

Projeto Final

PretaLab Ciclo 13 - Inteligência Artificial

Alunas:

Amanda da Silva Gonçalves
Camille dos Santos Nogueira
Griselda Karen Sillerico Justo
Manuela De Oliveira Souza Brito
Priscila Estevão Da Cunha
Raysa Leide De Oliveira

Relatório de Violência contra LGBTQIA+

Introdução ao tema e justificativa.

A violência contra pessoas LGBTQIA+ no Brasil é uma realidade persistente e alarmante, marcada por profundas desigualdades estruturais, heranças históricas de preconceito e a contínua ausência de políticas públicas efetivas de proteção e inclusão. De acordo com o relatório anual do Grupo Gay da Bahia (GGB), o Brasil continua entre os países que mais matam pessoas LGBTQIA+ no mundo. Em 2024, foram registradas 291 mortes violentas motivadas por LGBTfobia - um aumento de mais de 8% em relação ao ano anterior (Brasil de Fato, 2025). Esses números revelam não apenas a gravidade do problema, mas também a urgência de políticas comprometidas com a promoção dos direitos humanos e a garantia da vida e dignidade da população LGBTQIA+.

Este estudo se justifica pela necessidade de compreender os padrões da violência contra a população LGBTQIA+, identificar os grupos mais vulneráveis e contribuir para o desenvolvimento de estratégias de enfrentamento baseadas em evidências. A análise realizada utiliza dados do período de 2011 a 2019 e, embora não contemple os registros mais recentes, oferece um panorama relevante para identificar tendências e desigualdades históricas.

Como trabalho de conclusão do curso Pretalab - Ciclo 13, na trilha de Inteligência Artificial com foco em análise de dados, este projeto busca demonstrar a aplicabilidade de técnicas exploratórias e preditivas no enfrentamento de problemas sociais complexos. A proposta é utilizar abordagens quantitativas para organizar, visualizar e interpretar dados sobre homicídios contra a população LGBTQIA+, identificando padrões temporais, geográficos e demográficos. A análise também envolve a construção de um modelo preditivo com base em atributos demográficos,

a fim de investigar o potencial da inteligência artificial na geração de insights que possam apoiar futuras pesquisas e reflexões sobre políticas públicas.

Metodologia de análise.

Este estudo foi conduzido por meio de uma análise exploratória e preditiva em relação aos dados sobre homicídios de pessoas LGBTQIA+ no Brasil, com a utilização de quatro bases de dados distintas contemplando os diferentes recortes demográficos e geográficos. A metodologia adotada seguiu as seguintes etapas:

1. Coleta e carregamento dos dados

Quatro arquivos no formato CSV foram utilizados, contendo as seguintes informações:

- Dados gerais do Brasil
- Distribuição por local
- Distribuição por raça/cor
- Distribuição por grupos específicos da população LGBTQIA+

2. Pré-processamento dos dados

Em cada conjunto de dados foi realizada a inspeção para verificar a estrutura da base de dados, os tipos de dados, a presença de valores ausentes (tratados com o preenchimento com zero), a integridade das colunas relevantes para a análise posterior e a confirmação de valores únicos nas colunas categóricas.

3. Análise exploratória dos dados

Foram realizadas análises descritivas e visualizações gráficas para identificação de padrões e distribuições relevantes para a análise, como:

- Quantidade total de homicídios por local, utilizando-se gráficos de barras para identificar os locais com maior incidência.
- Distribuição de homicídios por raça/cor, destacando-se as diferenças demográficas através de gráficos de barras.
- Distribuição de homicídios entre grupos específicos da população LGBTQIA+, utilizando-se gráficos de barras para visualização.

4. Modelagem preditiva

Com o objetivo de prever o número de homicídios com base em atributos demográficos da população LGBTQIA+, foi desenvolvida uma modelagem preditiva utilizando algoritmos de aprendizado de máquina. A base de dados utilizada foi extraída da distribuição por raça/cor, e o processo envolveu as seguintes etapas:

- ❖ **Pré-processamento dos dados:** As variáveis categóricas foram transformadas em variáveis numéricas por meio da técnica de codificação one-hot (dummies), permitindo que fossem utilizadas por algoritmos de regressão.

- ❖ **Divisão dos dados:** O conjunto foi dividido entre subconjuntos de treino e teste, com o objetivo de avaliar o desempenho preditivo dos modelos em dados não vistos.
- ❖ **Treinamento de modelos:** Inicialmente, foi utilizado o algoritmo Random Forest Regressor, conhecido por sua robustez e capacidade de lidar com dados tabulares e padrões não lineares. Além disso, foram aplicados outros modelos de regressão para fins comparativos, como:
 - Regressão Ridge, que utiliza regularização L2 para controle de overfitting;
 - Regressão Linear simples, utilizada como modelo base.
- ❖ **Avaliação dos modelos:** Os modelos foram avaliados por meio de métricas como:
 - Erro médio absoluto (MAE);
 - Erro quadrático médio (MSE);
 - Coeficiente de determinação (R^2).
- ❖ **Validação cruzada:** Para o modelo de Regressão Ridge, foi aplicada validação cruzada com o uso do GridSearchCV, a fim de testar diferentes valores do hiperparâmetro alpha e otimizar a performance do modelo.
- ❖ **Visualização dos resultados:** A performance do modelo Random Forest foi analisada por meio de gráficos de densidade (KDE plots), comparando visualmente a distribuição dos valores reais e dos valores previstos.
- ❖ **Análise de importância das variáveis:** A partir do modelo Random Forest, foi possível extrair a importância relativa de cada variável demográfica, fornecendo insights sobre quais atributos possuem maior influência na predição do número de homicídios.

Essa abordagem permitiu não apenas realizar previsões, mas também compreender os fatores sociodemográficos que mais contribuem para a ocorrência de homicídios na população LGBTQIA+, promovendo uma análise quantitativa orientada por dados.

5. Interpretação e apresentação dos resultados

Os resultados da análise exploratória e do modelo preditivo foram apresentados por meio de gráficos e tabelas, evidenciando os principais insights a respeito da distribuição e evolução dos homicídios contra pessoas LGBTQIA+ no Brasil. Através da metodologia adotada, foi possível uma compreensão ampla dos dados e uma previsão inicial das variáveis que influenciam os homicídios, permitindo uma contribuição no embasamento de políticas públicas e posteriores estudos.

Visualizações (gráficos, mapas, tabelas).

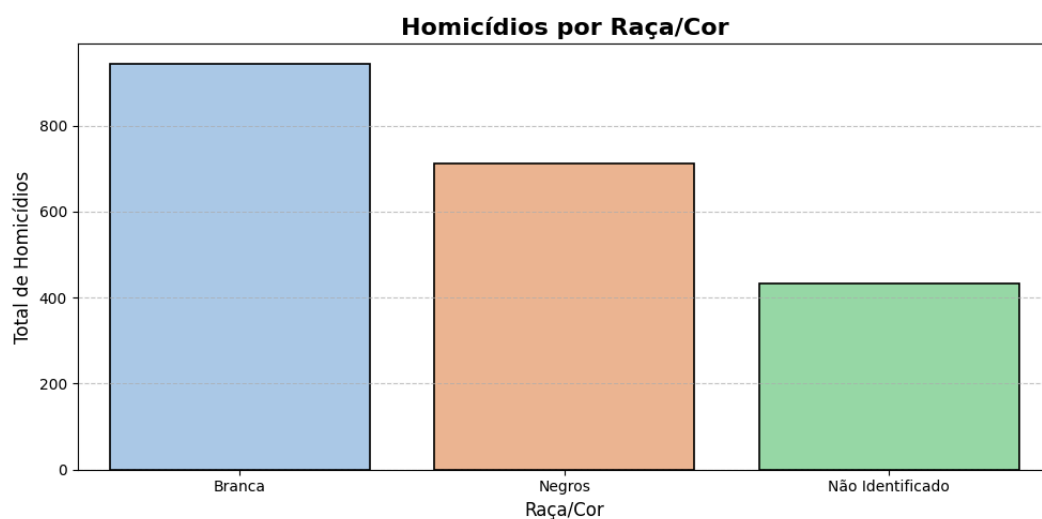


Figura 1 - Gráfico de Homicídios por Raça/cor com o grupo Negros sendo composto por Pardos e Pretos.

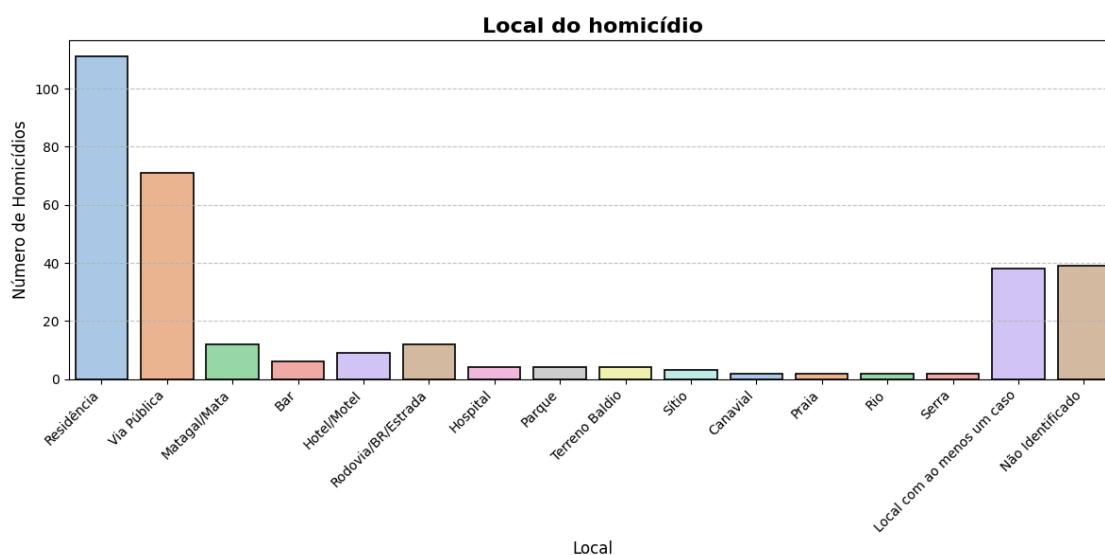
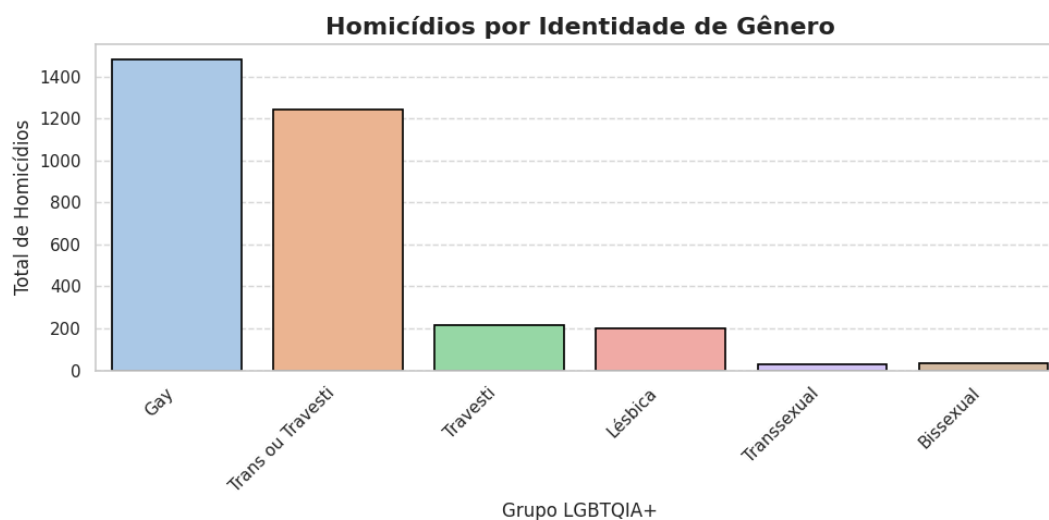
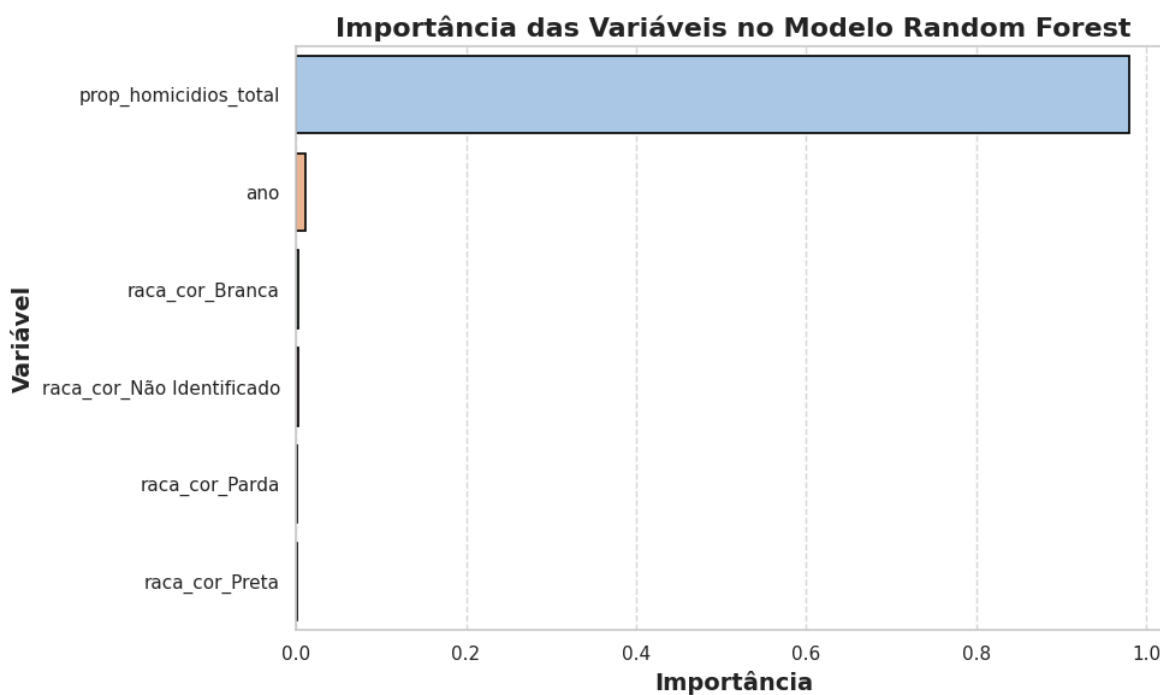
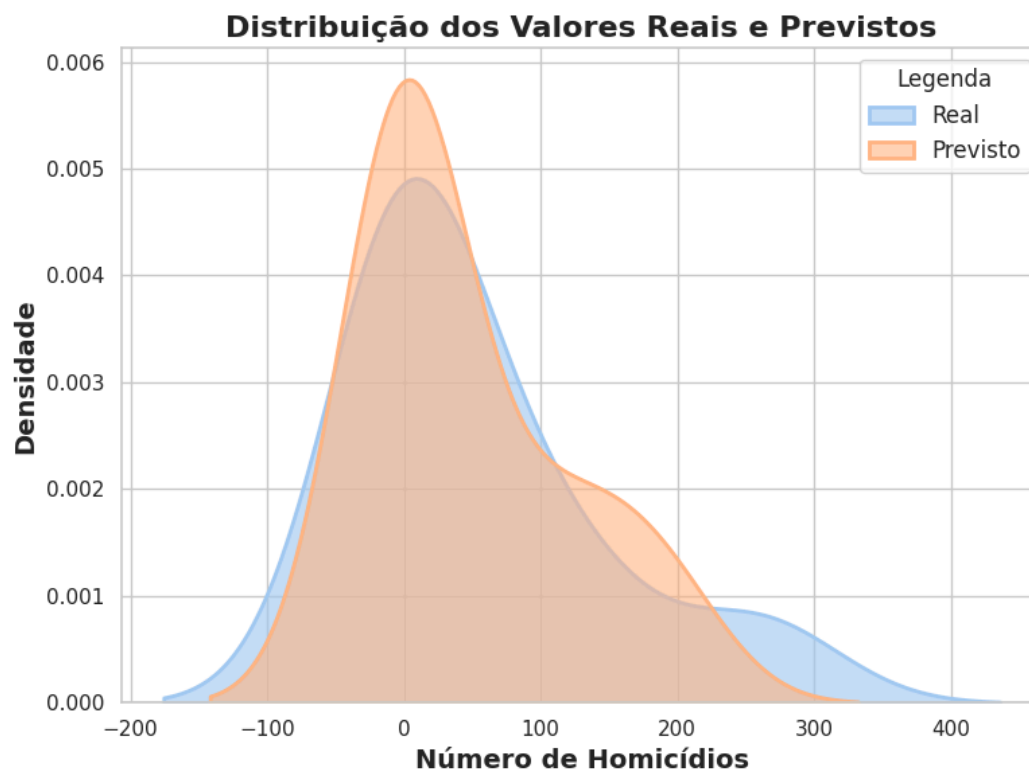


Figura 2 - Gráfico do Local do homicídio



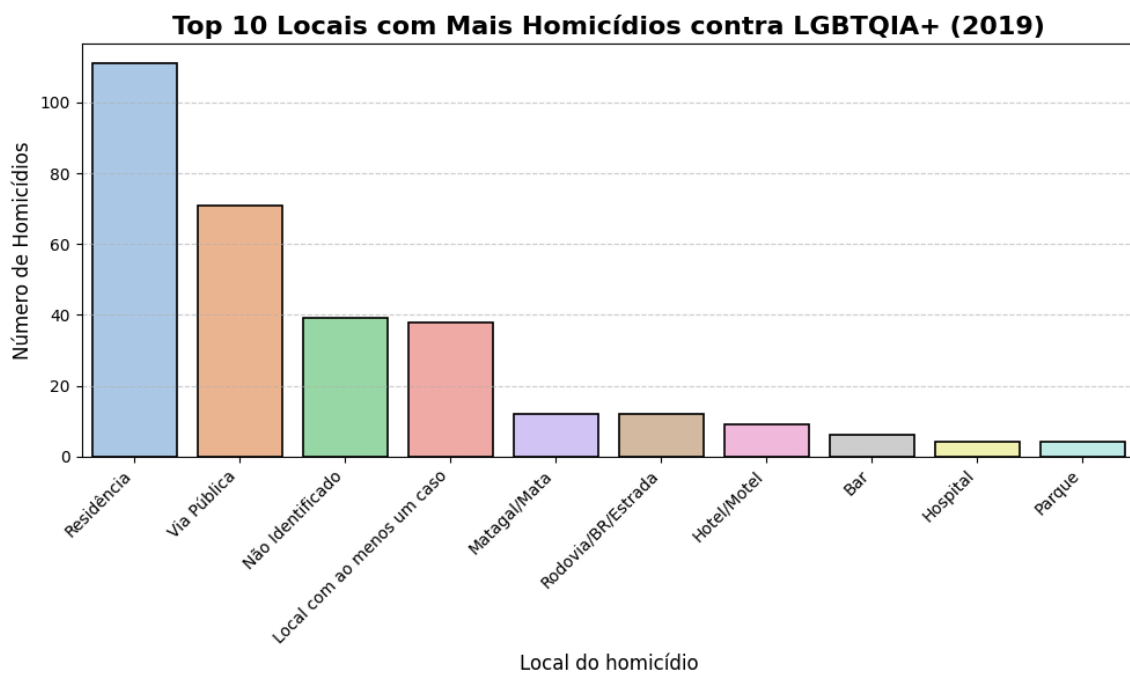


Respostas às perguntas orientadoras.

Realizou-se uma exploração da série temporal dos homicídios ao longo dos anos, a fim de identificar as tendências e mudanças ao longo do tempo, representadas graficamente. Dessa forma, foram respondidas as perguntas orientadoras propostas pela docente do curso.

- **Quais locais têm as maiores taxas de violência contra a população LGBTQIA+?**

Embora os gráficos analisados não apresentem diretamente os nomes dos estados brasileiros, ele revela os locais físicos onde os homicídios contra pessoas LGBTQIA+ ocorreram com maior frequência. A análise mostra que os principais cenários de violência são as residências e as vias públicas, seguidos por matagal ou mata, hotéis e motéis, e rodovias ou estradas. Esses dados indicam que a violência não está restrita a um tipo específico de ambiente, mas se manifesta tanto em espaços privados quanto em locais públicos.



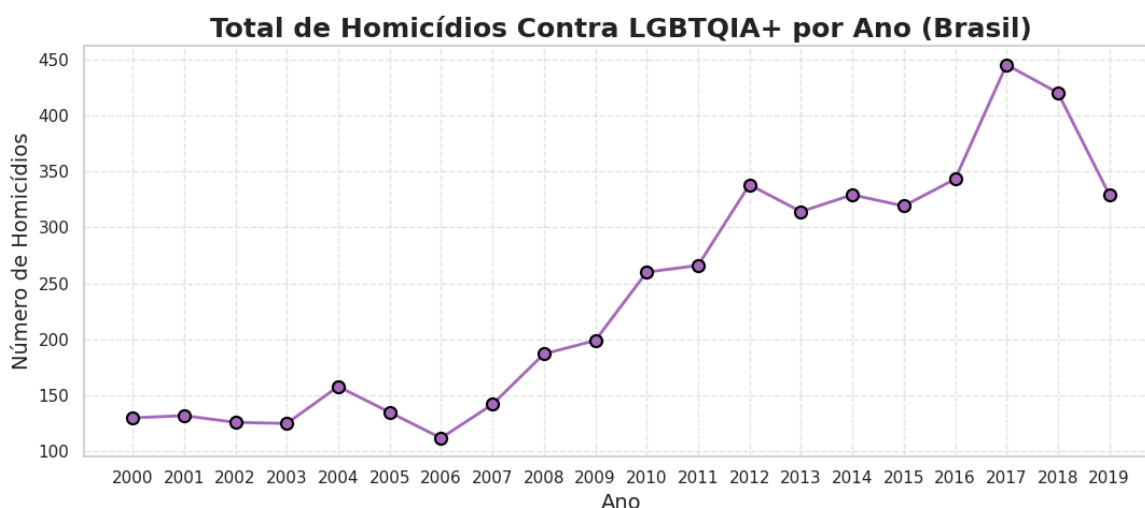
A presença significativa de homicídios em residências sugere que muitos desses crimes ocorrem em contextos íntimos ou familiares, o que pode estar relacionado à rejeição, violência doméstica ou conflitos interpessoais motivados por LGBTfobia. Já a alta incidência em vias públicas e áreas isoladas, como matas e estradas, aponta para a exposição da população LGBTQIA+ à violência urbana e à vulnerabilidade em espaços abertos, muitas vezes sem proteção institucional.

Portanto, embora não possamos identificar especificamente os estados com maior número de homicídios, é possível afirmar que os locais mais perigosos para essa população são aqueles onde há menor vigilância, acolhimento ou presença de políticas públicas de proteção.

- **Há padrões sazonais ou demográficos (idade, raça ou identidade de gênero) nos crimes?**

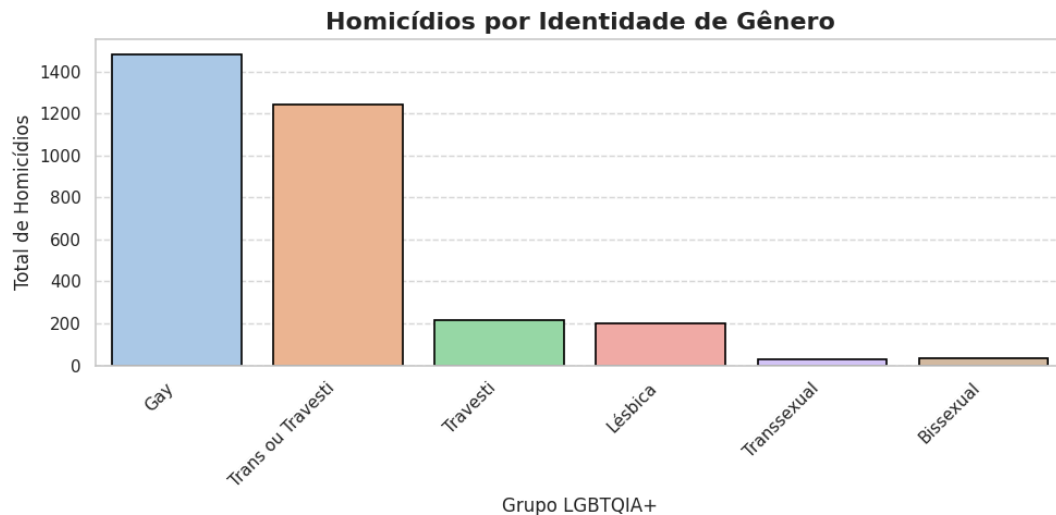
A análise da série histórica de homicídios contra pessoas LGBTQIA+ no Brasil, entre os anos de 2000 e 2019, revela uma tendência preocupante de crescimento ao longo do tempo. Os anos de 2017 e 2018 se destacam como os períodos com os maiores números de homicídios registrados, indicando um pico alarmante de violência nesse intervalo. Esse aumento pode estar relacionado a fatores como o acirramento de discursos de ódio, retrocessos em políticas públicas de proteção e a intensificação da intolerância em contextos políticos polarizados.

Apesar de um declínio observado em 2019, é importante destacar que essa redução não necessariamente representa uma melhora estrutural na segurança da população LGBTQIA+. A queda pode estar associada a fatores como subnotificação, mudanças na metodologia de coleta de dados ou até mesmo à ausência de registros oficiais em determinadas regiões. Portanto, o decréscimo em 2019 deve ser interpretado com cautela, pois não há evidências suficientes para afirmar que houve uma reversão consistente na tendência de violência.



Os dados por raça/cor revelam um resultado que merece atenção: embora o número de vítimas registradas na categoria “branca” seja o mais elevado individualmente, a soma das vítimas pardas e pretas, ou seja, pessoas negras, representa um contingente expressivo da violência letal contra a população LGBTQIA+. Além disso, chama a atenção o número significativo de casos em que a raça/cor da vítima não foi identificada, o que pode indicar falhas na coleta de dados ou negligência institucional na documentação desses crimes. Esse cenário sugere que os dados disponíveis podem não refletir com precisão a realidade da violência, especialmente no que diz respeito à sua dimensão racial. A discrepância entre os registros e a tendência observada em outros estudos, que apontam pessoas negras como as mais vulneráveis, pode estar relacionada a fatores como subnotificação, enviesamento na identificação racial, ou ainda à distribuição geográfica dos

homicídios, com maior incidência em regiões de predominância branca. Ainda assim, é fundamental reconhecer que a violência contra pessoas LGBTQIA+ é um fenômeno estrutural, seletivo e interseccional, exigindo políticas públicas que considerem as múltiplas dimensões da vulnerabilidade, incluindo raça, classe, gênero, território e identidade de gênero.



Conclusões e sugestões de políticas públicas.

Os dados mostram o que a população LGBTQIA+ já sente na pele há muito tempo: viver com medo, seja dentro de casa ou na rua, ainda é uma realidade no Brasil. A violência não é pontual nem isolada – ela é reflexo de uma estrutura que normaliza a exclusão, silencia denúncias e falha em proteger quem mais precisa.

O uso da inteligência artificial para entender padrões e prever riscos é um passo importante, mas ele só tem valor real se virar ação concreta. Dados sozinhos não salvam vidas.

Sugestões de políticas públicas

1. **Criação de centros comunitários LGBTQIA+ com apoio público**, em bairros periféricos e cidades menores, oferecendo acolhimento, atendimento jurídico e psicológico gratuito, além de espaço seguro para quem sofre violência ou rejeição familiar. Esses espaços devem ser geridos com participação da própria comunidade, valorizando saberes locais.
2. **Educação que acolhe e forma**, com políticas de formação continuada para professores(as) sobre gênero e sexualidade, e inclusão de temas LGBTQIA+ nos currículos escolares. Não se trata de “doutrinar”, mas de garantir que jovens LGBTQIA+ possam estudar sem medo – e que seus colegas aprendam a respeitar as diferenças desde cedo.

3. **Formação e sensibilização de profissionais de atuação na ponta** através da implementação de programas de capacitação continuada para profissionais da saúde, educação, segurança pública e justiça, com conteúdos sobre diversidade sexual e de gênero.
4. **Criar um sistema nacional de monitoramento da violência contra pessoas LGBTQIA+**, com recorte por raça/cor, identidade de gênero, orientação sexual e território, com a padronização dos registros em boletins de ocorrência, laudos médicos e certidões de óbito.

Mais do que punir a violência, o desafio é desmontar a lógica que a torna possível. A vida LGBTQIA+ precisa deixar de ser sobrevivência para ser, de fato, vivida com dignidade.