

TRABALHO PRÁTICO IBD

Análise de dados de violência doméstica e feminicídio em MG

Integrantes: Elen Stephane, João Victor Lisboa, Luma Guimarães, Mariana Ordoni e Nicolle Oliveira

1. Análise Crítica das Fontes de Dados

Os dados utilizados neste trabalho foram extraídos de duas fontes primárias com naturezas e escopos de coleta distintos: a Secretaria de Estado de Justiça e Segurança Pública (SEJUSP) e a Secretaria de Estado de Saúde (SES), através do Sistema de Informação de Agravos de Notificação (SINAN). A análise dessas fontes revela importantes dificuldades, omissões e limitações de qualidade.

A base de dados da SES (Violência Doméstica) é mais diversificada e abrangente, cobrindo múltiplas facetas do fenômeno. Esta base permite recortes detalhados como raça da vítima, recorrência da violência, data de nascimento e gênero do autor, dando mais identidade para as vítimas. No entanto, ela não captura casos que não chegam ao sistema de saúde, sendo uma amostra grande, mas não universal, da violência.

Por outro lado, a base de dados da SEJUSP (Feminicídio) é significativamente mais limitada, representando apenas os casos extremos e fatais. Essa base de ocorrências policiais, ao focar na classificação do crime, carece de profundidade. A disparidade de dados é evidente pela ausência de atributos cruciais na tabela de Feminicídio que existem na de Violência Doméstica. Um crime tão brutal, que requer uma análise profunda, é computado com poucas nuances, faltando dados como o relacionamento entre o autor e a vítima e o histórico de denúncias prévias, o que impede a construção de uma análise causal direta.

Outra omissão crítica percebida foi a ausência de indicadores de saúde mental ou histórico psicológico nas bases estaduais, o que inviabilizou a análise direta sobre o impacto emocional e psíquico da violência de gênero.

Em relação à manipulação dos dados, a integração das fontes revelou um problema clássico de qualidade de dados. Foi enfrentado um desafio de inconsistência na escrita das chaves de junção. A diferença no uso de letras maiúsculas e minúsculas e a presença de caracteres especiais nos nomes dos municípios dificultaram o cruzamento de dados. Para mitigar esse problema e garantir a integridade referencial, foi necessário implementar uma rotina de pré-processamento para normalizar os nomes dos municípios, removendo acentuação e padronizando as caixas de caracteres antes de realizar a junção.

A consequência mais crítica da limitação de atributos na base de Feminicídio é a perda da diversidade da SES após o merge. Visto que não foi possível cruzar os dados de feminicídio com aspectos demográficos ou sociais detalhados, a análise integrada teve que se restringir às variáveis comuns (município e ano), o que forçou o foco do trabalho na taxa proporcional.

Em síntese, a diferença de detalhamento entre as bases reflete um contraste no perfil de coleta de informações de cada órgão e expõe uma falta de interesse da SEJUSP em investigar mais os casos de feminicídios ou, ao menos, de computá-los com mais detalhes. Além disso, também impõe restrições severas à profundidade da análise integrada que poderia ser realizada.

