacks against prop
erty', 1138, 2021), ('Attacks against property', 1238, 2022), ('Threats', 31
6, 2016), ('Threats', 453, 2017), ('Threats', 334, 2018), ('Threats', 632, 20
19), ('Threats', 655, 2020), ('Threats', 636, 2021), ('Threats', 807, 2022),
('Violent attacks against people', 1052, 2016), ('Violent attacks against peo
ple', 1253, 2017), ('Violent attacks against people', 1154, 2018), ('Violent
attacks against people', 1400, 2019), ('Violent attacks against people', 123
0, 2020), ('Violent attacks against people', 868, 2021), ('Violent attacks ag
ainst people', 1210, 2022)]
```



```
[('Attacks against property', 2226, 2016), ('Attacks against property', 2811, 2017), ('Attacks against property', 2320, 2018), ('Attacks against property', 2631, 2019), ('Attacks against property', 3173, 2020), ('Attacks against property', 3316, 2021), ('Attacks against property', 3772, 2022), ('Threats', 1564, 2016), ('Threats', 1803, 2017), ('Threats', 1888, 2018), ('Threats', 2347, 2019), ('Threats', 3247, 2020), ('Threats', 3133, 2021), ('Threats', 3464, 2022), ('Violent attacks against people', 2211, 2016), ('Violent attacks against people', 2386, 2017), ('Violent attacks against people', 2561, 2018), ('Violent attacks against people', 2703, 2019), ('Violent attacks against people', 3198, 2020), ('Violent attacks against people', 4025, 2021), ('Violent attacks against people', 4116, 2022)]
```



A Europa mostra uma tendência ascendente geral nos ataques contra a propriedade, com um aumento notável em 2020, possivelmente influenciado pela tensão social e económica resultante da pandemia de COVID-19. A queda em 2021 pode refletir medidas de contenção

ou mudanças na dinâmica social, enquanto o aumento em 2022 indica uma recuperação na prevalência desse tipo de crime.

As ameaças na Europa aumentaram substancialmente de 2016 a 2017, estabilizaram-se em 2019 e 2020, e aumentaram novamente em 2022. O aumento contínuo até 2022 pode estar relacionado ao crescente clima de tensão política e social, bem como ao aumento do uso de plataformas digitais para disseminar ameaças.

A Europa viu um aumento nos ataques violentos contra pessoas até 2019, seguido de uma queda significativa em 2020 e 2021, possivelmente devido a restrições de mobilidade e lockdowns durante a pandemia. No entanto, o aumento em 2022 pode indicar uma retomada das atividades normais e, conseqüentemente, um ressurgimento das oportunidades para tais crimes.

Nos EUA, há um claro aumento nos ataques contra a propriedade ao longo dos anos, um crescimento contínuo até 2022. Este aumento pode ser atribuído a fatores como agitação social, protestos, e aumento das tensões políticas e raciais.

As ameaças nos EUA mostram um aumento constante, com um salto significativo em 2020, que pode estar ligado ao aumento do discurso de ódio online e polarização política exacerbada durante as eleições presidenciais e a pandemia. A ligeira queda em 2021, possivelmente devido à pandemia, foi seguida por um novo aumento em 2022, indicando uma tendência persistente.

Os ataques violentos contra pessoas nos EUA aumentaram de forma consistente, com um pico acentuado em 2020 e 2021, refletindo o aumento das tensões sociais e políticas, bem como o impacto da pandemia. O crescimento contínuo em 2022 destaca uma tendência alarmante de aumento da violência.

Em resumo, os dados indicam que, enquanto ambos os continentes enfrentam desafios significativos relacionados aos crimes de ódio, os EUA apresentam uma tendência mais acentuada de aumento, particularmente em ataques violentos e ameaças, refletindo um ambiente social e político mais polarizado e tenso.

## Em quais meses do ano os crimes de ódio são mais frequentes, e há diferenças significativas entre os EUA e a Europa?

A análise temporal dos crimes de ódio é essencial para entender se há períodos específicos em que esses incidentes são mais prevalentes. Identificar meses ou anos em que o número de crimes de ódio se destaca pode revelar padrões sazonais, impactos de eventos sociais, políticos ou económicos, e auxiliar na criação de políticas e medidas preventivas mais eficazes.

In [ ]:

```
query = """select count(*),month from crimes where year >= 2016 group by mo

cur = conn.cursor()
cur.execute(query)

rows = cur.fetchall()
europa = rows[:12]
labels= []
```

```

values = []
print(europa)
for eu in europa:

    labels.append(eu[1])
    values.append(eu[0])

plt.figure(figsize=(12, 8))

plt.pie(values, autopct='%1.1f%%', shadow=True, startangle=160, pctdistance=0.
plt.axis('equal')
plt.title('Distribuição dos crimes de ódio por mês - Europe')
plt.legend(labels,
            title='Mês',
            fontsize=8,
            handletextpad=0.1,

            loc='upper left')

eua = rows[13:]
print(eua)
labels= []
values = []

for eu in eua:

    labels.append(eu[1])
    values.append(eu[0])

plt.figure(figsize=(12, 8))

plt.pie(values, autopct='%1.1f%%', shadow=True, startangle=160, pctdistance=0.
plt.axis('equal')
plt.title('Distribuição dos crimes de ódio por mês - United States of America')
plt.legend(labels,
            title='Mês',
            fontsize=8,
            handletextpad=0.1,

            loc='upper left')

```

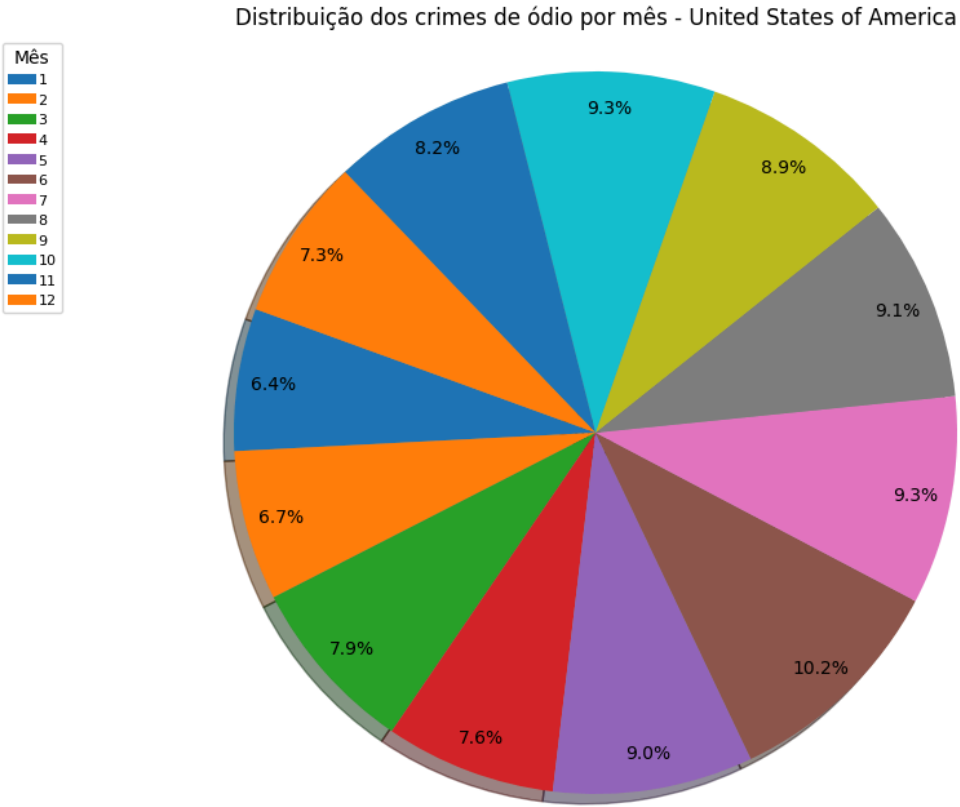
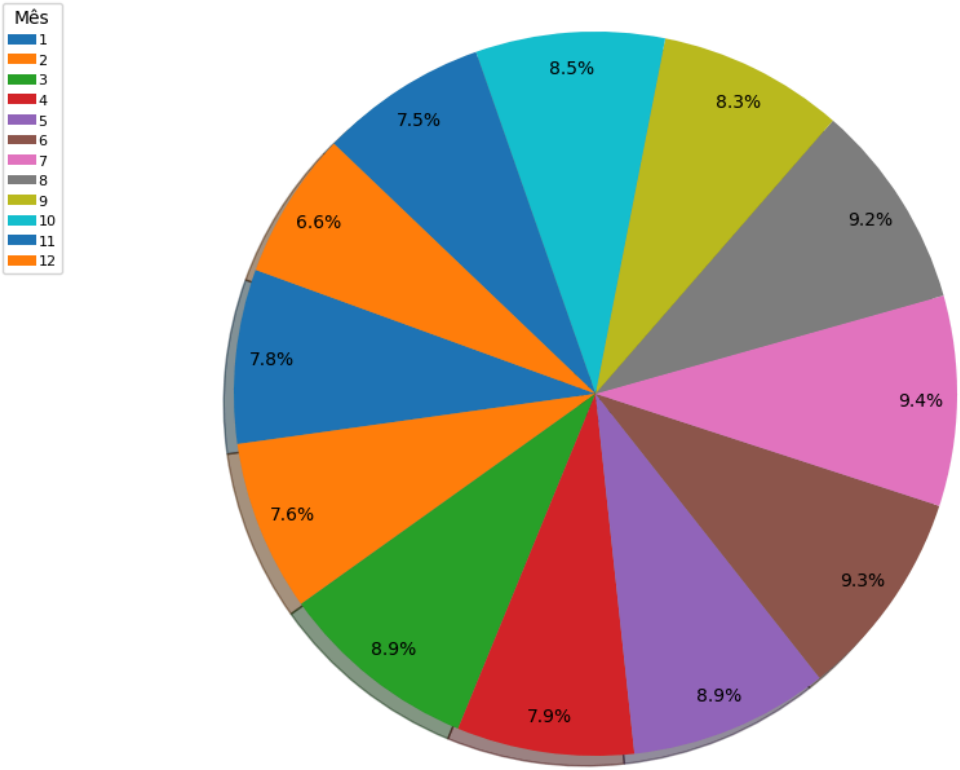
```

[(1432, 1), (1407, 2), (1645, 3), (1459, 4), (1643, 5), (1721, 6), (1729, 7),
(1699, 8), (1530, 9), (1557, 10), (1382, 11), (1211, 12)]
[(3753, 1), (3956, 2), (4677, 3), (4498, 4), (5276, 5), (6015, 6), (5499, 7),
(5356, 8), (5265, 9), (5464, 10), (4853, 11), (4283, 12)]

```

Out[ ]: <matplotlib.legend.Legend at 0x11b969b10>





Analisando os 2 gráficos, verificamos que os crimes de ódio entre os meses são bastante parecidos. Nos Estados Unidos, o número de crimes de ódio em junho e julho é significativamente maior do que na Europa. Junho, com 6015 incidentes, é o mês mais problemático nos EUA, seguido de julho com 5499 incidentes.

Na Europa, os números são menores em comparação com os EUA, com julho sendo o pico com 1729 incidentes, seguido de junho com 1721.

A sazonalidade dos crimes de ódio parece ser uma característica comum em ambas as regiões, com picos durante os meses de verão. Esse padrão pode estar relacionado a uma série de fatores, incluindo maior interação social, aumento de atividades ao ar livre e, possivelmente, eventos culturais ou políticos que podem intensificar tensões.

Nos EUA, os picos nos meses de verão podem estar relacionados a eventos específicos como o mês do orgulho LGBTQ+ em junho, que pode levar a um aumento tanto de visibilidade quanto de incidentes de ódio.

Na Europa, os fatores podem ser variados e mais difusos, incluindo eventos políticos ou sociais específicos que não são facilmente atribuíveis a um único fator como nos EUA.

Visto que queremos abordar mais esta questão, e visto que o mês do orgulho LGBTQ+ é realizado em junho, vamos analisar se o número de crimes relacionados a este **Anti-LGBTI hate crime**.

## Qual o mês com maior número de crimes com o crime Anti - LGBTI hate crime?

Tal como referido anteriormente, iremos analisar qual o mês com maior nº de crimes para esta sociedade.

```
In [ ]: query = """select count(*) as v,month from crimes where year >= 2016 and bia
import numpy as np
cur = conn.cursor()
cur.execute(query)

rows = cur.fetchall()
europa = rows[:12]
labels= []
values = []

for eu in europa:
    values.append(eu[0])
    labels.append(eu[1])

l = [229, 245, 271, 276, 287, 310, 311, 367, 379, 399, 402, 462]

max = np.max(values)
pos = np.argmax(values)

print("Europa: O mês com valor mais alto é: ",labels[pos])

eua = rows[14:]
labels= []
values = []

for eu in eua:
    values.append(eu[0])
    labels.append(eu[1])
max = np.max(values)
pos = np.argmax(values)
print("United States of America: O mês com valor mais alto é: ",labels[pos])
```

Europa: 0 mês com valor mais alto é: 6

United States of America: 0 mês com valor mais alto é: 6

Tanto nos Estados Unidos como na Europa o mês mais frequente deste tipo de crime é em Junho é frequentemente reconhecido como o mês do orgulho em muitas partes do mundo, celebrando a diversidade sexual e de género, promovendo a igualdade e a aceitação.

No entanto, também é um período em que vemos um aumento significativo na incidência de crimes de ódio e discriminação contra a comunidade LGBTI. Existem bastante marchas e desfiles, o que pode provocar reações negativas de grupos intolerantes.

## Como é a evolução entre meses nos tipos de crimes?

Ao respondermos à pergunta anterior, ficamos interessados em analisar em que meses ocorrem o maior número de crimes para os restantes crimes. Portanto iremos realizar a pesquisa, para verificarmos se existe algum padrão.

In [ ]:

```
query = """SELECT month AS mes,
               bias_desc,
               COUNT(*) AS total_crimes
FROM crimes where month >= 1
GROUP BY bias_desc, country, month
ORDER BY country, bias_desc, mes, total_crimes DESC;"""
cur = conn.cursor()
cur.execute(query)

rows = cur.fetchall()

listaCrimes = []
lb = []
anos = [1,2,3,4,5,6,7,8,9,10,11,12]
for i in range(0,108,12):
    lb.append(rows[i][1])
    m = []
    for j in range(12):
        m.append(rows[i+j][2])
    listaCrimes.append(m)
count = 0

plt.figure(figsize=(17, 8))
print('Europe')
for incident in listaCrimes:
    print(lb[count], incident)
    plt.plot(anos, incident, label=lb[count])
    count+=1

plt.xlabel('Mês')
plt.ylabel('Quantidade de Crimes')
plt.title('Evolução Mensal dos Tipos Crimes - Europe')
plt.legend()
plt.grid(True)

plt.show()

listaCrimes = []
lb = []
anos = [1,2,3,4,5,6,7,8,9,10,11,12]
for i in range(108, len(rows), 12):
```

```

lb.append(rows[i][1])
m = []
for j in range(12):
    m.append(rows[i+j][2])
listaCrimes.append(m)
count = 0

plt.figure(figsize=(17, 8))
print('\nUnited States of America')
for incident in listaCrimes:
    print(lb[count], incident)
    plt.plot(anos, incident, label=lb[count])
    count+=1

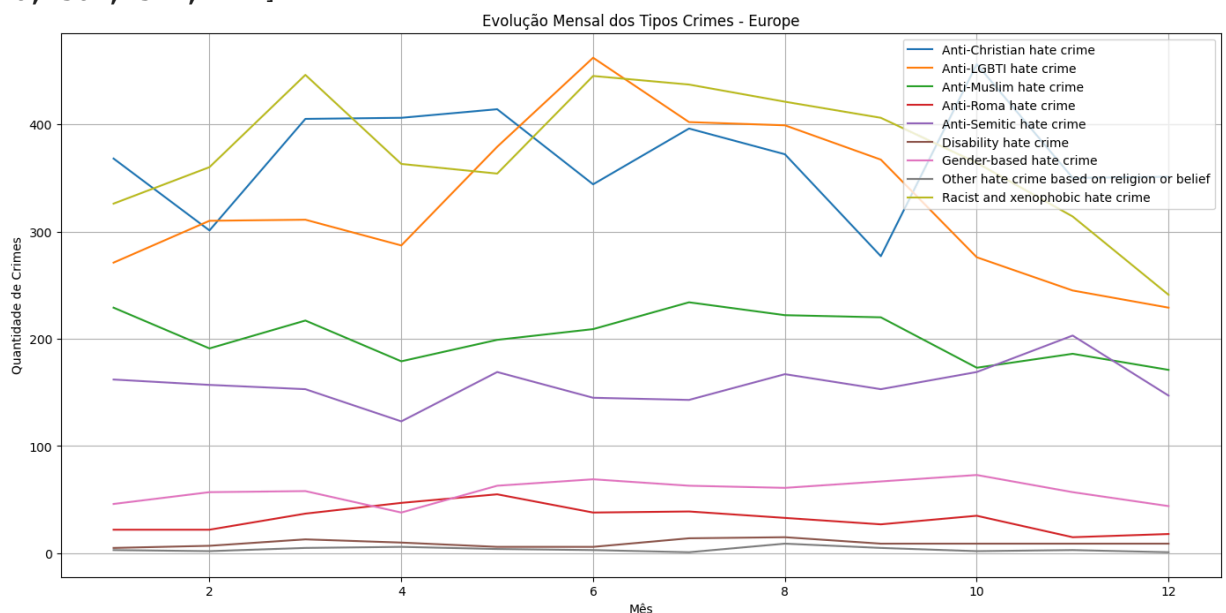
plt.xlabel('Mês')
plt.ylabel('Quantidade de Crimes')
plt.title('Evolução Mensal dos Tipos Crimes - United States of America')
plt.legend()
plt.grid(True)

plt.show()

```

### Europe

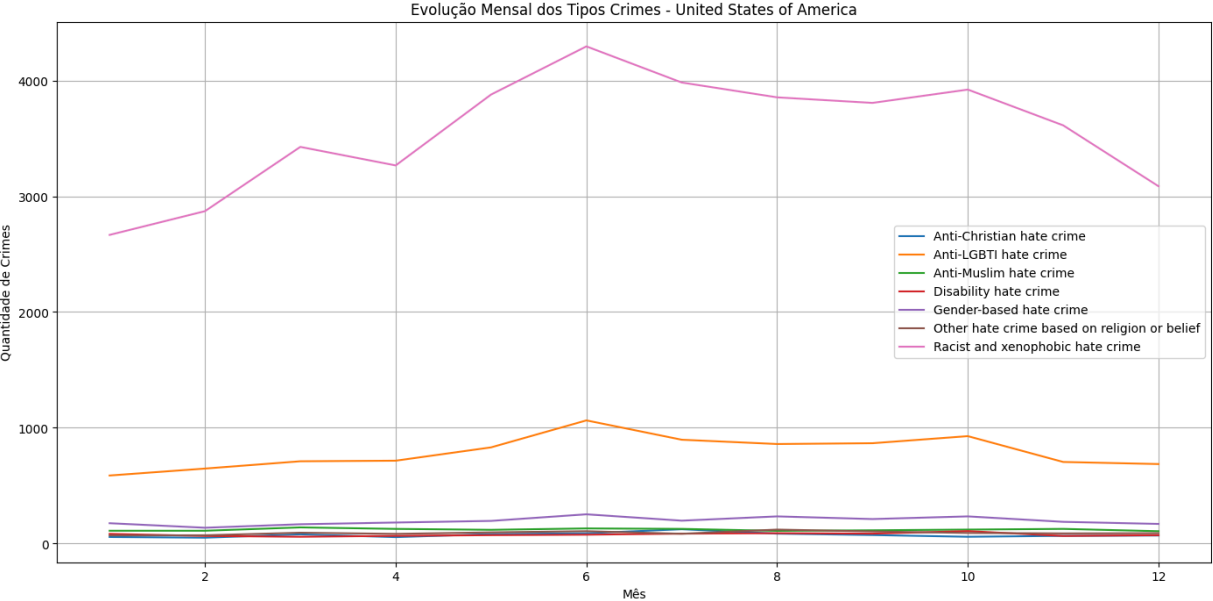
Anti-Christian hate crime [368, 301, 405, 406, 414, 344, 396, 372, 277, 456, 350, 351]  
 Anti-LGBTI hate crime [271, 310, 311, 287, 379, 462, 402, 399, 367, 276, 245, 229]  
 Anti-Muslim hate crime [229, 191, 217, 179, 199, 209, 234, 222, 220, 173, 186, 171]  
 Anti-Roma hate crime [22, 22, 37, 47, 55, 38, 39, 33, 27, 35, 15, 18]  
 Anti-Semitic hate crime [162, 157, 153, 123, 169, 145, 143, 167, 153, 169, 203, 147]  
 Disability hate crime [5, 7, 13, 10, 6, 6, 14, 15, 9, 9, 9, 9]  
 Gender-based hate crime [46, 57, 58, 38, 63, 69, 63, 61, 67, 73, 57, 44]  
 Other hate crime based on religion or belief [3, 2, 5, 6, 4, 3, 1, 9, 5, 2, 3, 1]  
 Racist and xenophobic hate crime [326, 360, 446, 363, 354, 445, 437, 421, 406, 364, 314, 241]



### United States of America

Anti-Christian hate crime [57, 51, 80, 56, 82, 89, 124, 86, 74, 59, 67, 70]  
 Anti-LGBTI hate crime [588, 648, 711, 716, 831, 1064, 897, 860, 867, 928, 705, 687]  
 Anti-Muslim hate crime [111, 111, 140, 127, 119, 131, 127, 110, 115, 121, 127, 107]  
 Disability hate crime [84, 66, 59, 67, 73, 77, 86, 90, 86, 108, 66, 73]  
 Gender-based hate crime [176, 137, 167, 182, 196, 253, 198, 235, 212, 235, 18]

8, 170]  
Other hate crime based on religion or belief [71, 72, 94, 84, 97, 107, 85, 121, 105, 92, 88, 90]  
Racist and xenophobic hate crime [2666, 2871, 3426, 3266, 3878, 4294, 3982, 3854, 3806, 3921, 3612, 3086]



A análise dos dados mensais de crimes de ódio nos Estados Unidos e na Europa revela diferenças e semelhanças significativas nos padrões de incidência.

Começando pelos Estados Unidos, os crimes de ódio anti-cristãos atingem o pico no verão (visualizando a tabela), especificamente em junho. Isso pode ser influenciado por eventos ao ar livre, celebrações que aumentam durante esta época do ano.

Tal como dito anteriormente a comunidade LGBTI tem um aumento em junho.

A maior incidência em dos crimes anti-muslim são realizados em Março, sendo que nesse mês é o Ramadão, um dos pontos mais importantes nesta religião, por isso estão mais expostos.

Para o crime de Disability, o pico é em Outubro, sendo que nos estados unidos, nesse mesmo mês ocorre, o evento "Mês Nacional de Emprego de Pessoas com Deficiências nos EUA", sendo que ocorre um aumento na visibilidade de pessoas com deficiências, possivelmente levando a um aumento nos crimes de ódio.

A correlação com o mês de orgulho LGBTI pode influenciar os crimes baseados em género. O aumento em agosto dos crimes sobre outras religiões, pode estar associados a eventos que aumentam a visibilidade de certos grupos religiosos.

O Racismo tem um pico no verão, provavelmente por conta das viagens das pessoas que vão para outras comunidades.

Na Europa, o aumento em outubro pode estar relacionado a eventos religiosos ou culturais específicos que elevam a visibilidade da comunidade cristã.

Similar aos EUA, junho é o Mês do Orgulho LGBTQ+ na Europa, com um aumento significativo nos crimes de ódio devido à maior visibilidade da comunidade.

A comunidade Islã também tem um valor alto no mês de março (ramadão).

Os anti semitic, o aumento em novembro pode estar associado à lembrança da Kristallnacht, um evento histórico que aumenta a visibilidade da comunidade judaica.

Para as outras religiões, semelhante aos EUA, pode haver um aumento devido a eventos religiosos específicos em agosto.

Concluindo, a análise dos meses com maior incidência de crimes de ódio nos EUA e na Europa revela tanto similaridades quanto diferenças significativas. Eventos culturais e históricos desempenham um papel crucial na dinâmica desses crimes, com a visibilidade dos grupos minoritários sendo um fator comum que aumenta a probabilidade de incidentes. As diferenças sazonais e a diversidade de influências refletem as complexidades sociais e culturais distintas de cada continente.

## Qual o top5 de estados com maior e menor taxa de criminalidade por população? (EUA vs Europe)

A seguir, será apresentada uma comparação entre os cinco estados americanos e os cinco países europeus com as maiores e menores taxas de criminalidade per capita, destacando as principais características que contribuem para esses rankings.

Europe:

In [ ]:

```
query = """select count(distinct(incident_id)),Cr.state_name,Cr.year,da.pop20
join datastates_country as da
on da.state_name = Cr.state_name where da.totalarea > 0 group by country,Cr.
"""

cur = conn.cursor()
cur.execute(query)

rowsTop = cur.fetchall()
listadados = []
c = True
num = 2016
CountryInci = []
for i in range(308):
    hel = []

    if num <= rowsTop[i][2]:
        num+=1
        CountryInci.append(rowsTop[i][0])
    else:
        num = 2016
        média = np.mean(CountryInci)
        médiaPop = np.mean([rowsTop[i-1][4:10]])

        CountryInci = []

        dens = médiaPop/int(rowsTop[i-1][-1])
        info = [(média/dens), rowsTop[i-1][1]]
        listadados.append(info)

lista_ordenada = sorted(listadados, key=lambda x: x[0])
menor = lista_ordenada[:5]
```

```

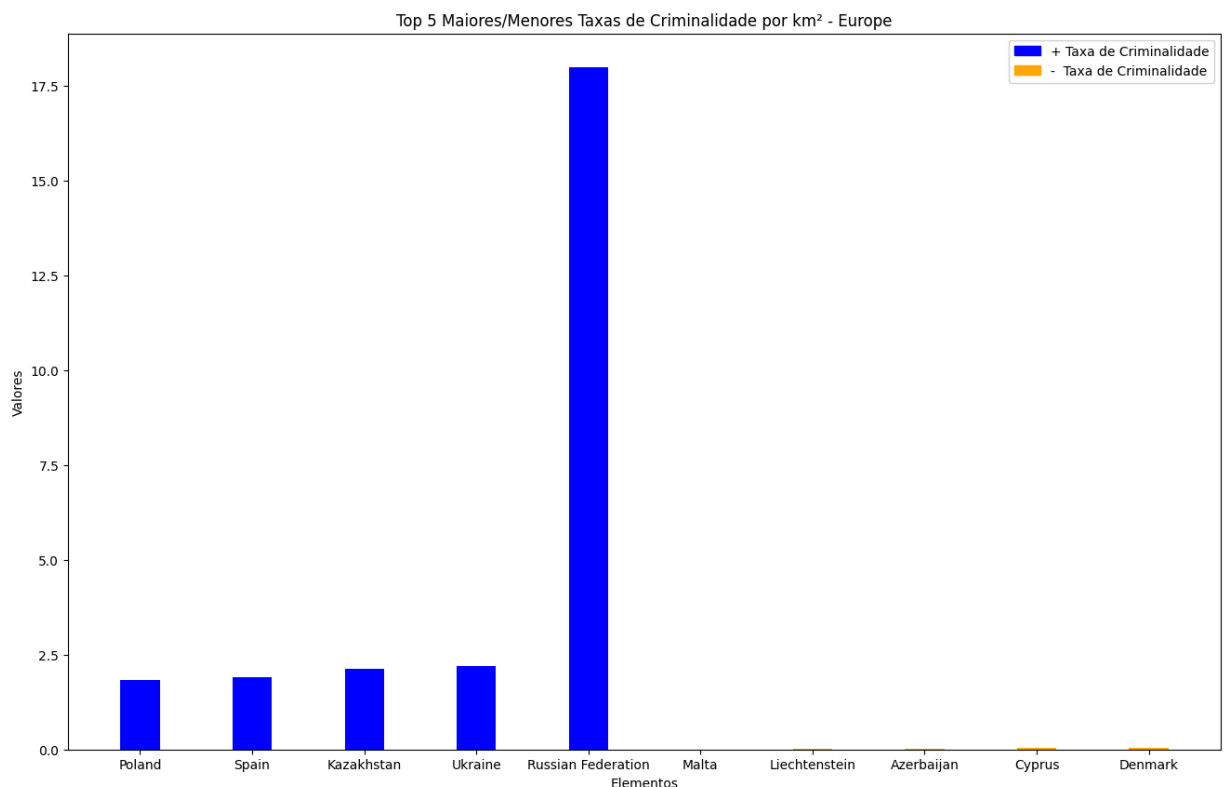
lista_ordenada = sorted(listados, key=lambda x: x[0])
maior = lista_ordenada[-5:]

lab = [x[1] for x in maior] + [x[1] for x in menor]
val = [x[0] for x in maior] + [x[0] for x in menor]

plt.figure(figsize=(16, 10))
largura = 0.35
plt.bar(np.arange(len(val)), val, width=largura, color=['blue'] * len(menor))
plt.xlabel('Elementos')
plt.ylabel('Valores')
plt.title('Top 5 Maiores/Menores Taxas de Criminalidade por km² - Europe')
plt.xticks(np.arange(len(val)), lab)
azul_patch = mpatches.Patch(color='blue', label='+ Taxa de Criminalidade')
laranja_patch = mpatches.Patch(color='orange', label='- Taxa de Criminalidade')
plt.legend(handles=[azul_patch, laranja_patch])
plt.show()
print("Países com maior Taxa de Criminalidade por km²\n")
print("País\tTaxa:")
for i in range(5):
    print(lab[i], "\t", val[i], " Crimes por km²")

print("\nPaíses com menor Taxa de Criminalidade por km²\n")
print("País\tTaxa:")
for i in range(5, len(lab)):
    print(lab[i], "\t", val[i], " Crimes por km²")

```



Países com maior Taxa de Criminalidade por km²

País	Taxa:
Poland	1.838901379668121 Crimes por km²
Spain	1.9030008716514937 Crimes por km²
Kazakhstan	2.1172885958764236 Crimes por km²
Ukraine	2.1900776569028095 Crimes por km²
Russian Federation	17.969094164687746 Crimes por km²

Países com menor Taxa de Criminalidade por km²

País	Taxa:	
Malta	0.001882401531337617	Crimes por km <sup>2</sup>
Liechtenstein	0.006215630584355603	Crimes por km <sup>2</sup>
Azerbaijan	0.023068674679619185	Crimes por km <sup>2</sup>
Cyprus	0.035012300723076005	Crimes por km <sup>2</sup>
Denmark	0.03546483472872204	Crimes por km <sup>2</sup>

Esta análise sugere que a Rússia tem a taxa de criminalidade muito alta em comparação com outros países, enquanto Malta possui uma das menores taxas de criminalidade por km<sup>2</sup>. No entanto, é importante considerar dados demográficos, tais como população, idade da população, nível de riqueza e percentagem de imigração.

De qualquer forma a Rússia tem uma taxa de criminalidade elevada tem um valor tão desnivelado em comparação ao resto do top 5 de maiores taxas. Uma justificação poderá ser que a Rússia possui características conservadoras em sua cultura e política. O governo, sob a liderança de Vladimir Putin, defende valores tradicionais, como a importância da família, da religião e do patriotismo. Há também restrições a algumas liberdades individuais, como a liberdade de expressão e os direitos LGBTQ+.

Comparando os países com maior taxa com os de menor taxa, a Ucrânia tem uma taxa 121666% superior em comparação a Malta, iremos agora analisar estes 10 países para tentarmos encontrar um padrão.

In [ ]:

```
q = "state_name = '" + lab[0]
for i in range(1,5):
    q += "' or state_name = '" + lab[i]

q += ""
query = ""select * from datastates_country where "" + q
cur = conn.cursor()
cur.execute(query)

infoCountry = cur.fetchall()
print("Países com maior Taxa:\n")
for c in infoCountry:
    p = np.mean([c[3:8]])
    gd = np.mean([c[10:16]])
    ag = np.mean(c[17:22])
    im = np.mean(c[-7:])
    print("País: ",c[1]," População: ",p," Área: ",c[9]," GDP ",gd," Age: ",a

q = "state_name = '" + lab[5]
for i in range(6,len(lab)):
    q += "' or state_name = '" + lab[i]

q += ""
query = ""select * from datastates_country where "" + q
cur = conn.cursor()
cur.execute(query)

infoCountry = cur.fetchall()
print("\n\nPaíses com menor Taxa:\n")
for c in infoCountry:
    p = np.mean([c[3:8]])
    gd = np.mean([c[10:16]])
    ag = np.mean(c[17:22])
```



```
im = np.mean(c[-7:])
print("País: ",c[1]," População: ",p," Área: ",c[9]," GDP ",gd," Age: ",a
```

Países com maior Taxa:

País: Kazakhstan População: 18516911.0 Área: 2699700.0 GDP 9347.178280  
289543333333333333 Age: 30.56 Imigração: 0.022999999999999996

País: Poland População: 37912249.0 Área: 306100.0 GDP 15210.9890122589  
16666666666667 Age: 41.04 Imigração: 0.03842857142857143

País: Russian Federation População: 144316896.0 Área: 16376870.0 GDP 1  
0829.2220052083333333333333333333 Age: 39.92 Imigração: 0.021999999999999995

País: Spain População: 47061455.2 Área: 499733.2 GDP 28692.80629175513  
333333333333333333 Age: 44.18 Imigração: 0.13

País: Ukraine População: 44358961.2 Área: 579400.0 GDP 3360.6093750000  
0166666666666667 Age: 40.80 Imigração: 0.043999999999999999

Países com menor Taxa:

País: Azerbaijan População: 10009791.6 Área: 82650.0 GDP 4535.23123346  
6871666666666667 Age: 31.32 Imigração: 0.021999999999999995  
País: Cyprus População: 1227583.0 Área: 9240.0 GDP 28499.3935546875 A  
ge: 37.20 Imigração: 0.2232857142857143  
País: Denmark População: 5812235.0 Área: 40000.0 GDP 60594.04172017523  
333333333333 Age: 41.84 Imigração: 0.09457142857142856  
País: Liechtenstein População: 38469.4 Área: 160.0 GDP 173676.54205122  
55 Age: 43.34 Imigração: 0.338  
País: Malta População: 498111.8 Área: 316 GDP 30404.8243853253 Age:  
41.82 Imigração: 0.19028571428571425

Analisando, os países com maiores taxas de criminalidade tendem a ter populações maiores e áreas geográficas extensas, o que pode criar desafios adicionais para aplicação da lei e controle da criminalidade.

No entanto, é interessante nota que a Espanha, com uma alta taxa de imigração e um PIB per capita relativamente alto, tem uma taxa de criminalidade comparativamente alta.

Por outro lado, os países com menores taxas de criminalidade tendem a ter populações menores e áreas geográficas menores, possivelmente facilitando o controle da criminalidade.

Liechtenstein destaca-se com o maior PIB per capita entre os países listados, o que pode contribuir para a qualidade de vida e, conseqüentemente, para uma menor incidência de crimes.

Podemos visualizar que nos países com menor taxa de crimes têm uma taxa de imigrantes relativamente alta, mas provavelmente é menor visto que são países mais pequenos, sendo que visto que grande parte são imigrantes estes não têm tanto preconceito uns com os outros, visto que todos os imigrantes conseguem se compreender uns aos outros.

## EUA

```
In [ ]: listadados = []
        for i in range(308, len(rowsTop)):
            hel = []

            if num <= rowsTop[i][2]:
                num+=1
                CountryInci.append(rowsTop[i][0])
            else:
                num = 2016
```

```

média = np.mean(CountryInci)
médiaPop = np.mean([rowsTop[i-1][4:10]])

CountryInci = []

dens = médiaPop/int(rowsTop[i-1][-1])
info = [((média/dens)), rowsTop[i-1][1]]
listados.append(info)

lista_ordenada = sorted(listados, key=lambda x: x[0])
menor = lista_ordenada[:5]

lista_ordenada = sorted(listados, key=lambda x: x[0])
maior = lista_ordenada[-5:]
# Pegar os rótulos e valores separadamente

lab = [x[1] for x in maior] + [x[1] for x in menor]
val = [x[0] for x in maior] + [x[0] for x in menor]
l = ['Menor', 'Maior']

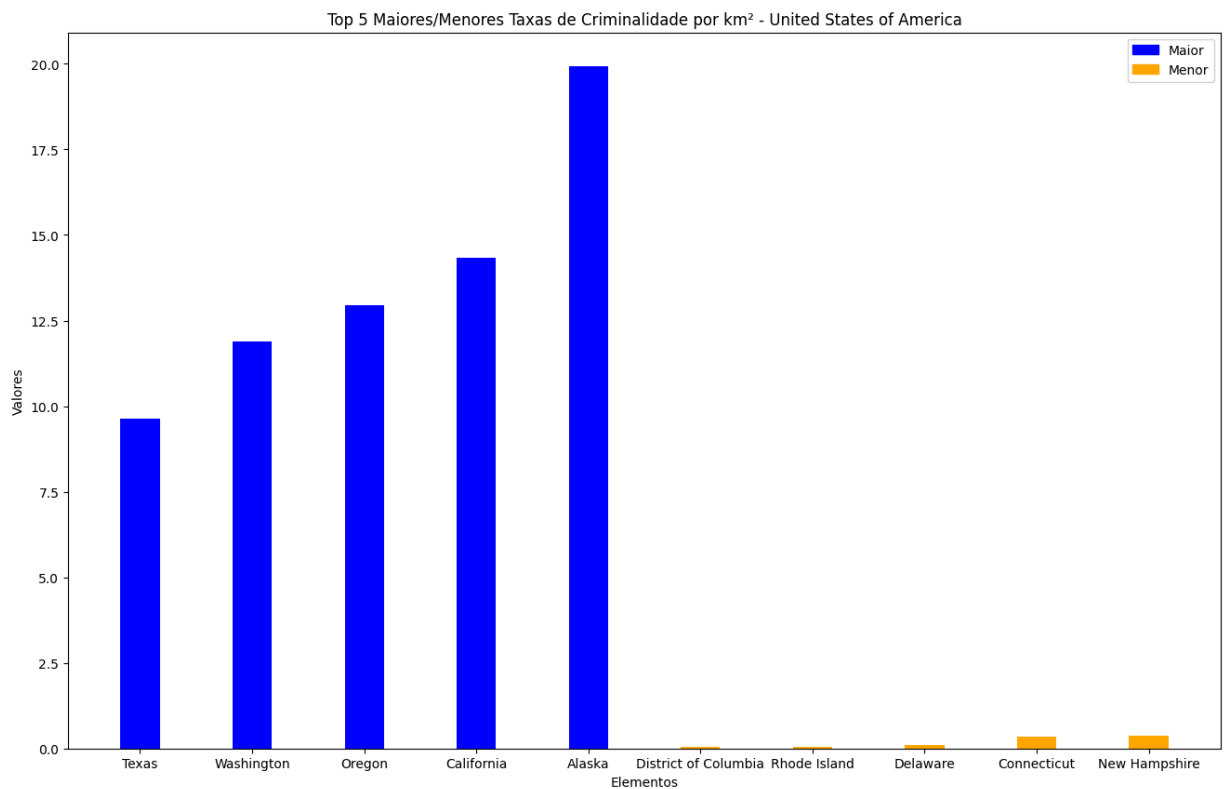
# Criar o gráfico de barras
plt.figure(figsize=(16, 10))
largura = 0.35
plt.bar(np.arange(len(val)), val, width=largura, color=['blue'] * len(menor))
plt.xlabel('Elementos')
plt.ylabel('Valores')
plt.title('Top 5 Maiores/Menores Taxas de Criminalidade por km² - United States')
plt.xticks(np.arange(len(val)), lab)

azul_patch = mpatches.Patch(color='blue', label='Maior')
laranja_patch = mpatches.Patch(color='orange', label='Menor')
plt.legend(handles=[azul_patch, laranja_patch])
# Mostrar o gráfico
plt.show()

print("Estados com maior Taxa de Criminalidade por km²\n")
print("País\tTaxa:")
for i in range(5):
    print(lab[i], "\t", val[i], " Crimes por km²")

print("\nEstados com menor Taxa de Criminalidade por km²\n")
print("País\tTaxa:")
for i in range(5, len(lab)):
    print(lab[i], "\t", val[i], " Crimes por km²")

```



Estados com maior Taxa de Criminalidade por km<sup>2</sup>

País Taxa:

Texas	9.638975365104447	Crimes por km <sup>2</sup>
Washington	11.9011988281605	Crimes por km <sup>2</sup>
Oregon	12.957673167446991	Crimes por km <sup>2</sup>
California	14.33119066432975	Crimes por km <sup>2</sup>
Alaska	19.91902928424229	Crimes por km <sup>2</sup>

Estados com menor Taxa de Criminalidade por km<sup>2</sup>

País Taxa:

District of Columbia	0.04016572563577833	Crimes por km <sup>2</sup>
Rhode Island	0.05449074105600186	Crimes por km <sup>2</sup>
Delaware	0.108925669467878	Crimes por km <sup>2</sup>
Connecticut	0.34714396846631324	Crimes por km <sup>2</sup>
New Hampshire	0.3824142091281466	Crimes por km <sup>2</sup>

Em resumo, os dados apresentados revelam uma ampla disparidade nas taxas de criminalidade por km<sup>2</sup> entre os estados dos Estados Unidos.

Estados como Alaska e Califórnia exibem taxas de criminalidade significativamente mais altas, enquanto o Distrito de Columbia e Rhode Island apresentam taxas excepcionalmente baixas.

Essa variação pode ser atribuída a uma combinação de fatores geográficos, socioeconômicos e demográficos. No entanto, uma análise mais aprofundada vai ser realizada para entender completamente as razões por trás dessas diferenças.

In [ ]:

```
q = "state_name = '" + lab[0]
for i in range(1,5):
    q += "' or state_name = '" + lab[i]

q += "'"

query = """select * from datastates_country where """ + q
cur = conn.cursor()
cur.execute(query)
```

Estado com maior Taxa:

Estado com menor Taxa:

A imigração é relativamente semelhante em estados com taxas de criminalidade mais altas e mais baixas.

A idade podemos verificar que quanto mais jovem é a população maior é a taxa de criminalidade em relação a populações mais velhas e mais conservadoras.

Em termos de área, verificamos que quanto maior é a área total do estado maior é a taxa de criminalidade, o que corresponde à realidade, visto que existe uma maior área a vigiar o que implica que as pessoas façam crimes de ódio, pois é complicado profissões responsáveis por garantir segurança ao estado, vigiarem toda a área.

## Conclusão:

Tanto na Europa como nos Estados unidos, as maiores taxas de criminalidade tendem a ser nas populações maiores e áreas geográficas extensas, que pode criar desafios para o controlo da criminalidade. Na Europa países mais ricos tendem a ter taxas de criminalidade mais baixas. Na Europa este top10 quando comparando a população de imigrantes.

Para avaliar o verdadeiro impacto destes indicativos iremos analisar a correlação nas próximas perguntas.

## Qual a correlação entre a população de imigrantes (populationimi) e o número de crimes de ódio (EUA vs Europa)?

A correlação entre a população de imigrantes e o número de crimes de ódio é um tema complexo e frequentemente debatido, especialmente nos Estados Unidos e na Europa, onde questões relacionadas à imigração e diversidade cultural têm desempenhado papéis significativos na política e na sociedade.

Nos Estados Unidos, a questão da imigração tem sido central em debates políticos, com posições polarizadas sobre seus impactos sociais, económicos e culturais. Alguns argumentam que uma população de imigrantes maior está associada a um aumento nos crimes de ódio, enquanto outros contestam essa afirmação, sugerindo que a imigração contribui para a diversidade cultural e enriquecimento social, sem necessariamente levar a um aumento nos crimes de ódio.

Na Europa, especialmente em países onde houve um aumento significativo na migração, como resultado de conflitos no Oriente Médio e na África, a relação entre imigração e crimes de ódio também é objeto de escrutínio. Alguns observam um aumento nos incidentes de intolerância e xenofobia em áreas com uma concentração maior de imigrantes, enquanto outros destacam a importância da integração e da educação cultural na prevenção desses crimes.

```
In [ ]: query = """WITH crime_counts AS (  
    SELECT  
        state_name,  
        year,  
        COUNT(*) AS total_crimes  
    FROM  
        crimes  
    GROUP BY  
        state_name, year  
),  
immigrant_populations AS (  
    SELECT  
        state_name,
```

```

        UNNEST(ARRAY[2016, 2017, 2018, 2019, 2020, 2021, 2022]) AS year,
        UNNEST(ARRAY[populationimi2016, populationimi2017, populationimi2018,
        populationimi2019, populationimi2020, populationimi2021, populationimi2022]) AS populationimi
    FROM
        datastates_country where euavseurope = 'United States of America'
),
combined_data AS (
    SELECT
        ic.state_name,
        ic.year,
        ic.total_crimes,
        ip.population_immigrants
    FROM
        crime_counts ic
    JOIN
        immigrant_populations ip ON ic.state_name = ip.state_name AND ic.year = ip.year
),
calculations AS (
    SELECT
        year,
        COUNT(*) AS n,
        SUM(total_crimes) AS sum_y,
        SUM(population_immigrants) AS sum_x,
        SUM(total_crimes * population_immigrants) AS sum_xy,
        SUM(total_crimes * total_crimes) AS sum_y2,
        SUM(population_immigrants * population_immigrants) AS sum_x2
    FROM
        combined_data
    GROUP BY
        year
)
SELECT
    year,
    (n * sum_xy - sum_x * sum_y) /
    SQRT((n * sum_x2 - sum_x * sum_x) * (n * sum_y2 - sum_y * sum_y)) AS correlation
FROM
    calculations
ORDER BY
    year;

```

```

cur = conn.cursor()
cur.execute(query)

rows = cur.fetchall()
print(rows)
df = pd.DataFrame(rows, columns=['year', 'correlation'])

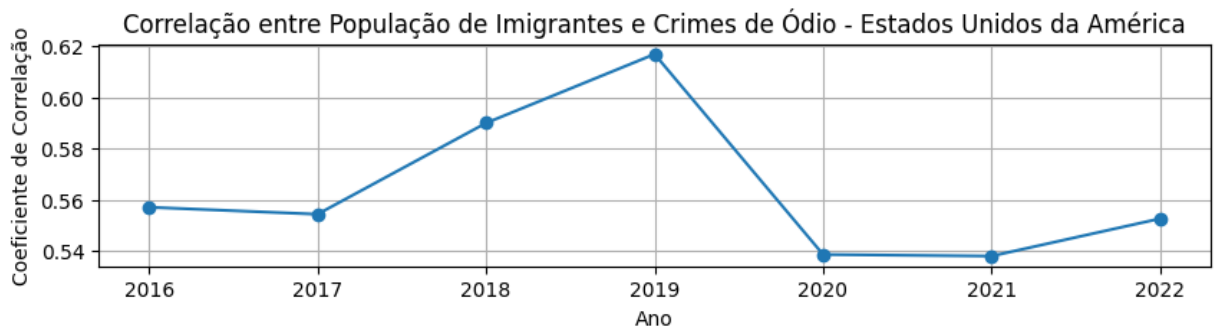
plt.figure(figsize=(10, 2))
plt.plot(df['year'], df['correlation'], marker='o')
plt.title('Correlação entre População de Imigrantes e Crimes de Ódio – Estado de New York')
plt.xlabel('Ano')
plt.ylabel('Coeficiente de Correlação')
plt.grid(True)
plt.show()

```

```

[(2016, 0.5571002620408528), (2017, 0.5543728250116345), (2018, 0.5900240235615178), (2019, 0.616948591936839), (2020, 0.5386170021982203), (2021, 0.5380031134426261), (2022, 0.5526348347724558)]

```



In [ ]:

```

query = """WITH crime_counts AS (
    SELECT
        state_name,
        year,
        COUNT(*) AS total_crimes
    FROM
        crimes
    GROUP BY
        state_name, year
),
immigrant_populations AS (
    SELECT
        state_name,
        UNNEST(ARRAY[2016, 2017, 2018, 2019, 2020, 2021, 2022]) AS year,
        UNNEST(ARRAY[populationimi2016, populationimi2017, populationimi2018,
        populationimi2019, populationimi2020, populationimi2021, populationim
    FROM
        datastates_country where euavseurope = 'Europe'
),
combined_data AS (
    SELECT
        ic.state_name,
        ic.year,
        ic.total_crimes,
        ip.population_immigrants
    FROM
        crime_counts ic
    JOIN
        immigrant_populations ip ON ic.state_name = ip.state_name AND ic.year
),
calculations AS (
    SELECT
        year,
        COUNT(*) AS n,
        SUM(total_crimes) AS sum_y,
        SUM(population_immigrants) AS sum_x,
        SUM(total_crimes * population_immigrants) AS sum_xy,
        SUM(total_crimes * total_crimes) AS sum_y2,
        SUM(population_immigrants * population_immigrants) AS sum_x2
    FROM
        combined_data
    GROUP BY
        year
)
SELECT
    year,
    (n * sum_xy - sum_x * sum_y) /
    SQRT((n * sum_x2 - sum_x * sum_x) * (n * sum_y2 - sum_y * sum_y)) AS corr
FROM
    calculations
ORDER BY
    year;

```

```

cur = conn.cursor()
cur.execute(query)

rows = cur.fetchall()
print(rows)
df = pd.DataFrame(rows, columns=['year', 'correlation'])

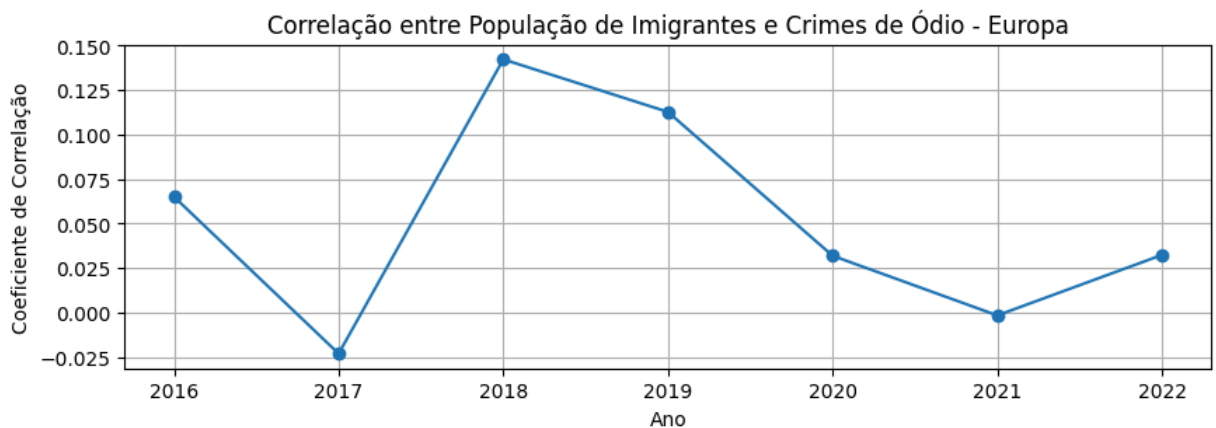
plt.figure(figsize=(10, 3))
plt.plot(df['year'], df['correlation'], marker='o')
plt.title('Correlação entre População de Imigrantes e Crimes de Ódio - Europa')
plt.xlabel('Ano')
plt.ylabel('Coeficiente de Correlação')
plt.grid(True)
plt.show()

```

```

[(2016, 0.06507328712615501), (2017, -0.02309054776328406), (2018, 0.14228675593591), (2019, 0.112738682537761), (2020, 0.0319027993025612), (2021, -0.0016188094256488147), (2022, 0.03236416357161542)]

```



Com bases nos resultados da análise da correlação entre a população de imigrantes e o número de crimes de ódio nos Estados Unidos e na Europa ao longo dos anos de 2016 a 2022, podemos elaborar uma conclusão:

- **Estados Unidos**

- Observamos uma correlação positiva moderada entre o número de imigrantes e o número de crimes de ódio ao longo dos anos analisados.
- Os coeficientes de correlação variaram de aproximadamente 0.54 a 0.62, indicando uma associação consistente e moderadamente forte entre o aumento do número de imigrantes e o aumento do número de crimes de ódio.
- Esses resultados sugerem que, em geral, há uma tendência de aumento nos crimes de ódio à medida que a população de imigrantes cresce nos Estados Unidos.

- **Europa**

- Na Europa, a relação entre o número de imigrantes e os crimes de ódio foi mais variável ao longo dos anos.
- Houve anos em que a correlação foi positiva, indicando uma associação entre o aumento do número de imigrantes e o aumento de crimes de ódio, embora com força variável.
- Em outros anos, a correlação foi negativa, sugerindo que em determinados períodos um aumento do número de imigrantes está associado a uma diminuição nos crimes de ódio.



- Visto que os valores da correlação são muito baixos, o número de imigrantes não tem impacto no aumento/diminuição do número de crimes de ódio.

Com base nos resultados analisados, podemos inferir que na Europa, o aumento do número de imigrantes não está associado a um aumento dos crimes de ódio, pelo contrário, os Estados Unidos, tem coeficientes de correlação positivos e moderadamente fortes, o que indicam uma tendência consistente nessa direção. Uma das razões para tal acontecer, é que o aumento do número de imigrantes significa que novas religiões, novas crenças e raças estão a ser colocadas na sociedade, sendo que os Estados Unidos poderá ser uma sociedade mais racista.

## Qual a correlação entre o Produto Interno Bruto e o número de crimes de ódio (EUA vs Europa)?

A correlação entre o Produto Interno Bruto (PIB) e o aumento do número de crimes é um tema complexo que envolve uma variedade de fatores económicos, sociais e políticos. Embora exista uma associação geral entre o crescimento económico e a diminuição da criminalidade, a relação entre o PIB e certos tipos de crimes, como os crimes de ódio, pode ser mais ambígua.

Em muitos casos, um aumento no PIB pode estar correlacionado com um aumento no número de crimes devido a vários fatores. Por exemplo, o aumento da riqueza económica pode levar a disparidades sociais mais amplas, aumentando o potencial de tensões entre grupos e conflitos.

In [ ]:

```
query = """WITH crime_counts AS (
    SELECT
        state_name,
        year,
        COUNT(*) AS total_crimes
    FROM
        crimes
    GROUP BY
        state_name, year
),
gdp_per_capita AS (
    SELECT
        euavseurope,
        state_name,
        UNNEST(ARRAY[2016, 2017, 2018, 2019, 2020, 2021, 2022]) AS year,
        UNNEST(ARRAY[gdp2016, gdp2017, gdp2018,
            gdp2019, gdp2020, gdp2021, gdp2022]) AS gdp_per_capita
    FROM
        datastates_country
)
SELECT

    cc.year,
    (COUNT(*) * SUM(cc.total_crimes * gdp.gdp_per_capita) - SUM(cc.total_crimes * gdp.gdp_per_capita) / COUNT(*)) /
    (SQRT((COUNT(*) * SUM(cc.total_crimes * cc.total_crimes) - SUM(cc.total_crimes * cc.total_crimes) / COUNT(*)) *
    (COUNT(*) * SUM(gdp.gdp_per_capita * gdp.gdp_per_capita) - SUM(gdp.gdp_per_capita * gdp.gdp_per_capita) / COUNT(*)))
    FROM
        crime_counts cc
    JOIN
        gdp_per_capita gdp ON cc.state_name = gdp.state_name AND cc.year = gdp.year
```

```

GROUP BY
    cc.year,gdp.euavseurope
ORDER BY
    gdp.euavseurope,cc.year;

-----

cur = conn.cursor()
cur.execute(query)

rows = cur.fetchall()

lista = rows[0:7]
print('Europa',lista)
df = pd.DataFrame(lista, columns=['year', 'correlation'])

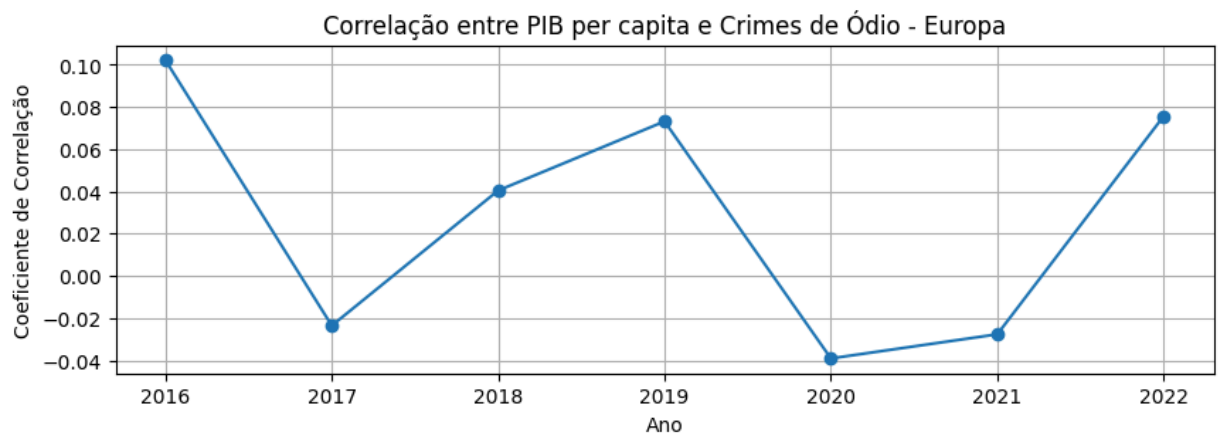
plt.figure(figsize=(10, 3))
plt.plot(df['year'], df['correlation'], marker='o')
plt.title('Correlação entre PIB per capita e Crimes de Ódio - Europa' )
plt.xlabel('Ano')
plt.ylabel('Coeficiente de Correlação')
plt.grid(True)
plt.show()

lista = rows[8:]
print('United States of America',lista)
df = pd.DataFrame(lista, columns=['year', 'correlation'])

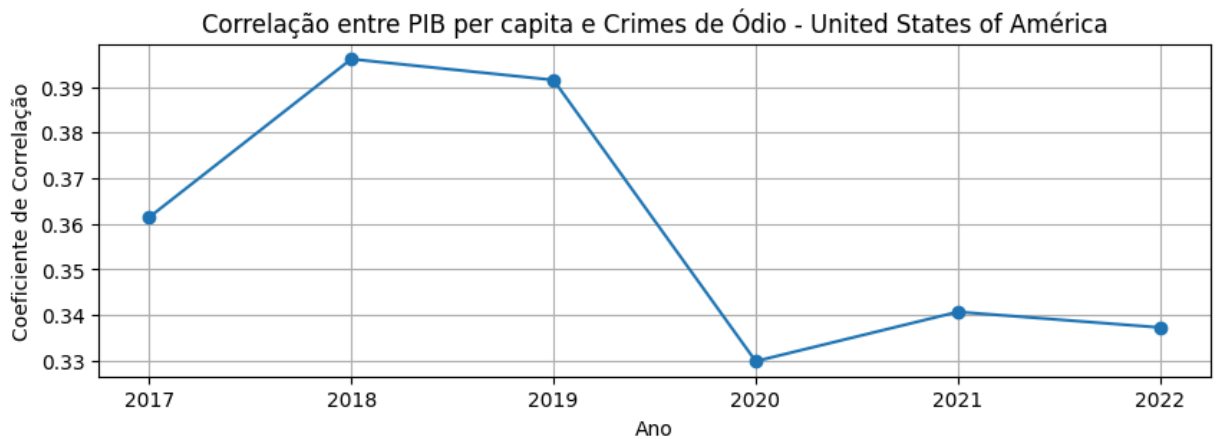
plt.figure(figsize=(10, 3))
plt.plot(df['year'], df['correlation'], marker='o')
plt.title('Correlação entre PIB per capita e Crimes de Ódio - United States o
plt.xlabel('Ano')
plt.ylabel('Coeficiente de Correlação')
plt.grid(True)
plt.show()

```

Europa [(2016, Decimal('0.1021812207099194550497')), (2017, Decimal('-0.0235844267449697507501')), (2018, Decimal('0.040434874745049446708103')), (2019, Decimal('0.073194885584812298062391')), (2020, Decimal('-0.039135426823792195731596')), (2021, Decimal('-0.027703156373672508046546')), (2022, Decimal('0.0756670455242743395698'))]



United States of America [(2017, Decimal('0.36139911038855784247')), (2018, Decimal('0.39610096466625033313')), (2019, Decimal('0.39151731980293639074')), (2020, Decimal('0.32986787725721045433')), (2021, Decimal('0.34068485046798555177')), (2022, Decimal('0.33726714770434096228'))]



Analisando os resultados fornecidos para a correlação entre o PIB per capita (gdp) e o número de crimes de ódio nos Estados Unidos da América e na Europa, podemos observar:

- **Estados Unidos**

- A correlação ao longo dos anos é bastante semelhante, e esta é positiva mas baixa. Isso sugere que existe uma associação fraca entre essas duas variáveis. Embora haja uma tendência para um maior PIB estar associado a um maior número de crimes de ódio, a relação não é forte o suficiente para fazer previsões precisas sobre os crimes de ódio com base apenas no PIB per capita.

- **Europa**

- A correlação entre o PIB per capita e o número de crimes de ódio na Europa é menos consistente ao longo dos anos, com flutuações significativas.
- Houve anos em que a correlação foi positiva e outros anos negativa.

Nos Estados Unidos, parece haver uma associação positiva mas fraca entre o PIB per capita e o número de crimes de ódio, com o aumento da riqueza económica muitas vezes acompanhado pelo aumento dos crimes de ódio.

## Qual a correlação entre o densidade populacional e o número de crimes de ódio (EUA vs Europa)?

Nesta análise, concentramo-nos em comparar os Estados Unidos da América (EUA) e a Europa, duas regiões que apresentam uma ampla diversidade populacional e uma variedade de dinâmicas sociais. Nosso objetivo é entender se existe uma relação significativa entre a densidade populacional - uma medida da concentração de pessoas em uma determinada área - e o número de crimes de ódio nessas regiões.

In [ ]:

```
query = """WITH crime_counts AS (
    SELECT
        state_name,
        year,
        COUNT(*) AS total_crimes
    FROM
        crimes
    WHERE year >= 2016
    GROUP BY
        state_name, year
),
population_density AS (
```

```

SELECT
    euavseurope,
    state_name,
    pop2016,
    UNNEST(ARRAY[pop2016/totalarea, pop2017/totalarea, pop2018/totalarea,
totalarea
FROM
    datastates_country

    WHERE totalarea > 0
)
SELECT
    cc.year,
    CASE
        WHEN SUM(pd.population) > 0 THEN
            (COUNT(*) * SUM(cc.total_crimes * pd.population) - SUM(cc.total_crimes) *
            Sqrt((COUNT(*) * SUM(cc.total_crimes * cc.total_crimes) - SUM(cc.total_cr
            (COUNT(*) * SUM(pd.population * pd.population) - SUM(pd.population) * SUM
            ELSE
                NULL
        END AS correlation
FROM
    crime_counts cc
JOIN
    population_density pd ON cc.state_name = pd.state_name

group by cc.year, pd.euavseurope

order by pd.euavseurope, cc.year

-----

cur = conn.cursor()
cur.execute(query)

rows = cur.fetchall()

lista = rows[0:7]
print('Europa', lista)
df = pd.DataFrame(lista, columns=['year', 'correlation'])

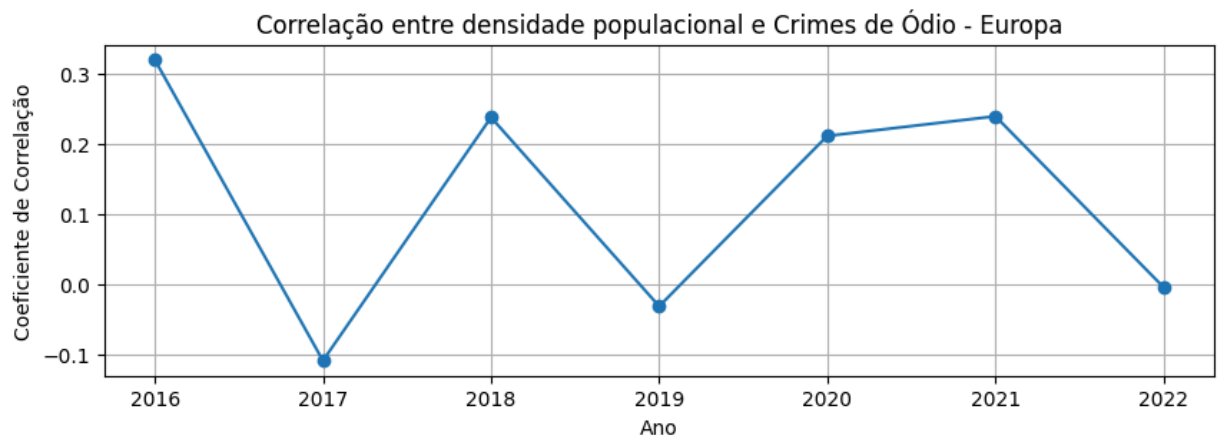
plt.figure(figsize=(10, 3))
plt.plot(df['year'], df['correlation'], marker='o')
plt.title('Correlação entre densidade populacional e Crimes de Ódio – Europa')
plt.xlabel('Ano')
plt.ylabel('Coeficiente de Correlação')
plt.grid(True)
plt.show()

lista = rows[8:]
print('United States of America', lista)
df = pd.DataFrame(lista, columns=['year', 'correlation'])

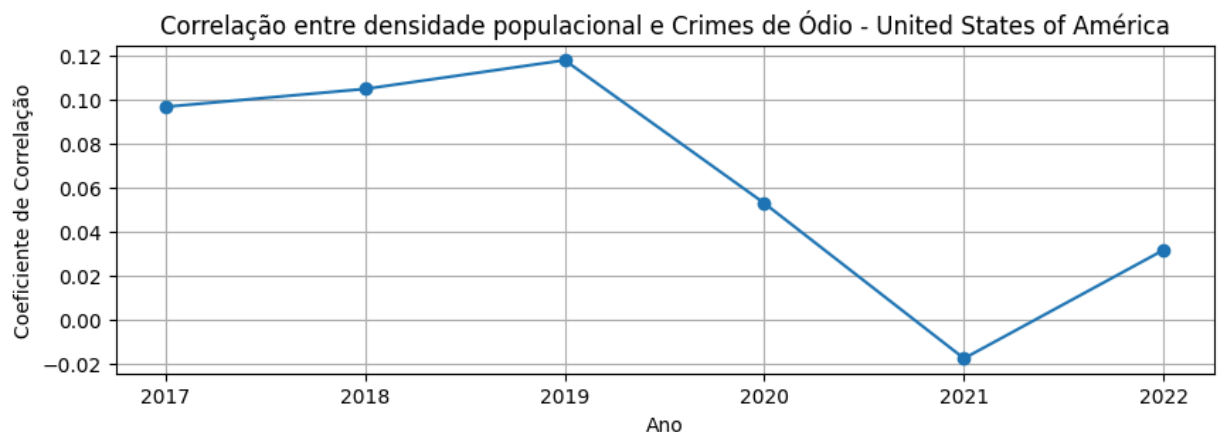
plt.figure(figsize=(10, 3))
plt.plot(df['year'], df['correlation'], marker='o')
plt.title('Correlação entre densidade populacional e Crimes de Ódio – United
plt.xlabel('Ano')
plt.ylabel('Coeficiente de Correlação')
plt.grid(True)
plt.show()

```

Europa [(2016, Decimal('0.32074015126338971577099436252508')), (2017, Decimal('-0.10962448165735972199438714538737')), (2018, Decimal('0.23833439412224389610527992161818')), (2019, Decimal('-0.03182707692629175499493660592552')), (2020, Decimal('0.21181763622727404342931935350050')), (2021, Decimal('0.23993705433182745012874605825072')), (2022, Decimal('-0.00454723386084640505175620267961'))]



United States of America [(2017, Decimal('0.0967603979517542903308757437726128584670')), (2018, Decimal('0.1048684148704837518455171949316541879754')), (2019, Decimal('0.1179061964805167414010178327925643975045')), (2020, Decimal('0.0529777270817925616233480713073629824432')), (2021, Decimal('-0.0174516215388243168959757040365527897633')), (2022, Decimal('0.0317845374991962559987886275168618323893'))]



As correlações entre a densidade populacional e o número de crimes de ódio tanto nos estados unidos quanto na europa são bastante baixas. As correlações variam de muito baixas a negativas, indicando uma correlação fraca entre estes dois fatores. A correlação fraca sugere que outros fatores podem estar a desempenhar um papel mais significativo na ocorrência de crimes de ódio do que simplesmente a densidade populacional.

## Qual a correlação entre a idade populacional e o número de crimes de ódio (EUA vs Europa)?

Nesta análise, concentramo-nos em comparar os Estados Unidos da América (EUA) e a Europa, nosso objetivo é entender se existe uma relação significativa entre a idade populacional - a média da idade de pessoas numa determinada área - e o número de crimes de ódio nessas regiões.

```
In [ ]: query = """WITH crime_counts AS (
        SELECT
            state_name,
            year,
            COUNT(*) AS total_crimes
        FROM
```

```

        crimes
        WHERE year >= 2016
    GROUP BY
        state_name, year
),
agesCorr AS (
    SELECT
        euavseurope,
        state_name,

        UNNEST(ARRAY[age2016, age2017, age2018, age2019, age2020, age2021, age2022]) AS
        totalarea
    FROM
        datastates_country
)
SELECT
    cc.year,

    (COUNT(*) * SUM(cc.total_crimes * ac.age) - SUM(cc.total_crimes) * SUM(ac.age)) /
    SQRT((COUNT(*) * SUM(cc.total_crimes * cc.total_crimes) - SUM(cc.total_crimes) *
    SUM(cc.total_crimes)) * (COUNT(*) * SUM(ac.age * ac.age) - SUM(ac.age) * SUM(ac.age)))
AS correlation
FROM
    crime_counts cc
JOIN
    agesCorr ac ON cc.state_name = ac.state_name

group by cc.year, ac.euavseurope

order by ac.euavseurope, cc.year""")

cur = conn.cursor()
cur.execute(query)

rows = cur.fetchall()

lista = rows[0:7]
print('Europa', lista)
df = pd.DataFrame(lista, columns=['year', 'correlation'])

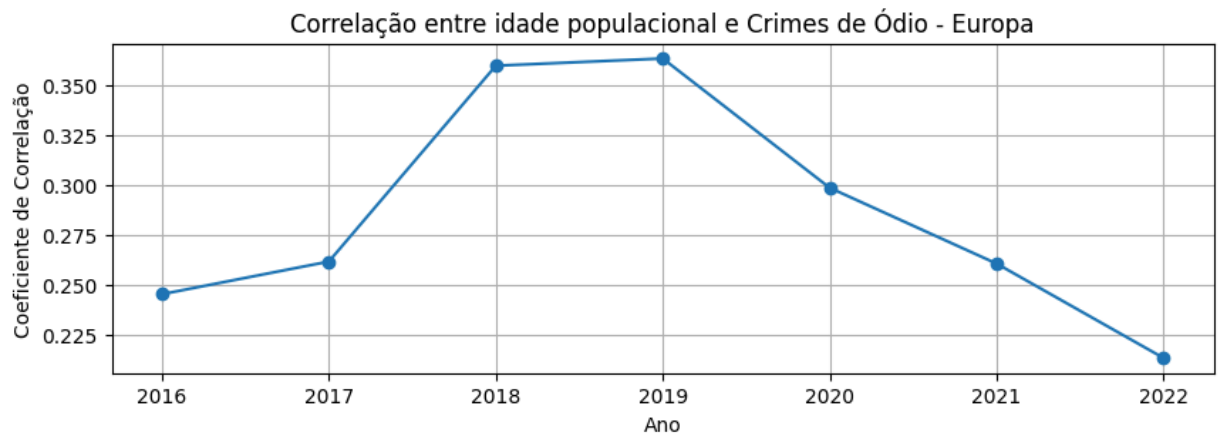
plt.figure(figsize=(10, 3))
plt.plot(df['year'], df['correlation'], marker='o')
plt.title('Correlação entre idade populacional e Crimes de Ódio – Europa')
plt.xlabel('Ano')
plt.ylabel('Coeficiente de Correlação')
plt.grid(True)
plt.show()

lista = rows[8:]
print('United States of America', lista)
df = pd.DataFrame(lista, columns=['year', 'correlation'])

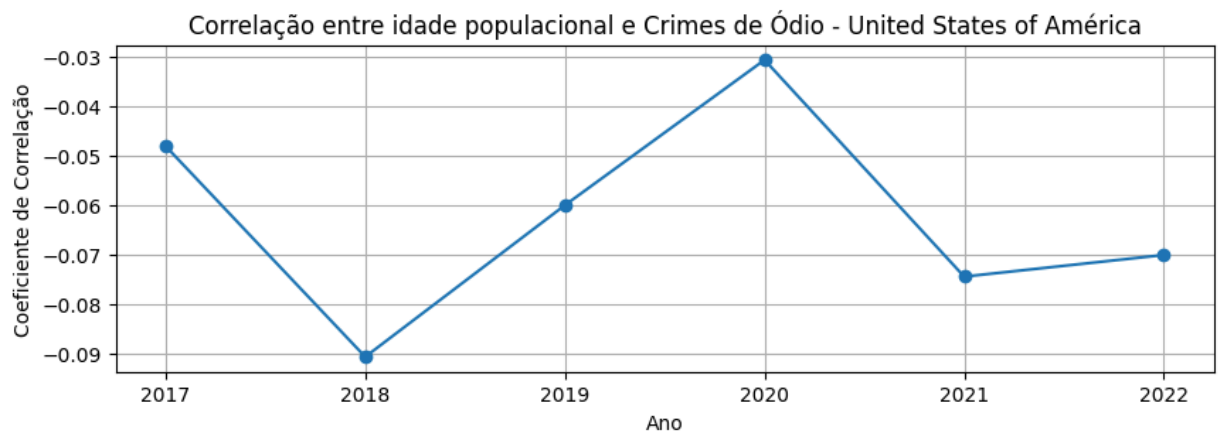
plt.figure(figsize=(10, 3))
plt.plot(df['year'], df['correlation'], marker='o')
plt.title('Correlação entre idade populacional e Crimes de Ódio – United States')
plt.xlabel('Ano')
plt.ylabel('Coeficiente de Correlação')
plt.grid(True)
plt.show()

```

Europa [(2016, Decimal('0.24550149360541538282')), (2017, Decimal('0.26180446418843633054')), (2018, Decimal('0.35968090189253582952')), (2019, Decimal('0.36328269192845744148')), (2020, Decimal('0.29865480204988983923')), (2021, Decimal('0.26076088341291104140')), (2022, Decimal('0.21359928193321490449'))]



United States of America [(2017, Decimal('-0.04799462015016359373')), (2018, Decimal('-0.09068366240206566704')), (2019, Decimal('-0.05997108400043677818')), (2020, Decimal('-0.03054068881474453845')), (2021, Decimal('-0.07443290876479574472')), (2022, Decimal('-0.07005184098523522960'))]



Analisando os resultados fornecidos para a correlação entre a idade média populacional e o número de crimes de ódio nos Estados Unidos da América e na Europa, podemos observar:

- **Estados Unidos**

- A correlação ao longo dos anos é bastante semelhante, e esta é negativa mas baixa. Isso sugere que existe uma associação fraca entre essas duas variáveis. Embora haja uma tendência para uma diminuição do número de idade estar associado a um maior número de crimes de ódio, a relação não é forte o suficiente para fazer previsões precisas sobre os crimes de ódio com base apenas no PIB per capita.

- **Europa**

- A correlação entre a idade populacional e o número de crimes de ódio na Europa é consistente ao longo dos anos, sendo que existe uma associação entre estas duas variáveis, mas mesmo assim é um número baixo. O que não nos permite ter uma conclusão mais detalhada

De qualquer forma, a correlação é mais alta na Europa do que nos Estados Unidos, o que implica que na Europa a idade tem um impacto na ocorrência apesar de ser um impacto bastante baixo.



## Quais países da Europa têm taxas de crimes de ódio semelhantes às dos estados dos EUA?

A questão sobre as taxas de crimes de ódio entre países da Europa e estados dos Estados Unidos desperta interesse devido à sua relevância para a compreensão das dinâmicas sociais e culturais que influenciam a incidência desses crimes em diferentes contextos geográficos e políticos. Os crimes de ódio representam uma violação grave dos direitos humanos e são uma preocupação global, requerendo uma análise abrangente e comparativa para identificar padrões e tendências em diferentes partes do mundo.

Nesta análise, exploraremos os países da Europa que apresentam taxas de crimes de ódio semelhantes às dos estados dos Estados Unidos, examinando dados quantitativos e qualitativos para identificar padrões e fatores comuns que possam contribuir para a ocorrência desses crimes em ambos os EUA e Europa.

Essa consulta fornecerá uma lista de estados dos EUA e países da Europa que têm taxas de crimes de ódio por km<sup>2</sup> semelhantes, considerando a faixa de comparação de  $\pm 1\%$ .

In [ ]:

```
query = """

WITH total_crimes_us AS (
    SELECT
        c.state_name AS state,
        COUNT(*) AS total_crimes,
        dsc.totalarea AS area
    FROM crimes c
    JOIN datastates_country dsc ON c.state_name = dsc.state_name
    WHERE dsc.euavseurope = 'United States of America'
    GROUP BY c.state_name, dsc.totalarea
),
total_crimes_europe AS (
    SELECT
        c.state_name AS country,
        COUNT(*) AS total_crimes,
        dsc.totalarea AS area
    FROM crimes c
    JOIN datastates_country dsc ON c.state_name = dsc.state_name
    WHERE dsc.euavseurope = 'Europe'
    GROUP BY c.state_name, dsc.totalarea
)
SELECT
    us.state AS us_state,
    us.total_crimes / us.area AS us_crimes_per_km²,
    eur.country AS europe_country,
    eur.total_crimes / eur.area AS europe_crimes_per_km²
FROM total_crimes_us us
JOIN total_crimes_europe eur ON
    eur.total_crimes / eur.area BETWEEN (us.total_crimes / us.area) * 0.99 AND
    (us.total_crimes / us.area) * 1.01
ORDER BY us.state, eur.country;
"""

cur = conn.cursor()
cur.execute(query)

rows = cur.fetchall()

print(rows)
```



```

states = []
europe_crime_rates = []
us_crime_rates = []
for i in range(len(rows)):
    s = rows[i][0] + "/" + rows[i][2]
    states.append(s)
    europe_crime_rates.append(rows[i][3])
    us_crime_rates.append(rows[i][1])

x = np.arange(len(states))
width = 0.25

fig, ax = plt.subplots(figsize=(14, 6))
bars1 = ax.bar(x - width/2, us_crime_rates, width, label='EUA')
bars2 = ax.bar(x + width/2, europe_crime_rates, width, label='Europa')

ax.set_xlabel('Estado / País')
ax.set_ylabel('Taxa de Crimes')
ax.set_title('Comparação de Taxas de Crimes de Ódio entre EUA e Europa')
ax.set_xticks(x)
ax.set_xticklabels(states)
ax.legend()

def add_labels(bars):
    for bar in bars:
        height = bar.get_height()
        ax.annotate(f'{height:.3f}',
                    xy=(bar.get_x() + bar.get_width() / 2, height),
                    xytext=(0, 3),
                    textcoords="offset points",
                    ha='center', va='bottom')

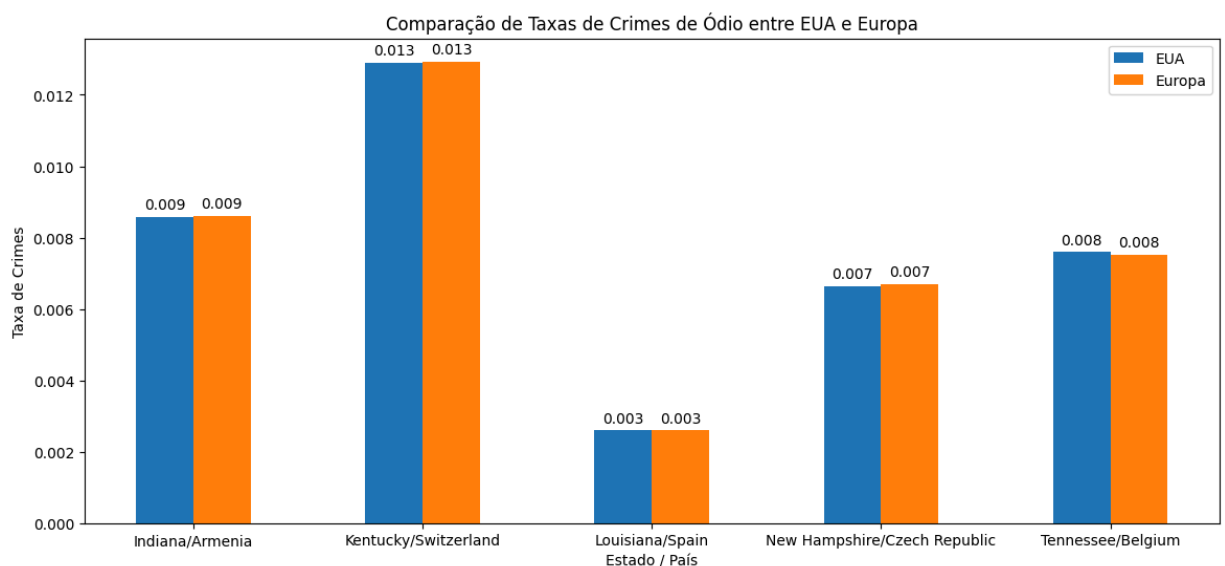
add_labels(bars1)
add_labels(bars2)

```

```

[('Indiana', Decimal('0.00858723999745563259')), ('Armenia', Decimal('0.0086055
4970144011240'))), ('Kentucky', Decimal('0.01290895887478978749')), ('Switzerlan
d', Decimal('0.01293356551319173062'))), ('Louisiana', Decimal('0.002594741226
16265784')), ('Spain', Decimal('0.00259738596515100458'))), ('New Hampshire', De
cimal('0.00664904600644255389')), ('Czech Republic', Decimal('0.006698036451828
55100'))), ('Tennessee', Decimal('0.00759484393466052239')), ('Belgium', Decima
l('0.00752972258916776750')))]

```



Podemos verificar que existem 5 matches entre países Europeus e estados dos Estados Unidos.

In [ ]:

```

for i in range(len(states)):
    query = "select * from datastates_country where state_name = '" + rows[i]
    cur = conn.cursor()
    cur.execute(query)
    res = cur.fetchall()
    print("State/Country ", res[0][1], " ", res[1][1])
    print("Population ", res[0][8], " ", res[1][8])
    print("Area ", res[0][9], " ", res[1][9])
    print("Densidade pop ", round(res[0][8]/res[0][9], 3), " ", round(res[1][8]
    print("GDP ", res[0][16], " ", res[1][16])
    print("Age ", res[0][23], " ", res[1][23])
    print("IMI pop ", res[0][-1], " ", res[1][-1])
    print("\n")

```

State/Country	Indiana	Armenia
Population	6832274	2780469
Area	94326	28470.0
Densidade pop	72.433	97.663
GDP	58329	7018.05150434798
Age	38.20	37.70
IMI pop	0.025	0.017

State/Country	Kentucky	Switzerland
Population	4511563	8775760
Area	104656	39509.6
Densidade pop	43.108	222.117
GDP	51929	93259.9057183024
Age	39.40	43.80
IMI pop	0.02	0.264

State/Country	Louisiana	Spain
Population	4588023	47778340
Area	135659	499733.2
Densidade pop	33.820	95.608
GDP	54527	29674.5442864413
Age	38.20	45.90
IMI pop	0.018	0.136

State/Country	New Hampshire	Czech Republic
Population	1399003	10672118
Area	24214	77186.8
Densidade pop	57.777	138.264
GDP	73711	27226.6156386023
Age	43.30	43.70
IMI pop	0.021	0.067

State/Country	Tennessee	Belgium
Population	7048976	11685814
Area	109153	30280.0
Densidade pop	64.579	385.925
GDP	58311	49926.8254295305
Age	39.20	41.70
IMI pop	0.023	0.14

### Indiana (EUA) vs. Armênia:

Indiana tem uma população significativamente maior do que a Armênia, mas a densidade populacional da Armênia é maior.

Indiana tem um GDP muito maior do que a Armênia, indicando um nível mais alto de desenvolvimento económico.

Ambos os locais têm uma idade média da população semelhante.

Indiana tem um IMI populacional mais alto do que a Armênia, sugerindo uma taxa mais alta de migração internacional.

Concluindo, a alta densidade populacional na Armênia pode levar a tensões sociais e conflitos semelhantes aos observados em áreas mais densamente povoadas, como Indiana nos EUA. Embora Indiana tenha um GDP mais alto, a presença de uma população migrante internacionalmente significativa pode contribuir para a ocorrência de crimes de ódio em ambas as regiões.

### **Kentucky (EUA) vs. Suíça:**

Kentucky tem uma população menor do que a Suíça, mas a densidade populacional da Suíça é muito maior.

A Suíça tem um GDP muito maior do que o Kentucky.

A Suíça tem uma idade média da população mais alta do que o Kentucky.

A Suíça tem um IMI populacional significativamente mais alto do que o Kentucky.

Resumindo, alta densidade populacional na Suíça pode levar a desafios semelhantes aos enfrentados em áreas menos densamente povoadas, como Kentucky. Embora a Suíça tenha um GDP mais alto, a presença de uma população mais envelhecida e uma taxa mais alta de migração internacional pode contribuir para a ocorrência de crimes de ódio em ambas as regiões.

### **Louisiana (EUA) vs. Espanha:**

A população da Espanha é significativamente maior do que a da Louisiana.

A área da Espanha é substancialmente maior do que a da Louisiana.

A densidade populacional da Espanha é maior em comparação com a Louisiana.

A Espanha tem um GDP mais baixo do que a Louisiana.

A idade média da população na Espanha é mais alta do que na Louisiana.

A Espanha tem um IMI populacional mais alto do que a Louisiana.

Ou seja, a maior densidade populacional e a área geográfica maior da Espanha podem contribuir para um ambiente mais propenso a crimes de ódio. Além disso, apesar de ter um GDP mais baixo, a presença de uma população mais envelhecida e uma taxa de migração internacional mais alta podem influenciar a ocorrência de crimes de ódio na Espanha, semelhante à Louisiana.

### **New Hampshire (EUA) vs. República Checa:**

A população da República Checa é substancialmente maior do que a de New Hampshire.

A área da República Checa é significativamente maior do que a de New Hampshire.

A densidade populacional da República Checa é mais alta em comparação com New Hampshire.

New Hampshire tem um GDP mais alto do que a República Checa.

Ambos os locais têm uma idade média da população semelhante.

A República Checa tem um IMI populacional mais alto do que New Hampshire.

A maior densidade populacional e a área geográfica maior da República Checa podem contribuir para um ambiente social mais complexo, semelhante ao de New Hampshire. Embora New Hampshire tenha um GDP mais alto, fatores como a densidade populacional e a idade média da população podem influenciar a ocorrência de crimes de ódio em ambas as regiões.

### **Tennessee (EUA) vs. Bélgica:**

A população da Bélgica é maior do que a de Tennessee.

A área de Tennessee é maior do que a da Bélgica.

A densidade populacional da Bélgica é substancialmente maior do que a de Tennessee.

A Bélgica tem um GDP mais alto do que Tennessee.

Ambos os locais têm uma idade média da população semelhante.

A Bélgica tem um IMI populacional mais alto do que Tennessee.

A maior densidade populacional da Bélgica pode contribuir para uma maior incidência de crimes de ódio, semelhante ao que ocorre em Tennessee. Apesar de Tennessee ter uma área geográfica maior, outros fatores, como a densidade populacional e o GDP mais alto da Bélgica, podem influenciar a ocorrência de crimes de ódio em ambas as regiões.

### **Conclusão:**

As comparações entre os estados dos EUA e os países europeus revelam que, embora haja diferenças significativas em termos de densidade populacional, GDP e taxas de migração internacional, esses fatores se equilibram de maneiras que resultam em taxas de crimes de ódio semelhantes.

A alta densidade populacional e a diversidade cultural emergem como fatores-chave que contribuem para a ocorrência de crimes de ódio, independentemente do nível económico geral. Essas observações sugerem que políticas focadas na coesão social e na integração cultural podem ser eficazes na redução de crimes de ódio em ambos os contextos.

Após realização e análise de dados, iremos apresentar uma conclusão da influência dos fatores demográficos e socioeconómicos nos crimes de ódio

## Conclusão

As taxas de crimes de ódio nos Estados Unidos e na Europa são influenciadas por uma complexa interação de fatores demográficos e socioeconômicos. Ao analisar os dados de crimes de ódio em ambas as regiões, podemos identificar padrões distintos e comuns que elucidam como essas variáveis impactam a ocorrência desses crimes.

Nos Estados Unidos, a alta diversidade étnica e racial é um dos principais fatores que influenciam a prevalência de crimes de ódio. A história de tensões raciais profundamente enraizadas, especialmente entre brancos e negros, tem contribuído para uma maior incidência de crimes raciais e xenofóbicos. Por exemplo, a análise dos dados mostra que a maioria dos crimes de ódio nos EUA é de natureza racial, com um aumento significativo de 4.169 incidentes em 2016 para 8.036 em 2022. Este aumento pode ser atribuído a várias razões, incluindo o aumento da visibilidade das questões raciais e a polarização política.

Na Europa, a diversidade étnica também é um fator crucial, embora a distribuição dos crimes de ódio seja mais variada. Os crimes racistas e xenofóbicos são prevalentes, mas há também uma incidência significativa de crimes anti-cristãos e anti-LGBTI. A diversidade de culturas, línguas e políticas de imigração entre os países europeus contribui para uma dinâmica complexa de crimes de ódio.

A presença de populações imigrantes é outro fator demográfico significativo. Nos Estados Unidos, existe uma correlação positiva moderada entre o número de imigrantes e o número de crimes de ódio, com coeficientes de correlação variando de aproximadamente 0,54 a 0,62. Isso sugere que à medida que a população imigrante cresce, há uma tendência de aumento nos crimes de ódio. Este fenômeno pode ser explicado pela resistência de certos segmentos da população nativa às mudanças demográficas e culturais trazidas pela imigração.

Na Europa, a relação entre o número de imigrantes e os crimes de ódio é mais variável. Em alguns anos, há uma correlação positiva, enquanto em outros a correlação é negativa. Isso indica que o impacto da imigração nos crimes de ódio na Europa é influenciado por outros fatores contextuais, como políticas de integração, tensões políticas e eventos específicos que podem aumentar ou diminuir as tensões sociais.

A idade da população também influencia as taxas de crimes de ódio. Nos Estados Unidos, há uma correlação negativa, embora fraca, entre a idade média da população e o número de crimes de ódio. Isso sugere que populações mais jovens tendem a ter uma maior incidência de crimes de ódio. Esse padrão pode ser explicado pelo fato de que os jovens estão mais frequentemente envolvidos em atividades sociais e interações onde podem ocorrer conflitos.

Na Europa, a correlação entre a idade populacional e os crimes de ódio é mais consistente, embora ainda fraca. Populações mais jovens na Europa também tendem a apresentar uma maior incidência de crimes de ódio, possivelmente devido a fatores semelhantes aos observados nos Estados Unidos.

O PIB per capita tem uma correlação fraca, mas positiva, com os crimes de ódio nos Estados Unidos. Isso indica que, em geral, estados mais ricos podem ter uma maior incidência de crimes de ódio. Esta relação pode ser complexa, envolvendo fatores como desigualdade económica, que pode exacerbar tensões sociais e preconceitos.

Na Europa, a relação entre o PIB per capita e os crimes de ódio é menos consistente, variando de ano para ano. Em alguns períodos, há uma correlação positiva, enquanto em outros, a correlação é negativa. Isso sugere que a riqueza económica sozinha não é um determinante claro dos crimes de ódio na Europa, onde outros fatores contextuais podem desempenhar um papel mais significativo.

Tanto nos Estados Unidos quanto na Europa, a densidade populacional não mostra uma forte correlação com os crimes de ódio. Isso indica que a quantidade de pessoas por quilômetro quadrado não é um fator determinante na ocorrência desses crimes.

A sazonalidade dos crimes de ódio também revela padrões interessantes. Nos Estados Unidos, os crimes de ódio atingem picos nos meses de verão, particularmente em junho e julho. Este padrão pode estar relacionado a eventos como o mês do orgulho LGBTQ+ em junho, que aumenta a visibilidade e, conseqüentemente, a vulnerabilidade da comunidade LGBTI. Na Europa, os picos também ocorrem durante o verão, embora em menor escala comparado aos EUA, e podem estar associados a diferentes eventos culturais e políticos.

Em suma, os fatores demográficos e socioeconómicos têm uma influência significativa nas taxas de crimes de ódio nos Estados Unidos e na Europa. A diversidade étnica e racial, a população de imigrantes, a idade populacional, o PIB per capita, a densidade populacional e a desigualdade económica interagem de maneiras complexas para moldar a ocorrência desses crimes. Embora existam padrões comuns, como a maior prevalência de crimes de ódio em populações mais diversas e jovens, as diferenças regionais e contextuais também desempenham um papel crucial. Políticas eficazes para combater os crimes de ódio devem, portanto, considerar essas variáveis multifacetadas e adaptar-se às especificidades de cada contexto regional.