

# O USO DA CIÊNCIA DE DADOS NA ANÁLISE DAS POLÍTICAS PÚBLICAS VOLTADAS À VIOLÊNCIA CONTRA A MULHER

Caio R. Rossini<sup>1</sup> (RA 2002651), Fabio Z. Silva<sup>1</sup> (RA 2015845), Gabriel Parussolo<sup>1</sup> (RA 2005273), Gessyca S. H. Suenaga<sup>1</sup> (RA 2010006), Guilherme P. Batista<sup>1</sup> (RA 2011386), Marcos Lohmann<sup>1</sup> (RA 2006775), Victoria S. I. Silva<sup>1</sup> (RA 2007625), Walter A. Junior<sup>1</sup> (RA 2000081)

<sup>1</sup>Universidade Virtual do Estado de São Paulo – Curso de Ciência de Dados, São Paulo, Brasil

Link para o vídeo: <https://youtu.be/GW9B1DehqaM>

## Resumo

A violência contra a mulher é um problema cada vez mais presente na sociedade, sendo considerado um problema de saúde pública. Mesmo com a existência de políticas públicas e leis de proteção, o Brasil ainda apresenta números alarmantes de violência contra a mulher, incluindo feminicídios. O objetivo deste artigo é identificar o perfil das vítimas e buscar correlações com políticas públicas existentes, através da análise dos dados do SINAN - Sistema de Informação de Agravos de Notificação e do SINESP - Sistema Nacional de Informações de Segurança Pública, Prisionais, de Rastreabilidade de Armas e Munições, de Material Genético, de Digitais e de Drogas. Essas bases são fontes confiáveis extraídas de Dados Abertos e que possibilitam entender a situação do aumento de casos de violência contra a mulher e auxiliar no planejamento de políticas preventivas. Para a análise de dados foram utilizadas ferramentas de exploração e visualização de dados: *Pandas*; *Matplotlib*; *Oracle*; *Seaborn*; *Plotly*; *Apache Spark*. Após análise exploratória, chegou-se à conclusão de que fatores socioeconômicos são os que mais se destacam no momento de perfilar mulheres vítimas de violência doméstica. Adicionalmente não foi possível identificar uma correlação de efetividade com a Lei nº 13.104/2015, referente ao feminicídio.

**Palavras-chaves:** Violência contra mulher. Políticas públicas. Análise de dados. SINAN. SINESP.

## 1. Introdução

A violência contra a mulher é um tema que vem ganhando cada vez mais espaço na sociedade devido ao constante aumento do número de casos, especialmente após o confinamento obrigatório decorrente da pandemia causada pela COVID-19 (MARCOLINO et al., 2021).

Infelizmente não é incomum ligar a televisão ou abrir um portal de notícias e se deparar com histórias de mulheres que foram vítimas de algum tipo de violência, seja física, sexual, patrimonial, psicológica ou moral. Mas não se pode manter a informação no âmbito do fato corriqueiro e normalizado, é preciso apresentar constantemente, com uso de dados, os impactos diretos e indiretos destes casos.

Com o amplo acesso que se tem aos dados em diversas fontes confiáveis, que não são poucos, fazer uso de recursos para facilitar a sumarização, análise e apresentação de forma clara e simplificada sobre as estatísticas da violência contra a mulher, para que se possa tratar a educação desta geração e das próximas sobre o flagelo da violência contra a mulher, já não é sem tempo.

No entanto, a preocupação com esse assunto remonta a década de 70 quando o movimento feminista já reivindicava a implantação de políticas públicas para enfrentamento da questão, o que culminou com a criação da primeira delegacia especializada no atendimento à mulher no Estado de São Paulo em agosto de 1985 (SANTOS, 2005, apud PASINATO e SANTOS, 2008).

Observa-se que essa delegacia foi criada antes mesmo da promulgação da Constituição Federal em 1988 que trouxe de forma expressa a igualdade de direitos e deveres entre homens e mulheres e a previsão de criação de mecanismos para coibir a violência no âmbito das relações familiares.

Posteriormente, o Brasil também ratificou a Convenção Interamericana para Prevenir, Punir e Erradicar a Violência contra a Mulher (Convenção de Belém do Pará), instituiu a Política Nacional de Enfrentamento à Violência contra as Mulheres, promulgou a Lei n. 11.340/2006 (Maria da Penha), alterou o Código Penal para prever penalidade mais grave nos casos em que o homicídio é praticado contra a mulher por razões da condição de sexo feminino (feminicídio), implementada pela Lei nº 13.104/2015, e criou Promotorias de Justiça Especializadas, Juizados de Violência Doméstica e Familiar contra a Mulher e Centros de Referência de Atendimento à Mulher.

Dentre todas essas políticas públicas, merece destaque a Lei Maria da Penha, reconhecida pela Organização das Nações Unidas (ONU) como uma das melhores legislações do mundo no enfrentamento à violência contra as mulheres (BRASIL, 2012), uma vez que ela reconheceu a hipossuficiência e a vulnerabilidade da mulher em contexto de violência doméstica e familiar.

As consequências da violência contra a mulher são nefastas e atingem não somente a própria vítima, como toda a sociedade, gerando impactos na saúde pública, na economia, na assistência social, no sistema judiciário etc.

Outrossim, não se pode olvidar a repercussão dessa violência no desenvolvimento de crianças e adolescentes que por vezes se tornam órfãos, o que motivou, inclusive, a criação do Auxílio Ampara pelo Município de São Paulo que prevê o pagamento de um salário mínimo até completar 18 anos àqueles que perderam a genitora em razão do feminicídio e se encontram em situação de vulnerabilidade social.

Portanto, a erradicação da violência contra a mulher perpassa pela eficácia das políticas públicas, as quais devem ser analisadas a fim de serem aperfeiçoadas e refletirem a real necessidade da sociedade.

### 1.1. Justificativa

O Brasil vem avançando na criação das políticas públicas no combate da violência contra a mulher desde a criação da Lei Maria da Penha, mas ainda mantém números alarmantes sobre o tema. A taxa de feminicídios é a quinta maior do mundo, segundo as NAÇÕES UNIDAS DO BRASIL (2016), e dados do Anuário Brasileiro de Segurança Pública mostram que houve um crescimento de todas as formas de violência contra a mulher, em 2022.

Dada a introdução ao tema da violência, em destaque à violência contra a mulher, é imperativo que se faça saber a toda a comunidade mundial sobre o caráter epidêmico deste verdadeiro flagelo, e é obrigação de quem possui as ferramentas próprias para o saber, que façam uso ilimitado destes recursos para que se possa produzir o efeito educacional sobre as pessoas, hoje, e nas gerações futuras acerca deste mal.

O uso da Ciência de Dados e suas ferramentas de sumarização, análise, visualização de dados, já amplamente utilizados como aceleradores de decisão nos diversos setores da economia, é sem sombra de dúvida, também aplicável para esclarecer e desvendar a sociedade sobre os números da violência; dados não faltam.

Diante desse cenário, o objetivo principal deste trabalho foi analisar os dados fornecidos pelo SINAN (Sistema de Informação de Agravos de Notificação) e pelo SINESP (Sistema Nacional de Informações de Segurança Pública, Prisionais, de Rastreabilidade de Armas e Munições, de Material Genético, de Digitais e de Drogas). Com base na análise dos dados do SINAN é possível traçar o perfil das vítimas e dos agressores. E com a análise de ambos os bancos é possível identificar em que circunstâncias ocorreram essas violações, a fim de verificar os resultados das políticas públicas em vigor, obtendo a partir de uma análise exploratória alguns *insights* acionáveis que podem ajudar na tomada de decisão pelo poder público.

Além disso, outras indagações surgiram durante a discussão dessa temática, como o motivo do aumento de registros/denúncias de agressões. Esse aumento pode estar relacionado ao apoio/incentivo que atualmente as mulheres têm para formalizar a denúncia.

### 1.2. Tema

O tema deste trabalho é “O uso da ciência de dados na análise e extração de conhecimento com base nos dados do SINAN e do SINESP e nas políticas públicas voltadas à violência contra a mulher”.

É consenso que a informação é a melhor ferramenta que se pode ter contra a ignorância sobre os fatos, e assim surge o interesse na construção de um artigo em que se empregam técnicas de ciência de dados para produção de informação sumarizada, analisada, tratada e exposta de forma que se possa ter simples entendimento a qualquer leitor, dos dados existentes sobre a violência contra as mulheres, para que não mais se passe despercebida a sua importância.

O interesse pelo tema surgiu pelo fato de que a popular *Lei Maria da Penha* (Lei 11.340/2006) teve sua criação motivada por uma condenação do Brasil imposta pela Comissão Interamericana de Direitos Humanos da Organização dos Estados Americanos (OEA). Esse fato chamou a atenção do grupo para a ineficiência de medidas de combate à violência contra a mulher.

A partir de 2011, com a Portaria nº 104 publicada em 25 de janeiro, os serviços de saúde, sejam eles públicos ou privados, em todo o Brasil, passaram a ter a obrigação de notificar casos de violência doméstica, sexual e outras formas de violência. Como resultado, o grupo decidiu estudar esse tema muito preocupante e relevante em nosso país, buscando entender melhor algumas questões e quantificá-las, na medida do possível, utilizando os dados do SINAN e do SINESP.

### 1.3. Objetivos

#### 1.3.1. Objetivo Geral

Analisar os dados abertos, as políticas nacionais e o desfecho das políticas públicas no combate à violência contra a mulher.

#### 1.3.2. Objetivos específicos

1. Levantar as políticas públicas e ações de enfrentamento e combate à violência contra a mulher criadas no Brasil;
2. Extração, transformação e carregamento (ETL) das bases do SINAN e SINESP, utilizando a linguagem *Python* com *frameworks Pandas, Apache Spark, Matplotlib, Seaborn* e armazenamento em banco de dados *Oracle*;
3. Identificar o perfil das mulheres que sofreram algum tipo de violência no período de 2009 a 2021 por meio de técnicas de análise de dados (análise exploratória, filtros, agregações e correlações);
4. Investigar através dos dados abertos se uma das políticas levantadas pode ter seu resultado avaliado.

## 2. Metodologia

### 2.1. Método de pesquisa e procedimento de coleta de dados

O método de pesquisa e o procedimento de coleta de dados desempenham um papel fundamental na produção de conhecimento científico confiável e na consecução dos objetivos de qualquer estudo. Para o desenvolvimento deste artigo foram utilizados dois métodos de pesquisa: qualitativo e quantitativo. Para a abordagem qualitativa foi realizada uma entrevista em profundidade. E para a parte quantitativa foi realizada uma exploração de dados estruturados por meio da linguagem *Python* e suas bibliotecas, produzindo o conhecimento para avaliar o resultado de políticas públicas já implementadas no combate à violência contra a mulher. Portanto, a produção deste material trata-se de um método de pesquisa descritiva.

Para que a delimitação do problema ocorresse, houve a consulta de fontes oficiais, publicações sobre o tema e entrevista com profissionais da área. As publicações tiveram o papel de contextualizar o grupo nos diversos desafios presentes neste tema e as dificuldades existentes em combater esse mal que atinge a sociedade. Durante a entrevista foram identificados problemas que os departamentos responsáveis enfrentam na implementação de políticas eficazes. Esses problemas apresentados auxiliarão na definição de quais ações serão trabalhadas durante este artigo.

Após consultas em órgãos responsáveis, matérias e reportagens já publicadas, o grupo identificou duas bases que serão utilizadas como principais fontes de conhecimento para desenvolver as análises: o banco de dados do DATASUS, especificamente o Sistema de Informação de Agravos de Notificação (SINAN) e o banco de dados do Ministério da Justiça e Segurança Pública, o Sistema Nacional de Informações de Segurança Pública, Prisionais, de Rastreabilidade de Armas e Munições, de Material Genético, de Digitais e de Drogas (SINESP).

O SINAN é um banco de dados alimentado principalmente por meio da notificação e investigação de casos de doenças e agravos de notificação compulsória, conforme estabelecido pela legislação (Portaria de Consolidação nº 4, de 28 de setembro de 2017, Anexo).

O abastecimento deste banco de dados possibilita um diagnóstico dinâmico da ocorrência de eventos na população. Isso fornece informações valiosas para compreender as causas subjacentes aos agravos notificados e pode indicar os riscos enfrentados pela população. Dessa forma, o Sinan desempenha um papel crucial na identificação da situação epidemiológica de uma área geográfica específica. É também uma ferramenta relevante que auxilia no planejamento de políticas de saúde, na definição de prioridades de intervenção e na avaliação do impacto das medidas implementadas. Ao utilizar o SINAN, é possível obter uma visão abrangente e atualizada da situação epidemiológica, o que é essencial para promover a saúde e o bem-estar da população. Para realizar as análises, o grupo utilizará dentro dessa fonte de dados os registros relacionados à Violência Interpessoal/Autoprovocada.

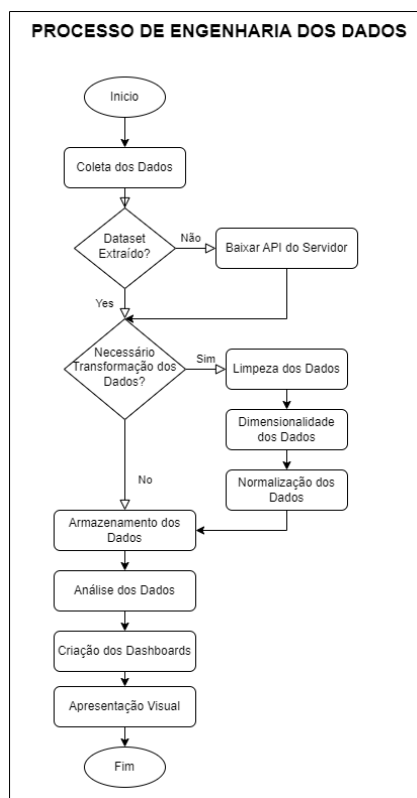
Já o SINESP é uma plataforma de informações integradas, que possibilita consultas operacionais, investigativas e estratégicas sobre segurança pública, implementado em parceria com os entes federados. Ele foi criado através pela Lei 12.681, de 04 de julho de 2012 e, com a publicação da Lei 13.675, de 11 de junho de 2018, o Sinesp firmou-se como um dos meios e instrumentos para a implementação da Política Nacional de Segurança Pública e Defesa Social (PNSPDS) instituindo-se o Sistema Único de Segurança Pública (Susp).

Ambas as fontes foram apresentadas aos profissionais do setor tendo sua confiabilidade validada, possibilitando a exploração das bases para o desenvolvimento da pesquisa.

Visando o atendimento aos objetivos propostos por este artigo, os seguintes passos foram seguidos:

1. Acesso aos dados das bases
2. Seleção de Variáveis
3. Filtragem e Limpeza
4. Análise Descritiva
5. Análise Estatística
6. Interpretação e Discussão

Adicionalmente, foram enriquecidas as informações com base na metodologia estruturada pelo fluxo de processos a seguir:



1. Fonte: Próprio autor (2023).

Nas sessões seguintes serão detalhados os métodos utilizados para o cumprimento dos tópicos expostos.

## 2.2. Acesso aos dados e seleção de variáveis

### 2.2.1. SINAN

É possível acessar os dados provenientes do SINAN de diversas formas, porém para a elaboração deste estudo, preferiu-se o acesso direto por meio da biblioteca *Python* disponibilizada pelo SUS (Sistema Único de Saúde). Para isso, foi necessário instalar a biblioteca *PySUS* por meio do comando `pip` e importar as funções “*download*” e “*parquets\_to\_dataframe*”, além do sistema SINAN.

Com as instalações e importações concluídas, foi possível extrair dados de 2009 a 2021 resultando em um *dataframe* de 3.035.965 linhas e 160 colunas, levantando a necessidade da seleção das variáveis que serão úteis para o trabalho em questão.

Foram excluídas basicamente colunas que competem apenas a informações de sistema para o cadastro das ocorrências e informações duplicadas, mantendo os dados referentes às características censitárias e atributos da violência sofridas, restando 100 colunas para compreender melhor as informações necessárias.

### 2.2.2. SINESP

Por meio do portal “Dados Abertos” do Governo Federal, o Ministério da Justiça e Segurança Pública (MJSP) disponibiliza publicamente a base sumarizada em formato XLSX de Ocorrências Criminais do período de 2015 a 2023 provida pelo Sistema Nacional de Informações de Segurança Pública (SINESP) agrupadas por município e por unidade federativa.

Tratando-se de uma fonte de dados previamente sumarizada, não foi necessário fazer redução de dimensionalidade ou higienização, restando o trabalho em filtros e agrupamento dos registros.

## 2.3. Limpeza e tratamento de dados

### 2.3.1. SINAN

A fim de garantir a qualidade dos dados para as etapas de análise subsequente, os dados extraídos do SINAN foram submetidos a uma limpeza e tratamento.

Utilizando como base o dicionário de dados disponibilizado pelos mantenedores da base, a princípio os dados foram filtrados excluindo valores duplicados, por meio da coluna “NDUPLIC” já constante na tabela. Em seguida, foram mantidas apenas as linhas referentes a vítimas do sexo feminino e deletadas as ocorrências derivadas de lesões autoprovocadas. Logo após, as colunas que não estão presentes no dicionário de dados ou foram utilizadas para filtrar o *dataframe*, foram excluídas. Os demais tratamentos realizados foram os seguintes:

- Transformação de campos vazios em nulos;
- Substituição de valores numéricos de acordo com o dicionário de dados, como por exemplo colunas que aceitam respostas como “Sim”, “Não”, “Não aplicável” e “Ignorado”, mas que recebiam valores numéricos na base bruta, isso se aplica também a outros tipos de dados categóricos;
- Conversão de dados em tipo *datetime*;
- Substituição de valores numéricos referentes ao estado onde ocorreu o atendimento e a profissão da vítima de acordo com lista presente no IBGE (Instituto Brasileiro de Geografia e Estatística) e a lista do CBO (Classificação Brasileira de Ocupações);
- Substituição de código referente a idade de acordo com dicionário de dados

Esta etapa permitiu que 1.700.365 de registros e 92 colunas fossem capturadas e as primeiras análises já fossem realizadas.

### 2.3.2. SINESP

Conforme apresentado no item 2.1.2, a base de dados fornecida pelo MJSP tem a função de informar, de forma sumarizada, dados de Ocorrências Criminais agrupados ano, mês, uf, município, tipo-de-crime e sexo-da-vítima, de forma que esta base de dados não demandou higienização ou redução de dimensionalidade, restando o trabalho de filtragem e agrupamentos para extração de visualizações.

## 2.4. Análise dos dados e validação da hipótese

Com os dados estruturados e de fácil acesso, análises aprofundadas foram realizadas visando entender ainda mais o contexto em que está sendo elaborada a pesquisa, bem como a validação da hipótese de que a política pública que criou o Feminicídio, promulgada pela Lei nº 13.104/2015, obteve um resultado positivo.

Nesse contexto, foram realizadas análises descritivas das principais variáveis dentro de nossa base que caracterizam um conjunto populacional, utilizando o DATASUS para identificação do perfil das vítimas de violência e agregando o de Segurança Pública para refinamento do perfil da agressão sofrida.

Após isso foi realizada uma análise estatística aplicando-se a matriz densa de correlação entre atributos através do cálculo de correlação de Pearson, observando-se os atributos que qualificam as ocorrências por causa e efeito, tais como a relação da vítima com o agressor, o uso de força, arma ou recursos imateriais como agressão moral, bem como a influência do álcool, a idade da vítima ou o emprego de violência sexual.

Em vista da hipótese preliminar de se analisar a correlação entre as políticas públicas e as ocorrências, observa-se, pela leitura dos dados, que as políticas públicas em geral buscam tratar os efeitos da violência, por exemplo, na segmentação do tipo de crime e de seus agravantes com intuito de realizar um julgamento mais técnico e penalizar os

autores de forma consistente, porém não trata as causas da criminalidade, tais como questões relacionadas à família, educação e saúde mental.

### 2.4.1 Matriz Densa e Correlação de Pearson

A correlação de Pearson é uma medida estatística que avalia a força e direção da relação linear entre duas variáveis. O indicador tem variação entre -1 e 1, onde 1 representa uma correlação positiva perfeita, -1 uma correlação negativa perfeita e 0 nenhuma correlação.

A matriz densa de correlação entre atributos é uma tabela que mostra as correlações de Pearson entre todas as combinações de atributos. É uma ferramenta visual útil para analisar como diferentes variáveis se relacionam entre si.

Para a análise de causa e efeito, utilizamos a matriz de correlação para identificar quais atributos têm forte correlação com um determinado objetivo. Aqueles indicadores com correlações mais altas podem ser considerados como tendo um potencial efeito maior no objetivo, podendo ser mais influentes ou relevantes na sua determinação, no entanto, a correlação não implica causalidade direta, apenas uma relação estatística entre variáveis, portanto, é importante realizar análises mais aprofundadas para estabelecer relações de causa e efeito entre variáveis, de modo que a interpretação deve-se levar em consideração o contexto dos dados.

## 2.5. Ferramentas e tecnologias em nuvem

Pensando em emergir grande massa de dados para um centro unificado, foram selecionadas soluções que tornassem as aplicações e serviços mais acessíveis ao projeto. Tendo este conceito tornamos aplicáveis diretamente em nuvem sem utilização prévia de infraestrutura ou, como a centralização de homogeneidade entre serviços, arquitetura chamado *data centers* (PEDROSA e NOGUEIRA, 2011, p.1). Para o presente projeto foi elaborada a estruturação de serviços e aplicações diretamente em nuvem para ser um facilitador no momento da modelagem de dados.

Visando alta disponibilidade, foi utilizada computação em nuvem como uma forma de uso da infraestrutura de um fornecedor para implantar a sua solução (ORACLE, 2023). A ideia foi conseguir direcionar esforços da solução diretamente na ideia sintetizada, retirando o escopo de sustentação dos serviços e manutenção da Tecnologia da Informação e Comunicação (TIC). Com isso, foi possível maximizar o planejamento da modelagem de dados através de sua estruturação e uso para aplicação no caso de uso.

Em linhas gerais, os provedores de computação em nuvem fornecem três modelos de solução para as aplicações:

- IaaS: Em *Infrastructure as a Service* o consumidor é responsável por instalar, configurar, monitorar e gerenciar os serviços contratados pelo provedor, sendo que eles são acessados pela internet sob demanda (ORACLE, 2023);
- PaaS: Em *Platform as a Service* o consumidor usufrui dos serviços de desenvolvedores e seus apetrechos em conjunto com APIs e *frameworks* essenciais para suas atividades. Esta característica de serviço determina as responsabilidades do consumidor para os conceitos mencionados e o provedor se responsabiliza pela solução IaaS;
- SaaS: Em *Software as a Service* o consumidor se encaixa na ideia de centralizar seus recursos na linha de negócio, utilizando todos os serviços mencionados em PaaS pela responsabilidade do provedor sendo uso, exclusivamente, de suas aplicações e serviços com a infraestrutura completamente preparada. O Projeto presente utiliza essa solução por ser ideal para a proposta e ser a solução mais inovadora disponível (ORACLE, 2023).

A decisão de uso desta solução reflete na proposta de fornecimento de algo adaptável a dados que são altamente mutáveis, sendo que, eles podem determinar diversas informações dependendo da necessidade. Com isto, tem-se um ambiente ideal para a capacidade de personalização dos recursos da nuvem (PEDROSA e NOGUEIRA, 2011, p.2).

### 2.5.1. Benefícios do Cloud Computing

Nesta modalidade, o principal objetivo é a alta disponibilidade, pois faz uso da internet e é propagada em qualquer localidade, contando que tenha acesso à rede. Com isso, foram listadas as principais vantagens do uso:

- Velocidade – Provisionamento de recursos e serviços rápido para desenvolvimento e teste (ORACLE, 2023);
- Escalabilidade – Capacidade de estender/expandir a performance dos recursos sem a perda do desempenho (DIAN e PAULA, 2021, p.06);

- Alta Disponibilidade – Disponibilização e armazenamento dos dados de forma que sejam consistentes mesmo em condições adversas, como desastres naturais, mau funcionamento de hardware e afins (DIAN e PAULA, 2021, p.06);
- Confiabilidade – Sistemas distribuídos tolerantes a falhas e escaláveis (ORACLE, 2023).

Pensamos em inovar a solução implantando em nuvem para maximizar o tratamento dos dados para consumo ainda em nuvem. Foi adotada a solução em nuvem primeiramente pelo baixo custo, já pensando em minimizar os gastos com a infraestrutura e a abstração, no qual permite a facilidade pelo uso do acesso pela internet aos usuários que podem consumir diretamente e/ou indiretamente os recursos em nuvem (MACHADO, MOREIRA e SOUSA, 2010).

### 2.5.2. Sistema Gerenciador de Banco de Dados (SGBD)

É notado a cada ano um aumento de volume estrondoso de dados que são consumidos e gerados diariamente devido a alta competição no mercado, onde que jaz a importância de gerenciar essa massa de forma eficiente. Segundo Medeiros (2013), o armazenamento e processamento adequado de informação é atribuído um papel fundamental no âmbito de um sistema de informação. Entende-se como um Sistema Gerenciador de Banco de Dados como um recurso obrigatório para futuras sustentações de dados e onde serão armazenadas as informações que serão consumidas pelo projeto, permitindo aos usuários criar e manter um banco de dados (AMADEU, 2014).

Pensando no contexto atual, a proposta do banco de dados é trazer uma imagem de abstração do mundo real em formato de objetos (MEDEIROS, 2013), realizando a conjunção de áreas responsáveis que são pertinentes as análises de levantamento sobre os dados e suas assimilações com as políticas envolvidas.

### 2.5.3. Banco de Dados Oracle Autonomous

Foi utilizado um banco de dados relacional da empresa ORACLE, que oferece as funcionalidades propostas por Amadeu (2014) em possibilitar consultas e recuperar dados específicos, atualizações para refletir mudanças no minimundo e geração de relatórios com base nos dados. Aliando essa prática com uso de ferramentas atuais, a implantação do *Autonomous* reflete na construção de um banco de dados robusto com condicionamento do *Cloud Computing*, oferecendo a dinâmica de independência de complexidade, escalabilidade ou criticidade (ORACLE, 2023).

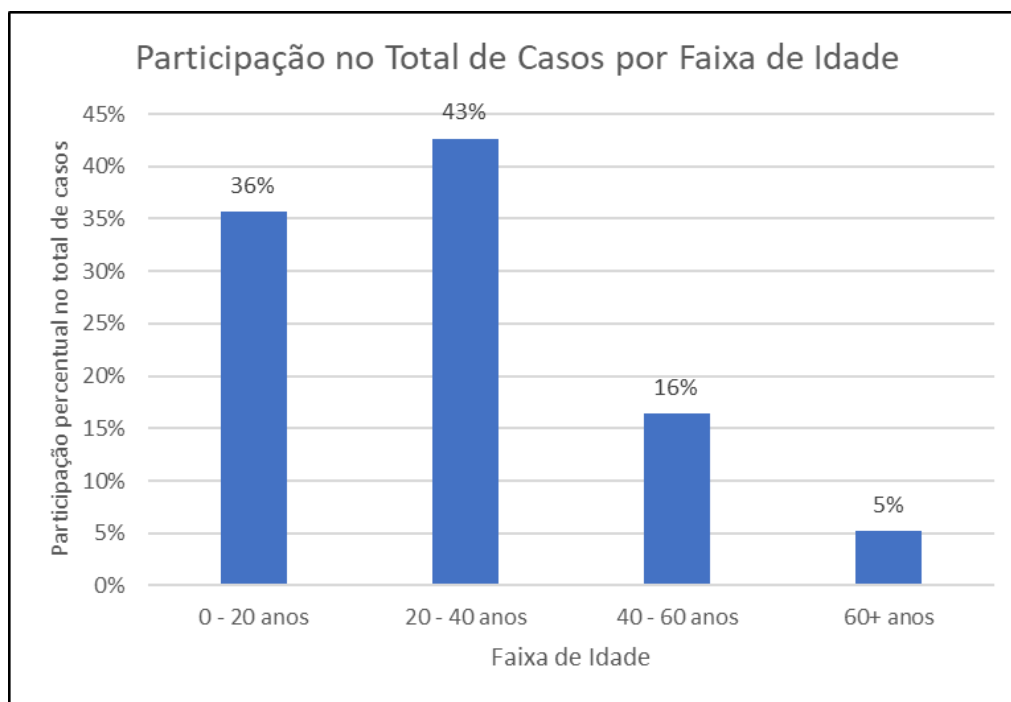
Aprofundando a conceituação de uso, entendemos que a grande quantidade de dados requer a infraestrutura de bancos de dados relacionais abrangentes, sendo implantado para esta proposta o uso do produto *Autonomous Data Warehouse* da empresa ORACLE. Permite a criação da massa de dados atribuída com diversas ferramentas de automatização e modelagem de dados, contendo os principais *frameworks* atuais que permite a descoberta e uso de dados de forma rápida e fácil independente do tipo e tamanho do dado (ORACLE, 2023), de modo a viabilizar a construção de *Dashboards* e a exploração de novos formatos de interação com a base.

## 3. Resultados e discussão

### 3.1. Perfil das mulheres que sofreram algum tipo de violência

O perfil da mulher que enfrenta violência no Brasil é multifacetado, refletindo uma triste realidade enraizada em diversas camadas da sociedade. Segundo dados do SINAN de 2009 a 2019, mais de 15 mulheres sofrem alguma violência por hora no Brasil.

Mulheres de todas as idades estão sujeitas a ser vítimas, 36% dos casos são de crianças ou mulheres de até 20 anos. Em 43% dos casos a vítima é jovem, na faixa etária de 20 a 40 anos, uma fase em que muitas mulheres estão na idade reprodutiva e, por muitas vezes, envolvidas em relacionamentos íntimos. Além disso, adolescentes também estão vulneráveis, enfrentando formas específicas de violência, como abuso no namoro. 16% dos casos têm como vítima mulheres de 40 a 60 anos; idosas a partir dos 60 anos representam 5% dos registros.



2. Fonte: Próprio autor (2023).

Em um país tão vasto e diverso como o Brasil, a localização geográfica também desempenha certo papel. Regiões urbanas densamente povoadas podem apresentar desafios únicos em comparação com áreas rurais, onde o acesso a serviços pode ser mais limitado. É possível ver que alguns estados se destacam negativamente nos índices, com destaque para o estado do Mato Grosso do Sul que possui o maior registro de casos por 100 mil habitantes em toda a série histórica, mas isso não interfere no fato de a violência contra a mulher ser um problema em todo território nacional.

A classe social não é uma barreira para a violência de gênero, afetando mulheres de todas as condições econômicas. Contudo, mulheres em situação de vulnerabilidade econômica podem ter menos recursos para buscar ajuda e apoio. Mas essa não é uma regra, mulheres adultas com o ensino médio concluído representam 24% das vítimas de violência.

Mulheres pardas, negras e indígenas também enfrentam uma interseção de desigualdades, experimentando níveis mais elevados de violência e obstáculos adicionais para acessar serviços de apoio, elas foram vítimas em mais de 50% dos casos no ano de 2019.

Um dos fatores que também preocupa autoridades é o local em que o crime ocorre, sendo em 64% dos casos o próprio local onde a vítima reside, seguido pela via pública com 15%. Isso evidencia que a maior parte dessa violência acontece fora dos olhos da segurança pública e exige que a vítima se disponha a denunciar, mesmo sob risco de retaliação por parte do agressor.

Por meio dos estudos realizados sobre a influência de fatores circunstanciais mais presentes nas ocorrências de violência contra a mulher, sob a ótica dos registros públicos classificados especificamente como “violência motivada por sexismo” (onde a vítima é agredida pelo simples fato de ser mulher), observam-se alguns padrões que podem ser utilizados para a tratativa da causa da violência. Estes padrões são observados por meio do cruzamento quantitativo entre cada aspecto da ocorrência com todos os demais aspectos, analisando um-a-um, a sua relação com os demais, e assim compreendendo o quanto cada um deles influencia no outro. A matriz de correlação de Pearson nos revela tais fatores de influência circunstancial através de indicadores numéricos destacados no gráfico a seguir, pela intensidade das cores “azul” e “vermelho”, onde o azul representa a influência oposta e o vermelho representa a influência favorável entre os fatores. A grosso modo, podem-se enxergar como “zonas frias” e “zonas quentes” de influência.



Fonte: Próprio autor (2023).

[illegible]

Fonte: Próprio autor (2023).

O recorte apresentado na figura 4 demonstra a influência de fatores acerca do perfil do agressor e das características da ocorrência em relação ao tipo de agressão. Observa-se uma correlação óbvia entre a violência física e o emprego de força, além de indicadores importantes como a questão do agressor ser o cônjuge nos casos recorrentes de violência física.

	Agressos é cônjuge	Agressor é ex-cônjuge	Agressor é namorado	Agressor é ex-namorado	Agressor é filho	Agressor desconhecido	Influência de Álcool	Ocorrência na residência	Vítima Homossexual	Vítima adulta	Vítima jovem	Vítima idosa
<b>Vítima casada</b>	0,6	-0,1	-0,1	-0,1	0,0	-0,1	0,1	0,2	0,3	0,1	0,0	0,0
<b>Vítima viúva</b>	0,0	0,0	0,0	0,0	0,2	0,0	0,0	0,0	0,0	0,0	0,0	0,0
<b>Vítima solteira</b>	-0,4	0,0	0,2	0,2	0,0	0,2	0,0	-0,1	0,1	-0,1	0,1	-0,1
<b>Vítima divorciada</b>	-0,1	0,4	0,0	0,0	0,0	0,0	0,0	0,0	0,1	0,1	0,0	0,0

5. Zonas quentes de correlação de perfil de agressor vs. perfil de vítima II

Fonte: Próprio autor (2023).

No recorte da figura 5 destaca-se a influência de características do agressor, a influência de álcool e o perfil da vítima, em relação ao estado civil da vítima. Novamente se vê uma correlação inversa óbvia entre vítima solteira e agressor cônjuge, mas evidencia a correlação forte entre o agressor cônjuge e a vítima casada ou do ex-cônjuge com a vítima divorciada, assim como o namorado ou o ex-namorado em relação a vítima solteira.

	Pornografia	Agressor é o pai	Agressor é o padrasto	Agressor desconhecido	Agressor é conhecido	Agressão em grupo	Vítima solteira	Vítima bissexual	Vítima adolescente
<b>Assédio Sexual</b>	0,2	0,1	0,2	0,1	0,2	0,0	0,1	0,0	0,1
<b>Estupro</b>	0,1	0,1	0,2	0,4	0,3	0,1	0,2	0,1	0,2
<b>Exploração sexual</b>	0,2	0,0	0,0	0,0	0,1	0,0	0,0	0,0	0,0

6. Zonas quentes de correlação de perfil de agressor e vítima em crimes sexuais

Fonte: Próprio autor (2023).

Existem indicadores de correlação óbvios, como a relação entre “violência sexual”, “assédio sexual” e “estupro”, porém observam-se outros fatores que influenciam diretamente (zona quente; positiva) na violência sexual: a figura paterna (pai ou padrasto), infelizmente, o que reforça a questão de políticas ligadas à família.

	Violência recorrente	Violência física	Violência Psicológica	Emprego de força	Enforcamento	Uso de objeto contundente	Uso de facas	Agressos é cônjuge	Ocorrência em bar	Vítima casada	Vítima Homossexual	Vítima adulta
Influência de Álcool	0,1	0,2	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1

7. Zonas quentes de correlação entre o uso de álcool pelo agressor e fatores agravantes  
 Fonte: Próprio autor (2023).

Outros fatores sociais, tais como o uso de bebida alcoólica, estão relacionados a fatores de amplitude de perfil de vítimas, observando-se a influência em ocorrências onde se qualificam em maior ou menor grau: Violência recorrente, Violência física, Violência Psicológica, Tortura, Violência financeira, Emprego de força, Enforcamento, Uso de objeto contundente, Uso de facas, Provocar queimaduras, Envenenamento, Ameaça, Agressor é cônjuge, Agressor é ex-cônjuge, Agressor é filho, Agressão em grupo, Vítima transgênero, Ocorrência em abrigo público, Ocorrência em bar, Ocorrência na residência, Ocorrência em local/via pública, Vítima preta, Vítima indígena, Vítima asiática, Vítima parda, Vítima com escolaridade fundamental incompleta, Vítima com escolaridade fundamental completa, Vítima com escolaridade básica completa, Vítima com escolaridade média incompleta, Vítima com escolaridade média completa, Vítima com escolaridade superior incompleta, Vítima analfabeta, Vítima casada, Vítima viúva, Vítima divorciada, Vítima heterossexual, Vítima Homossexual, Vítima bissexual, Vítima adulta, Vítima jovem. O amplo espectro de influência do alcoolismo sobre a violência contra a mulher reforça a necessidade de atenção a políticas ligadas à família e à saúde.

A tabela a seguir apresenta os principais aspectos presentes nas ocorrências registradas nas bases de dados analisadas:

FATOR OBSERVADO	FATOR CONTRIBUINTE (Potencial de influência %)
Agressor colega de trabalho	Agressor é patrão ou chefe (68%),
Agressor conhecido	Violência sexual (66%), Estupro (64%),
Agressor cônjuge	Violência Física (65%), Emprego de força física (60%), Ocorreu na residência (62%), Vítima casada (81%),
Agressor desconhecido	Violência sexual (67%), Uso de arma de fogo (60%), Estupro (68%), Ocorreu em local público (66%),
Agressor ex-cônjuge	Vítima divorciada (68%),
Agressor padrasto	Violência sexual (61%),
Ameaça	Ocorrência recorrente (61%), Violência Psicológica (72%),
Emprego de força física	Violência Física (84%), Agressor cônjuge (60%),
Exploração sexual	Pornografia (61%),
Ocorreu na residência	Ocorrência recorrente (62%), Agressor cônjuge (62%),
Ocorrência recorrente	Violência Psicológica (62%), Ameaça (61%), Ocorreu na residência (62%),
Violência Física	Emprego de força física (84%), Agressor cônjuge (65%), Vítima casada (61%),
Violência Psicológica	Ocorrência recorrente (62%), Ameaça (72%),

Violência sexual	Assédio sexual (72%), Estupro (92%), Agressor padrasto (61%), Agressor desconhecido (67%), Agressor conhecido (66%),
Vítima casada	Violência Física (61%), Agressor cônjuge (81%), Vítima Homossexual (63%),
Vítima Homossexual	Vítima casada (63%),
Vítima negligenciada	Agressor é mãe (63%),

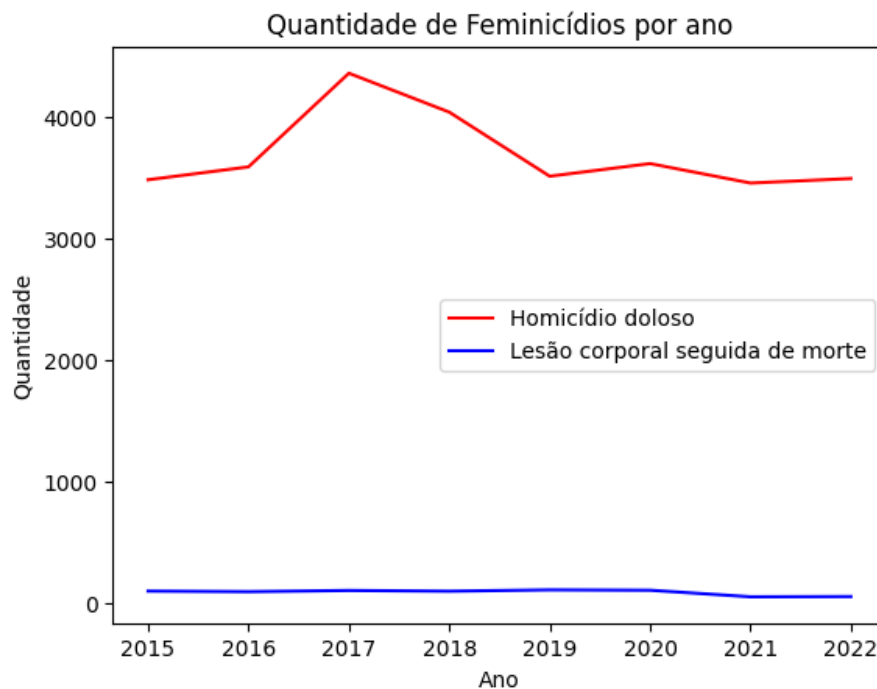
8. Tabela de indicadores de influência mais relevantes no agravante e aumento de casos  
Fonte: Próprio autor (2023).

O estudo completo está disponível no repositório Github (<https://t.ly/mPgqA>) e pode ser utilizado como ponto de partida para outros fins exploratórios e analíticos sobre esta base de dados.

Enfrentar a violência de gênero requer uma abordagem abrangente que leve em consideração todos esses fatores. É imperativo continuar promovendo a educação, conscientização e políticas públicas eficazes para criar uma sociedade onde todas as mulheres possam viver livres do medo da violência. O Brasil, como tantos outros países, precisa unir esforços para construir um futuro mais seguro e igualitário para todas as suas cidadãs.

### 3.2. Avaliação da política pública

Mesmo com a Lei nº 13.104/2015, que implementou uma penalidade mais grave nos casos em que o homicídio é praticado contra a mulher por razões da condição de sexo feminino, não foi possível observar uma redução na quantidade de feminicídios, conforme o gráfico abaixo, que relaciona a quantidade de feminicídios (Homicídio doloso e lesão corporal seguida de morte) após o ano de 2015, no Brasil.



9. Fonte: Próprio autor (2023).



#### 4. Conclusão

O uso da ciência de dados na análise das políticas públicas voltadas à violência contra a mulher pode ser uma ferramenta importante para a compreensão e quantificação desse problema.

Este estudo visou compreender as estratégias de combate à violência contra as mulheres, através da análise do perfil das vítimas e sua relação com as políticas públicas existentes. Para realizar essa análise, foram utilizados dados provenientes do SINAN e do SINESP.

Com base nas políticas públicas estudadas durante a elaboração deste artigo, o grupo teve a oportunidade de conhecer não apenas as políticas utilizadas na hipótese preliminar a avaliar, como também as outras criadas desde o ano de 2009 até o ano de 2023 todas elas pensadas pelo poder público com o objetivo de impactar de alguma forma na corrente dos casos de violência contra a mulher, defendendo a hipótese de sua redução.

Observa-se, porém, que as políticas públicas visam a judicialização dos casos, quando eles já ocorreram, bem como o processo legal decorrente, com ferramentas legais para a criminalização e punição mais severa dos agressores. No entanto, os dados da saúde e da segurança pública não apresentam redução na quantidade de casos registrados, o que leva à conclusão de que tais políticas não impactam na causa, mas no efeito, quando a justiça se faz pelo devido processo legal.

Ao analisar os elementos qualificadores dos casos de violência, pela leitura das características de cada caso, observam-se os fatores que mais influenciam na causa do cenário de violência, e estes são fatores socioeconômicos já conhecidos de todos os brasileiros: saúde, educação, dignidade e família. Ainda sobre os dados, a análise realizada não foi capaz de revelar uma correlação de efetividade da política pública (Lei nº 13.104/2015), desta forma, os dados evidenciam a hipótese sustentada pela opinião pública, de que as políticas mais efetivas são aquelas que fortalecem as bases da comunidade, e não apenas aquelas que trabalham na correção dos desvios de conduta de uns. Acredita-se que, enquanto as políticas públicas não forem direcionadas à causa específica da violência contra a mulher, o impacto no combate e prevenção desse tipo de violência tende a ser insignificante. Essa suposição é respaldada por dados estatísticos que indicam um crescimento contínuo da violência contra as mulheres, mesmo diante da implementação de tais políticas.

## Referências

AMADEU, Claudia Vicci (org.). **Banco de dados**. São Paulo: Pearson, 2014. E-book. Disponível em: <https://plataforma.bvirtual.com.br>. Acesso em: 18 out. 2023.

BRASIL. Ministério da Justiça e Segurança Pública. **O SINESP**. Disponível em: <<https://www.gov.br/mj/pt-br/assuntos/sua-seguranca/seguranca-publica/sinesp-1/>>. Acesso em: 27 out 2023.

\_\_\_\_\_. Secretaria de Políticas para as Mulheres. **Conheça a lei que protege as mulheres da violência doméstica e familiar: Lei Maria da Penha Lei nº 11.340/2006**. Brasília. 2012. Disponível em: <<https://assets-compromissoeatitude-igp.sfo2.digitaloceanspaces.com/2012/08/SPM-publicacao-Lei-Maria-da-Penha-edicao-2012.pdf>>. Acesso em: 09 set. 2023.

BUENO, Samira; MARTINS, Juliana; LAGRECA, Amanda; SOBRAL, Isabela; BARROS, Betina; BRANDÃO, Juliana. **O crescimento de todas as formas de violência contra a mulher em 2022**. In: FÓRUM BRASILEIRO DE SEGURANÇA PÚBLICA. 17º Anuário Brasileiro de Segurança Pública. São Paulo: Fórum Brasileiro de Segurança Pública, p. 136-145, 2023. Disponível em: <https://forumseguranca.org.br/wp-content/uploads/2023/07/anuario-2023.pdf>. Acesso em: 15 set. 2023.

DIAN, Maurício; PAULA, Laís. **COMPUTAÇÃO EM NUVEM: os desafios das empresas ao migrar para a nuvem**. *Interface Tecnológica*, v.18, n. 2, p.12, dez.2021. Disponível em: [https://www.academia.edu/download/4660345/computacao\\_nuvem.pdf](https://www.academia.edu/download/4660345/computacao_nuvem.pdf). Acesso em 16 out. 2023.

MACHADO, Javam; MOREIRA, Leonardo; SOUSA, Flávio; **Computação em Nuvem: Conceitos, Tecnologias, Aplicações e Desafios**. ERCEMAPI, p.26, set.2010. Disponível em: <https://www.ic.unicamp.br/~ducatte/mo401/1s2011/T2/Artigos/G04-095352-120531-t2.pdf>. Acesso em 16 out. 2023.

MARCOLINO, E. DE C. et al. O distanciamento social em tempos de Covid-19: uma análise de seus rebatimentos em torno da violência doméstica. *Interface - Comunicação, Saúde, Educação*, v. 25, p. e200363, 2021. Disponível em: <<https://www.scielo.br/j/icse/a/jCBm4rn9TSXDzrZQ5JXKnNL/#>>. Acesso em: 09 set. 2023.

MEDEIROS, Luciano Frontino de. **Banco de dados: princípios e prática**. 1. ed. Curitiba: Intersaberes, 2013. E-book. Disponível em: <https://plataforma.bvirtual.com.br>. Acesso em: 18 out. 2023.

NAÇÕES UNIDAS DO BRASIL. **ONU: Taxa de feminicídios no Brasil é quinta maior do mundo; diretrizes nacionais buscam solução**. Brasília, 09 abril 2016. Disponível em: <<https://brasil.un.org/pt-br/72703-onu-taxa-de-feminic%C3%ADdios-no-brasil-%C3%A9-quinta-maior-do-mundo-diretrizes-nacionais-buscam>>. Acesso em: 15 set. 2023.

PASINATO, Wânia e; SANTOS, Cecília MacDowell. **Mapeamento das Delegacias da Mulher no Brasil**. Campinas, 2008. Disponível em: <[https://www12.senado.leg.br/institucional/omv/copy\\_of\\_acervo/outras-referencias/copy2\\_of\\_entenda-a-violencia/pdfs/mapeamento-das-delegacias-da-mulher-no-brasil](https://www12.senado.leg.br/institucional/omv/copy_of_acervo/outras-referencias/copy2_of_entenda-a-violencia/pdfs/mapeamento-das-delegacias-da-mulher-no-brasil)>. Acesso em: 12 set. 2023.

SINAN - Sistema de Informação de Agravos de Notificação. **O Sinan**. Disponível em: <<https://portalsinan.saude.gov.br/>>. Acesso em: 10 ago. 2023.

NOGUEIRA, Tiago; PEDROSA, Paulo. **Computação em Nuvem**. UNICAMP, p.04, 2011. Disponível em: <https://www.ic.unicamp.br/~ducatte/mo401/1s2011/T2/Artigos/G04-095352-120531-t2.pdf>. Acesso em 16 out. 2023.

ORACLE CORPORATION. **Oracle - Autonomous Database**. Disponível em: <<https://www.oracle.com/br/autonomous-database/>>. Acesso em 18 out. 2023.

\_\_\_\_\_. **Oracle - Autonomous Data Warehouse.** Disponível em:  
<<https://www.oracle.com/br/autonomous-database/autonomous-data-warehouse/>>. Acesso em: 18 out. 2023.

\_\_\_\_\_. **Oracle – O que é Computação em Nuvem.** Disponível em:  
<<https://www.oracle.com/br/cloud/what-is-cloud-computing/#cloud-computing-benefits>>. Acesso em: 17 out. 2023.

### Apêndice 1 – Cronograma

ETAPAS	CRONOGRAMA / MESES / QUINZENA									
	JUL		AGO		SET		OUT		NOV	
	1	2	1	2	1	2	1	2	1	2
Formação dos grupos de TCC		X								
Discussão de possíveis temas			X	X						
Escolha do tema e tipo de TCC					X					
Levantamento bibliográfico					X					
Definição dos objetivos					X					
Definição do método de pesquisa					X					
Coleta de dados					X					
Tratamento dos dados					X					
Primeira entrega						X				
Análise dos dados						X	X			
Discussão dos resultados							X	X		
Segunda entrega								X		
Conclusão do artigo								X	X	
Elaboração do vídeo									X	
Entrega final									X	

Fonte: Próprio autor (2023).

### Apêndice 2 – Tabela de atributos do *Dataset* estudado

Tabela de atributos do *Dataset* estudado, analisados pela correlação de Pearson e posteriormente aplicados na matriz densa completa.

ATRIBUTO	DESCRIÇÃO	ATRIBUTO	DESCRIÇÃO
CS_GESTANT	Vítima Gestante	AUTOR_ALCO	Influência de Álcool
OUT_VEZES	Violência recorrente	REL_TRAB	Agressor é colega de trabalho
VIOL_FISIC	Violência física	AUTOR_MULT	Agressão em grupo
VIOL_PSICO	Violência Psicológica	GEN_TRANS	Vítima transgênero
VIOL_TORT	Tortura	PF_LOC_ABRIGO	Ocorrência em abrigo público
VIOL_SEXU	Violência sexual	PF_LOC_COMERC	Ocorrência em comércio
VIOL_FINAN	Violência financeira	PF_LOC_BAR	Ocorrência em bar
VIOL_LEGAL	Violência legal	PF_LOC_ESPORT	Ocorrência em clube/academia
VIOL_NEGLI	Negligência	PF_LOC_INDUST	Ocorrência em indústria/construção
VIOL_INFAN	Violência Infantil	PF_LOC_ESCOLA	Ocorrência em escola
AG_FORCA	Emprego de força	PF_LOC_RESID	Ocorrência na residência
AG_ENFOR	Enfocamento	PF_LOC_PUBLIC	Ocorrência em local/via pública
AG_OBJETO	Uso de objeto contundente	PF_RACA_PRT	Vítima preta
AG_CORTE	Uso de facas/objetos cortantes	PF_RACA_IND	Vítima indígena
AG_QUEENTE	Provocar queimaduras	PF_RACA_AMR	Vítima asiática
AG_ENVEN	Envenenamento	PF_RACA_BRC	Vítima caucasiana
AG_FOGO	Arma de fogo	PF_RACA_PRD	Vítima parda



<b>AG_AMEACA</b>	Ameaça	<b>PF_ESCOL_F1I</b>	Vítima com escolaridade fundamental incompleta
<b>SEX_ASSEDI</b>	Assédio Sexual	<b>PF_ESCOL_F1C</b>	Vítima com escolaridade fundamental completa
<b>SEX_ESTUPR</b>	Estupro	<b>PF_ESCOL_F2I</b>	Vítima com escolaridade básica incompleta
<b>SEX_PUDOR</b>	Atentado ao pudor	<b>PF_ESCOL_F2C</b>	Vítima com escolaridade básica completa
<b>SEX_EXPLO</b>	Exploração sexual	<b>PF_ESCOL_MI</b>	Vítima com escolaridade média incompleta
<b>SEX_PORNO</b>	Pornografia	<b>PF_ESCOL_MC</b>	Vítima com escolaridade média completa
<b>REL_PAI</b>	Agressor é o pai	<b>PF_ESCOL_SI</b>	Vítima com escolaridade superior incompleta
<b>REL_PAD</b>	Agressor é o padrasto	<b>PF_ESCOL_SC</b>	Vítima com escolaridade superior completa
<b>REL_MAD</b>	Agressor é a madrasta	<b>PF_ESCOL_AN</b>	Vítima analfabeta
<b>REL_CONJ</b>	Agressor é cônjuge	<b>PF_CONJ_CAS</b>	Vítima casada
<b>REL_EXCON</b>	Agressor é ex-cônjuge	<b>PF_CONJ_VIUV</b>	Vítima viúva
<b>REL_NAMO</b>	Agressor é namorado	<b>PF_CONJ_SOLT</b>	Vítima solteira
<b>REL_EXNAM</b>	Agressor é ex-namorado	<b>PF_CONJ_SEP</b>	Vítima divorciada
<b>REL_FILHO</b>	Agressor é filho	<b>PF_SEX_HET</b>	Vítima heterossexual
<b>REL_DESCO</b>	Agressor desconhecido	<b>PF_SEX_HOMO</b>	Vítima Homossexual
<b>REL_IRMAO</b>	Agressor é irmão	<b>PF_SEX_BI</b>	Vítima bissexual
<b>REL_CUIDA</b>	Agressor é cuidador	<b>PF_IDADE_CRI</b>	Vítima criança
<b>REL_CONHEC</b>	Agressor é conhecido	<b>PF_IDADE_ADL</b>	Vítima adolescente
<b>REL_PATRAO</b>	Agressor é patrão ou chefe	<b>PF_IDADE_ADT</b>	Vítima adulta
<b>REL_MAE</b>	Agressor é a mãe	<b>PF_IDADE_JOV</b>	Vítima jovem
<b>REL_POL</b>	Agressor é policial	<b>PF_IDADE_IDS</b>	Vítima idosa

Fonte: Próprio autor (2023).