

Influência do futebol nos índices de violência doméstica

Marcelo Paulino Simões¹, Cristiane N. Targa¹, Daniel Conrado¹

¹Departamento de Informática – Instituto Federal de Minas Gerais (IFMG)
CEP – 34590-390 – Sabará – MG – Brasil

ppaulinosimoess@outlook.com, {cristiane.targa,daniel.conrado}@ifmg.edu.br

Abstract. *Among the many human rights violations present in Brazil, the high rates of domestic violence against women stand out as a persistent concern. This study examines the association between Clube de Regatas do Flamengo matches and daily records of domestic violence against women in the state of Rio de Janeiro during 2020 and 2021, a period of social isolation due to the COVID-19 pandemic. Using microdata from the Dossiê Mulher (ISP-RJ) and official competition calendars, Negative Binomial Regression was applied across four analytical scenarios controlling for day of the week, public holidays, and year. Sundays with a match recorded 10.3% more complaints than Sundays without one. Sundays on which Flamengo suffered a defeat showed a significant increase of 19.1% in complaints. Across the full weekly series, the match effect was absorbed by the structural Sunday pattern, which systematically concentrates more occurrences. Public holidays consistently increased complaint volumes across all scenarios. The findings suggest that football acts as a situational catalyst for domestic violence, with effects concentrated on Sundays and amplified by defeat.*

Resumo. *Dentre os diversos problemas de violação de direitos humanos presentes no Brasil, destaca-se a elevada taxa de violência doméstica contra a mulher. Este estudo analisa a associação entre partidas do Clube de Regatas do Flamengo e registros diários de violência doméstica contra a mulher no estado do Rio de Janeiro em 2020 e 2021, período de isolamento social decorrente da pandemia de COVID-19. A partir dos microdados do Dossiê Mulher (ISP-RJ) e dos calendários oficiais de competições, aplicou-se Regressão Binomial Negativa em quatro cenários analíticos, controlando dia da semana, feriados e ano. Domingos com jogo registraram 10,3% mais denúncias do que domingos sem jogo. Domingos em que há derrota do Flamengo mostraram um aumento significativo de 19,1% nas denúncias. Na série semanal completa, o efeito do jogo foi absorvido pelo padrão estrutural do domingo, que concentra sistematicamente mais ocorrências. Nota-se que feriados elevaram consistentemente o volume de denúncias em todos os cenários analisados. Os resultados sugerem que o futebol atua como catalisador situacional da violência doméstica, com efeito concentrado nos domingos e potencializado pela derrota.*

1. Introdução

A violência doméstica contra a mulher constitui um dos principais desafios contemporâneos em saúde pública e direitos humanos, com elevada incidência no Brasil. Estudos nacionais e internacionais indicam que partidas de futebol podem atuar como gatilhos para comportamentos agressivos, especialmente quando combinadas à frustração emocional, à rivalidade esportiva e ao consumo de álcool.

A pandemia de COVID-19 introduziu um cenário singular: com o isolamento social e os jogos realizados sem público, a carga emocional das partidas passou a ser absorvida predominantemente no espaço privado, justamente onde se manifestam as dinâmicas de violência intrafamiliar. Apesar de estudos anteriores terem identificado associações entre eventos esportivos e violência doméstica, ainda há lacunas sobre como esse fenômeno se comporta em situações de isolamento compulsório.

Diante desse contexto, este estudo analisa o impacto das partidas de futebol nos índices de violência doméstica contra a mulher no Rio de Janeiro em 2020 e 2021, investigando tanto a ocorrência dos jogos quanto seus resultados como possíveis fatores explicativos.

Este artigo está dividido da seguinte maneira: na seção 2 faz-se um diálogo com estudos da literatura sobre a influência do esporte na violência doméstica; na seção 3 apresenta-se a metodologia, a base de dados e o processo de mineração; na seção 4 apresenta-se a análise de resultados; na seção 5 realiza-se uma discussão; e a seção 6 conclui com as considerações finais.

2. Fundamentação Teórica

Esta seção apresenta os conceitos e *frameworks* teóricos que embasam o estudo. Organiza-se em eixos temáticos: (i) violência doméstica contra a mulher e seu aparato legal no Brasil; (ii) os efeitos da pandemia de COVID-19 sobre a violência intrafamiliar; (iii) os efeitos da pandemia no futebol; (iv) a relação entre eventos esportivos e comportamento violento; (v) o processo *Knowledge Discovery in Databases* (KDD).

2.1. Violência Doméstica contra a Mulher

A violência doméstica contra a mulher constitui um fenômeno social de ampla prevalência no Brasil, reconhecido tanto como violação de direitos humanos quanto como problema de saúde pública. A Organização Mundial da Saúde (OMS) estima que, globalmente, uma em cada três mulheres já sofreu violência física ou sexual por parte de um parceiro íntimo, padrão que se reproduz no contexto brasileiro com intensidade expressiva [World Health Organization 2021].

No ordenamento jurídico nacional, a Lei nº 11.340/2006 conhecida como Lei Maria da Penha, representa o principal marco normativo de enfrentamento da violência doméstica e familiar contra a mulher [Brasil 2006]. O diploma legal tipifica cinco formas de violência: física, psicológica, sexual, patrimonial e moral; cria mecanismos de proteção, como as medidas protetivas de urgência; e estabelece juizados especializados para processamento e julgamento das causas. Segundo [Dias 2010], a lei representou uma ruptura paradigmática no tratamento jurídico do tema, ao conferir caráter público e incondicionado à ação penal nos casos de lesão corporal dolosa em contexto doméstico.

Para que a violência seja classificada como doméstica sob amparo da Lei nº 11.340/2006, é imperativa a presença de ao menos um dos seguintes critérios relacionais:

- I. **Relação Íntima de Afeto:** Compreende vínculos afetivos atuais ou extintos, formais ou não. A coabitação e o tempo de relacionamento são fatores irrelevantes; o que determina a incidência da norma é a existência de intimidade e a assimetria de poder entre as partes. Assimetria de poder significa que, numa relação íntima, os dois lados raramente ocupam posições de igual força. Um dos parceiros, na esmagadora maioria dos casos, a mulher, tende a estar em posição de maior vulnerabilidade.
- II. **Âmbito da Unidade Doméstica:** Define-se pelo espaço físico de convívio permanente, incluindo pessoas que compartilham a residência sem necessariamente possuírem vínculos familiares.
- III. **Vínculo Familiar:** Caracteriza-se pela relação entre indivíduos que se consideram aparentados por laços naturais (consanguíneos), por afinidade (parentes por parte de cônjuges) ou por vontade expressa, independentemente do local de residência das partes.

2.2. Pandemia de COVID-19 e Violência Doméstica

A decretação da Emergência em Saúde Pública de Importância Internacional (ESPII) pela Organização Mundial da Saúde, em março de 2020, impôs medidas de isolamento social em escala global que, embora necessárias do ponto de vista sanitário, produziram consequências severas sobre a segurança das mulheres no ambiente doméstico [World Health Organization 2020].

[Evans et al. 2020], analisando o cenário global durante a pandemia, mostraram que o isolamento social ampliou a exposição das vítimas aos seus agressores e, ao mesmo tempo, dificultou o acesso aos serviços de denúncia e proteção. Para as autoras, a ruptura das redes de apoio e o estresse financeiro decorrente do lockdown tornaram as vítimas mais vulneráveis, na medida em que o controle coercitivo passou a ser exercido longe de qualquer vigilância externa.

No Brasil, o Fórum Brasileiro de Segurança Pública registrou aumento de 22% nos feminicídios e crescimento de 27% nas denúncias ao Ligue 180 nos meses de março e abril de 2020, em comparação ao mesmo período do ano anterior [Fórum Brasileiro de Segurança Pública 2020]. No Rio de Janeiro especificamente, as chamadas para a Polícia Militar referentes a violência doméstica cresceram 5% em abril de 2020 em relação ao mesmo mês de 2019. [Garcia and Santana 2020], em análise sobre o cenário brasileiro publicada na revista *Ciência & Saúde Coletiva*, caracterizam a violência doméstica durante a crise sanitária como a 'pandemia das sombras'. As autoras argumentam que a coexistência forçada e o estresse econômico foram agravados pela fragilização das redes de apoio e dos serviços de assistência social, que enfrentaram limitações operacionais severas no momento em que as vítimas mais necessitavam de proteção estatal.

Em resposta a esse cenário, o Congresso Nacional aprovou a Lei nº 14.022/2020, que classificou os serviços de atendimento às vítimas de violência doméstica como essenciais durante o estado de calamidade pública, determinando sua continuidade inin-

terrupta e a priorização das denúncias recebidas pelos canais Ligue 180 e Disque 100 [Brasil 2020].

Do ponto de vista metodológico, a pandemia introduz desafios analíticos importantes para estudos de violência doméstica: por um lado, o isolamento pode ter aumentado a ocorrência de crimes; por outro, as restrições de mobilidade podem ter reduzido o acesso das vítimas às delegacias, gerando subnotificação.

2.3. Impactos da Pandemia de COVID-19 no Futebol Brasileiro

A pandemia de COVID-19 impôs transformações profundas na dinâmica do futebol brasileiro, alterando as condições em que o esporte era produzido e consumido. As competições foram disputadas total ou parcialmente sem a presença de torcedores, alterando variáveis centrais da experiência esportiva e, sobretudo, os mecanismos pelos quais os torcedores acompanhavam e se relacionavam com o futebol.

2.3.1. Paralisação Total e Retomada com Portões Fechados (2020)

Em 15 de março de 2020, a Confederação Brasileira de Futebol (CBF) anunciou a suspensão por tempo indeterminado das competições nacionais sob sua coordenação, determinando que as Federações Estaduais deliberassem autonomamente sobre seus torneios [GloboEsporte.com 2020]. A Federação de Futebol do Estado do Rio de Janeiro (FERJ) adotou medida semelhante, interrompendo o Campeonato Carioca na fase de retorno inicialmente por 15 dias [ESPN Brasil 2020c], prazo sucessivamente prorrogado por meio de deliberações da própria entidade e de recomendações do Ministério Público do Estado do Rio de Janeiro diante do agravamento da pandemia [Agência Brasil 2020b, Agência Brasil 2020c].

A primeira partida do período pós-paralisação ocorreu em 18 de junho de 2020, com o confronto entre Flamengo e Bangu no Estádio Jornalista Mário Filho, válido pela quarta rodada da Taça Rio (campeonato carioca), após 96 dias consecutivos sem futebol no estado [Agência Brasil 2020a]. O retorno foi realizado sem a presença de público e sob protocolos sanitários rigorosos. O Campeonato Brasileiro da Série A, por sua vez, foi reiniciado em 9 de agosto de 2020, igualmente com portões fechados e sob diretrizes estabelecidas pela CBF, que incluíam testagem periódica dos profissionais envolvidos [ESPN Brasil 2020a, ESPN Brasil 2020b].

2.3.2. Reabertura Progressiva e Conflitos Institucionais (2021)

O retorno gradual do público aos estádios do Rio de Janeiro ocorreu de forma escalonada e permeada por disputas entre entidades esportivas e instâncias administrativas. Em 29 de julho de 2021, a Prefeitura do Rio de Janeiro divulgou o plano de retomada "Rio de Novo", prevendo a permissão de público com esquema vacinal completo em três fases: inicialmente com 50% da capacidade a partir de 2 de setembro e a liberação total para 100% a partir de 17 de outubro [Prefeitura do Rio de Janeiro 2021e].

Na prática, a viabilização desse cronograma ocorreu por meio de três *eventos-teste* no Maracanã, autorizados pela Secretaria Municipal de Saúde em 6 de setembro de

2021. Esses jogos estabeleceram uma progressão de ocupação: 35% da capacidade na partida contra o Grêmio pela Copa do Brasil (15 de setembro); 40% no confronto pelo Campeonato Brasileiro (19 de setembro); e 50% no duelo com o Barcelona de Guayaquil pela Copa Libertadores (22 de setembro) [Prefeitura do Rio de Janeiro 2021g]. Para além do limite de público, o protocolo sanitário exigia cadastro prévio em sistema *online*, comprovante de vacinação completa e resultado negativo de teste RT-PCR ou de antígeno realizado em até 48 horas antes do evento [Prefeitura do Rio de Janeiro 2021g].

A decisão gerou tensão no âmbito do futebol nacional. No dia 8 de setembro, reunidos em Conselho Técnico Extraordinário convocado pela CBF, os representantes de 19 dos 20 clubes da Série A deliberaram pela manutenção do Campeonato Brasileiro sem público e pleitearam junto ao STJD a reconsideração da liminar que autorizava o Flamengo a receber torcedores [Agência Brasil 2021a]. O Flamengo, único clube ausente da reunião, manteve a posição de que a deliberação sobre presença de público em estádios competia exclusivamente às autoridades governamentais locais, e manteve a liminar obtida junto ao Superior Tribunal de Justiça Desportiva (STJD) [Exame 2021].

Em 15 de setembro de 2021, o Maracanã registrou a presença de cerca de 6.400 torcedores na partida entre Flamengo e Grêmio, válida pelas quartas de final da Copa do Brasil, o primeiro evento com público em estádios do Rio de Janeiro desde o início da pandemia [CNN Brasil 2021], encerrando um hiato de 549 dias sem público nos estádios do estado, contados a partir da suspensão oficial das competições em 15 de março de 2020. Em 17 de setembro, a Prefeitura formalizou, por meio do Decreto Rio n.º 49.411, a liberação de até 50% da capacidade dos estádios municipais para competições esportivas, condicionada ao esquema vacinal completo de todos os presentes.

Cabe registrar que o primeiro jogo de uma competição de primeira divisão do futebol brasileiro com presença de torcedores após o início da pandemia não ocorreu no Rio de Janeiro, mas em Brasília. Em 21 de julho de 2021, o Flamengo recebeu o Defensa y Justicia no Estádio Nacional Mané Garrincha, pelas oitavas de final da Copa Libertadores da América, com capacidade liberada de 25% do estádio (até 15 mil torcedores), após autorização do Governo do Distrito Federal [Agência Brasil 2021b]. A partida foi transferida do Maracanã para Brasília justamente porque, à época, o município do Rio de Janeiro ainda não havia autorizado o retorno de público aos seus estádios [Agência Brasil 2021b]. [Prefeitura do Rio de Janeiro 2021c, Prefeitura do Rio de Janeiro 2021f]. A autorização para ocupação integral dos estádios foi concedida por meio do Decreto Rio n.º 49.692, de 26 de outubro de 2021 [Prefeitura do Rio de Janeiro 2021d], mantendo-se a exigência de comprovação de esquema vacinal completo [Prefeitura do Rio de Janeiro 2021e].

2.3.3. Restrições ao Funcionamento de Bares e Restaurantes (2020–2021)

As medidas de controle sanitário adotadas pelo município do Rio de Janeiro afetaram diretamente os espaços de sociabilidade vinculados ao consumo do futebol, em especial os bares e restaurantes frequentados por torcedores antes e durante as partidas. Em 21 de março de 2020, o Decreto Rio n.º 47.282 determinou a suspensão do atendimento presencial nesses estabelecimentos, autorizando a operação exclusivamente nas moda-

lidades de entrega em domicílio (*delivery*) e *drive-thru*, vedado o consumo no local [Prefeitura do Rio de Janeiro 2020a].

A reabertura para atendimento à mesa ocorreu em 2 de julho de 2020, quando a Prefeitura anunciou o ingresso na Fase 3A do plano municipal de flexibilização. A partir dessa data, bares e restaurantes puderam receber clientes presencialmente, com limitação de 50% da capacidade de lotação na área interna, distanciamento mínimo de dois metros entre as mesas e uso obrigatório de máscara facial, permitida sua retirada apenas no momento da alimentação [Prefeitura do Rio de Janeiro 2020b].

O agravamento do quadro epidemiológico no início de 2021 impôs nova rodada de restrições. O Decreto Rio n.º 48.641, de 19 de março de 2021, proibiu novamente o atendimento presencial em bares e restaurantes, restringindo a operação às modalidades de *delivery*, *take away* e *drive-thru* [Prefeitura do Rio de Janeiro 2021a]. Tais restrições foram consolidadas e prorrogadas pelo Decreto Rio n.º 48.644, de 22 de março de 2021 [Prefeitura do Rio de Janeiro 2021b]. A normalização progressiva do setor acompanhou o avanço da campanha de vacinação, sendo o atendimento presencial integralmente restabelecido, com distanciamento mínimo de 1,0 metro entre conjuntos de mesa e cadeiras nas áreas internas, pelo Decreto Rio n.º 49.411, de 16 de setembro de 2021, o mesmo instrumento normativo que formalizou a liberação parcial de público nos estádios municipais [Prefeitura do Rio de Janeiro 2021c].

2.4. Eventos Esportivos e Violência Doméstica

A hipótese de que eventos esportivos de grande intensidade emocional podem atuar como gatilhos para a violência doméstica tem recebido crescente atenção na literatura. Um dos mecanismos explicativos apontados é o de que resultados negativos inesperados geram frustração e emoções negativas nos torcedores. Em situações de forte envolvimento emocional com a equipe, muitas vezes associadas ao consumo de álcool, esses sentimentos podem aumentar a probabilidade de comportamentos agressivos no ambiente doméstico.

O consumo de álcool ocupa papel relevante como variável mediadora nessa relação. [Ivandić et al. 2021] identificaram que o aumento mais expressivo de violência doméstica ocorre em dias de jogo quando há registro de consumo de álcool pelo agressor, especialmente nas horas subsequentes ao término da partida. No contexto brasileiro, a centralidade do álcool no ritual de assistir aos jogos de futebol em bares, residências e estádios amplifica os riscos de comportamentos agressivos, embora a literatura seja enfática ao distinguir o álcool como fator agravante e não como causa primária da violência [Instituto Avon and Fórum Brasileiro de Segurança Pública 2022].

O estudo de [Card and Dahl 2011] analisou partidas da *National Football League* (NFL) ao longo de 12 temporadas (1995–2006), envolvendo seis equipes em suas respectivas cidades. Como estratégia de identificação, os autores utilizaram o *point spread* pré-jogo, uma medida do mercado de apostas que indica a margem de vitória esperada para o time favorito. Os resultados mostram que derrotas inesperadas, definidas como aquelas em que o time local era favorito por quatro ou mais pontos, elevaram em aproximadamente 10% o número de registros policiais de violência física de homens contra parceiras íntimas. O efeito concentrou-se em uma janela temporal próxima ao término das partidas, sendo mais pronunciado em jogos de maior rivalidade e inexistente quando o resultado esperado se confirmou. O trabalho tornou-se uma referência metodológica na

literatura ao demonstrar empiricamente que eventos esportivos podem atuar como gatilhos de violência doméstica.

No Brasil, o Instituto Avon, em parceria com o Fórum Brasileiro de Segurança Pública, realizou a pesquisa *Violência Contra Mulheres e o Futebol* [Instituto Avon and Fórum Brasileiro de Segurança Pública 2022], analisando dados de boletins de ocorrência de cinco capitais (Rio de Janeiro, São Paulo, Salvador, Belo Horizonte e Porto Alegre) em todos os dias de jogos do Campeonato Brasileiro da Série A entre 2015 e 2018. O estudo identificou aumento de 23,7% nos registros de ameaça e de 20,8% nas lesões corporais dolosas em dias de jogo, com elevação de 25,9% nos dias em que o time local atuou como mandante.

Do ponto de vista da estratégia analítica, a pesquisa empregou Modelos Lineares Generalizados (GLM) com distribuição Binomial Negativa, abordagem também adotada no presente trabalho, estimando se o número de ocorrências em determinado dia e cidade se eleva em função da realização de partidas. Foram incluídas como variáveis de controle: dia da semana, mês, ano, temperatura máxima, índice pluviométrico, fase do campeonato (início, meio ou final) e a condição de feriado local. Os autores ressaltam que o futebol não constitui a causa da violência, mas pode atuar como um catalisador de desequilíbrios de poder preexistentes, ao interagir com valores de masculinidade, competitividade e frustração [Instituto Avon and Fórum Brasileiro de Segurança Pública 2022].

Complementarmente, [Árabe 2022] desenvolveu a dissertação de mestrado *Gol Contra: Impacto das Partidas de Futebol na Violência Contra a Mulher*, na Universidade de São Paulo, analisando aproximadamente 3.800 partidas do Campeonato Brasileiro das Séries A e B em conjunto com registros de ocorrência de 737 municípios dos estados de São Paulo e Rio de Janeiro no período de 2015 a 2019. A autora utilizou probabilidades de vitória de sites de apostas para definir o resultado esperado de cada jogo, identificando que derrotas inesperadas estão associadas a aumentos nos registros de violência física e psicológica contra mulheres, em padrão consistente com o evidenciado por [Card and Dahl 2011] para o futebol americano.

Os trabalhos de [Árabe 2022] e [Instituto Avon and Fórum Brasileiro de Segurança Pública 2022] são, portanto, os mais próximos metodologicamente do presente estudo no contexto nacional. Diferenciam-se, contudo, por analisarem um período pré-pandêmico e um conjunto mais amplo de equipes e localidades.

O presente estudo diferencia-se das análises anteriores ao investigar um cenário de “isolamento compulsório”, no qual os jogos foram disputados majoritariamente com portões fechados e sob restrições severas ao funcionamento de bares e espaços de convívio. Esse contexto singular funciona como um experimento natural para avaliar se o efeito dos eventos esportivos sobre a violência doméstica persiste, ou até se intensifica, quando a experiência do torcedor é concentrada no ambiente privado, concentrando a carga emocional do jogo no núcleo familiar.

2.5. Knowledge Discovery in Databases (KDD)

O *Knowledge Discovery in Databases* (KDD) é definido como o processo não trivial de identificação de padrões válidos, novos, potencialmente úteis e compreensíveis em dados [Fayyad et al. 1996]. Proposto na segunda metade dos anos 1990, o *framework*

consolidou-se como referência metodológica para projetos que envolvem a extração de conhecimento a partir de grandes volumes de dados estruturados.

O processo KDD é composto por cinco etapas sequenciais e iterativas:

1. **Seleção:** identificação e extração das fontes de dados relevantes para o problema investigado;
2. **Pré-processamento:** limpeza dos dados brutos, tratamento de valores ausentes, inconsistências e ruídos;
3. **Transformação:** conversão das variáveis para formatos adequados à aplicação dos algoritmos de mineração, incluindo engenharia de atributos;
4. **Mineração de dados:** aplicação de algoritmos e modelos estatísticos para identificar padrões e relações nos dados transformados;
5. **Interpretação e avaliação:** análise dos padrões descobertos à luz do domínio de aplicação, avaliação de sua validade e relevância.

A adoção do KDD como *framework* metodológico neste estudo justifica-se pela heterogeneidade das fontes de dados utilizadas base de violência doméstica do ISP-RJ e base de partidas do Flamengo e pela necessidade de etapas sistemáticas de integração e transformação antes da aplicação dos modelos estatísticos. O processo garante rastreabilidade e reprodutibilidade das decisões metodológicas em cada fase da análise.

3. Metodologia

Este estudo investigou a influência das partidas de futebol nos índices de violência doméstica nos anos de 2020 e 2021. O período analisado caracteriza-se pela realização da maioria dos jogos com portões fechados e pela imposição de regras de isolamento social, eliminando a presença física de torcedores nos estádios e bares e alterando, assim, a dinâmica de exposição aos eventos esportivos.

A abordagem metodológica fundamentou-se no framework Knowledge Discovery in Databases (KDD), que compreende as etapas sequenciais de seleção, pré-processamento, transformação, mineração de dados e, por fim, interpretação e avaliação dos resultados [Fayyad et al. 1996]. Na fase de mineração, foram empregados testes de hipóteses (t de Student e Mann-Whitney) e Modelos Lineares Generalizados (MLG) com distribuição binomial negativa, com o objetivo de mensurar o efeito das partidas sobre as ocorrências de violência.

A primeira etapa do KDD, a seleção de dados, consistiu na identificação e extração das bases utilizadas neste estudo. As características dessas bases, incluindo sua origem, período de abrangência e critérios de inclusão, são detalhadas na subseção 3.1.

3.1. Base de dados

Para o desenvolvimento do estudo, foi necessário selecionar um estado que atendesse a dois critérios principais: dispor de uma base de dados pública sobre violência doméstica com discriminação das datas dos crimes e ser sede de um clube de futebol com grande torcida.

Nesse contexto, foi escolhido o estado do Rio de Janeiro, que disponibiliza dados estatísticos de violência doméstica discriminados por data do crime por meio

do *Dossiê Mulher*, mantido pelo Instituto de Segurança Pública do Rio de Janeiro (ISP). A base abrange o período de 2014 a 2021 e está disponível em formato `.csv` [Instituto de Segurança Pública 2026]. Além disso, o estado é sede do Clube de Regatas do Flamengo, cujo torcedor é majoritário na capital fluminense: segundo levantamento do Instituto Prefab Future [INSTITUTO PREFAB FUTURE 2025], realizado nos dias 3 e 4 de novembro de 2025 com 1.200 entrevistados no Rio de Janeiro e divulgado pelo portal *Lance!*, 56% dos moradores declararam torcer para o Flamengo.

Todos os registros utilizados protegem a identidade das vítimas. A Tabela 1 apresenta um exemplo de registro retirado da base de dados, contendo o título do campo, o valor empregado e uma breve explicação de cada variável.

Tabela 1. Registro de Ocorrência de Violência Doméstica com Explicação dos Campos

Título	Exemplo	Explicação
Ano	2020	Ano em que a ocorrência foi registrada
RGOCRONU	001-00014/2020	Número do registro da ocorrência no sistema
ETIT	25	Código interno do evento
ESEQ	2	Variável de controle interno da ISP
Região	Baixada Fluminense	Região administrativa do RJ onde ocorreu o fato
RISP	3ª RISP	Região Integrada de Segurança Pública responsável
AISP	21ª AISP	Área Integrada de Segurança Pública correspondente
CISP	64ª CISP	Circunscrição Integrada de Segurança Pública correspondente
Delegacia de Registro	OUTRAS DELEGACIAS	Local onde a ocorrência foi registrada
Descrição	Lesão Corporal (outros)	Tipo de crime registrado
Tipo de Violência	Violência Física	Classificação geral da violência
Título	Lesão corporal dolosa	Categoria legal do delito
Detalhamento	Sem informação	Descrição adicional do fato
Meio Empregado	Sem informação	Instrumento ou método utilizado no crime
Idade	27	Idade da vítima
Faixa Etária	18 a 29 anos	Faixa etária da vítima
Cor	Parda	Autodeclaração ou classificação da vítima
Relação Vítima-Autor	Companheiro ou ex	Grau de relação entre vítima e autor
Estado Civil	Sem informação	Estado civil da vítima
Município	São João de Meriti	Município onde ocorreu o fato
Escolaridade	Sem informação	Nível educacional da vítima
Local do Fato	Residência	Local em que ocorreu a violência
Data da Comunicação	03/01/2020	Data em que a ocorrência foi comunicada à polícia
Hora da Comunicação	11:17:43	Hora em que a ocorrência foi comunicada
Faixa Horária Comunicação	Manhã - 06h às 11h59	Intervalo de tempo da comunicação
Hora do Fato	09:00	Hora aproximada em que o fato ocorreu
Data do Fato	27/12/2019	Data em que o fato ocorreu
Faixa Horária do Fato	Manhã - 06h às 11h59	Intervalo de tempo em que o fato ocorreu
Lei Aplicada	Lei 11.340/06	indica se há aplicação de lei específica

Em relação à base de jogos do Flamengo, foram selecionados os campeonatos: Brasileirão (nacional), Libertadores (continental) e Carioca (estadual), por serem os campeonatos com maior número de clássicos (partidas contra rivais do mesmo estado) e por serem as competições mais valorizadas pelos torcedores.

Os dados de partidas do Flamengo foram coletados manualmente, por meio da verificação das tabelas oficiais de cada campeonato, abrangendo o período de 2017 a 2021. A Tabela 2 apresenta um exemplo de registro retirado da base de dados, contendo o título do campo, o valor empregado e uma breve explicação de cada variável.

Tabela 2. Registro de Partida de Futebol - Exemplo

Campo	Exemplo	Explicação
Rodada	28	Rodada do campeonato
Data	06/01/2021	Data em que a partida foi realizada
Hora	21:30	Hora de início da partida
Mandante	Flamengo	Time que jogou em casa
Visitante	Fluminense	Time visitante
Mandante Placar	1	Número de gols marcados pelo time mandante
Visitante Placar	2	Número de gols marcados pelo time visitante
Competição	Brasileiro	Nome da competição

A etapa de pré-processamento compreende a preparação da base, fase em que ocorre os procedimentos de limpeza e padronização dos dados para a pesquisa, o detalhamento é apresentado na 3.2.

3.2. Pré-processamento e tratamento de dados

Nesta etapa, realizaram-se procedimentos de limpeza e preparação dos dados, visando assegurar a integridade e a qualidade das informações analisadas. O recorte temporal considerou as diretrizes da Portaria GM/MS nº 913 [BRASIL 2022], que delimitou a Emergência em Saúde Pública de Importância Nacional (ESPIN) entre fevereiro de 2020 e maio de 2022.

Os procedimentos de tratamento incluíram:

- Filtragem temporal: Como a base de dados de denúncias vai até 31/12/2021 o recorte temporal analisado foi fixado nos anos de 2020 e 2021.
- Limpeza de dados: Exclusão dos dados sem registro da data em que o crime ocorreu.
- Padronização de formatos: Harmonização das variáveis de data das suas bases para um formato único, permitindo a correlação consistente entre as bases de dados.
- Recorte de sazonalidade: Exclusão do período entre janeiro e 16 de junho de 2020, por representar o cenário pré-pandêmico e a fase de suspensão das competições esportivas, evitando assim vieses nos resultados.
- Seleção de variáveis: Identificação e extração dos atributos de interesse para as análises estatísticas, cujas definições e justificativas de escolha são detalhadas nesta mesma seção.

Na base do Dossiê Mulher, foram mantidos os atributos listados na Tabela 6, imprescindíveis para calcular o número de ocorrências ao longo de um dia e para a aplicação de critérios de filtragem nas categorias criminais. Na base dos jogos foram mantidos os atributos apresentados na Tabela 4.

Tabela 3. Atributos selecionados – Violência contra a mulher

Coluna	Tipo	Descrição
tipo_violencia	Alfanumérico	Tipo de violência (Física, Psicológica, Moral, Sexual ou Patrimonial)
data_fato	Data	Data em que o crime ocorreu
local_fato	Alfanumérico	Local em que o crime ocorreu
relacao_vitima_autor	Alfanumérico	Grau de relação entre vítima e autor

Tabela 4. Atributos selecionados - Tabela de jogos

Campo	Tipo de Dado	Descrição
Rodada	Alfanumérico	Rodada do campeonato
Data	Data	Data em que a partida foi realizada
Hora	Data	Hora de início da partida
Mandante	Alfanumérico	Time que jogou em casa
Visitante	Alfanumérico	Time visitante
Mandante Placar	Numérico	Número de gols marcados pelo time mandante
Visitante Placar	Numérico	Número de gols marcados pelo time visitante
Competição	Alfanumérico	Nome da competição

A análise preliminar dos dados relativos à violência contra a mulher permite identificar padrões críticos, conforme ilustrado na Figura 1 sobre a distribuição dos locais dos crimes, observa-se uma concentração crítica de ocorrências no ambiente doméstico. O local 'Residência' responde por 60,1% dos registros. Em contraste, a 'Via pública' e 'Outros locais' apresentam índices significativamente menores, com 16,5% e 14,0%, respectivamente. A baixa incidência em 'Estabelecimentos comerciais' (2,6%) e o registro de 2,9% em 'Ambientes virtuais' indicam que, embora a violência transite por diferentes esferas, a residência permanece como o espaço de maior vulnerabilidade para as vítimas.

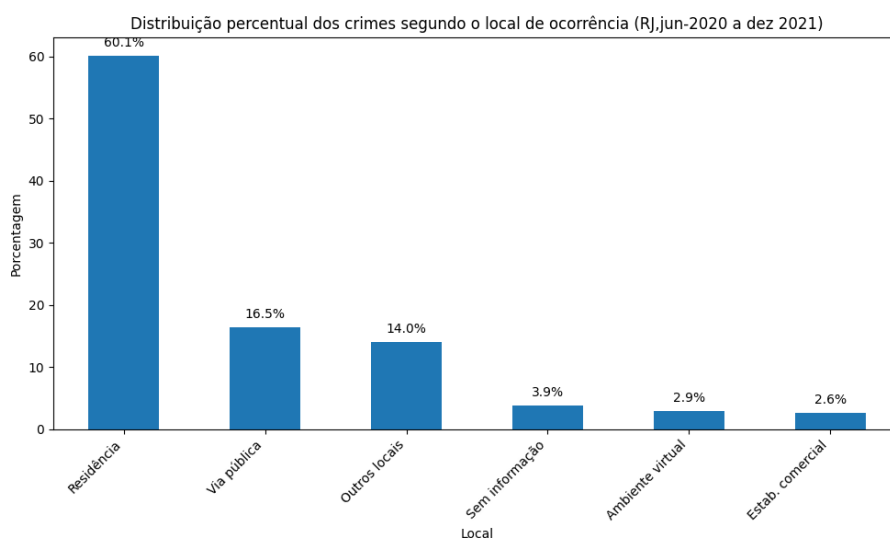


Figura 1. Distribuição percentual dos crimes contra a mulher por local de ocorrência no Estado do Rio de Janeiro (Jun/2020 – Dez/2021)
Fonte: Elaboração própria com base em dados da SEPOL/ISP-RJ.

Adicionalmente, a análise do vínculo entre vítima e agressor (Figura 2) revela que a violência de gênero é em sua maioria, dentro âmbito das relações afetivas, sendo a categoria "Companheiro ou ex" a mais expressiva, com 46,1% das ocorrências.

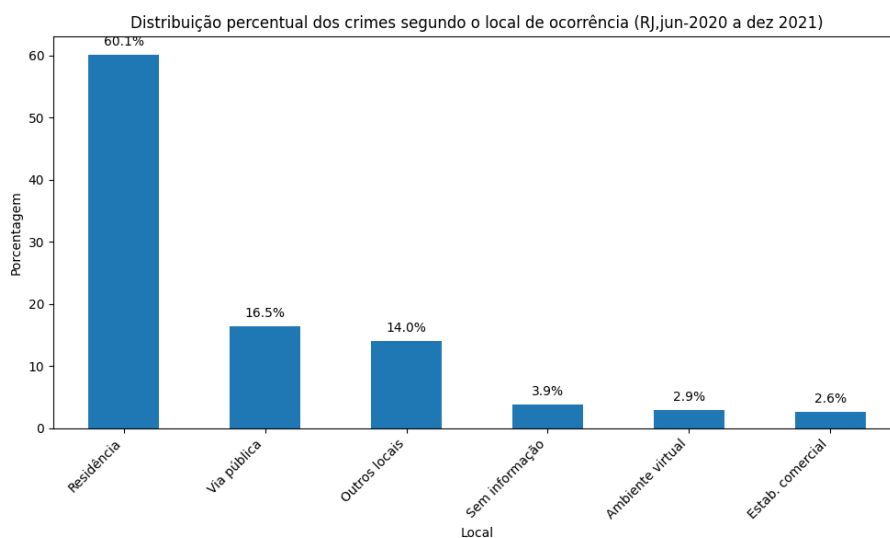


Figura 2. Distribuição percentual do vínculo entre vítima e autor dos delitos registrados no Rio de Janeiro (Jun/2020 – Dez/2021)
Fonte: Elaboração própria com base em dados da SEPOL/ISP-RJ.

A interpretação conjunta das Figuras 1 e 2 indicam que violência contra a mulher no Rio de Janeiro se trata de um fenômeno predominantemente intrafamiliar, enquanto o primeiro indicador aponta que 60,1% dos delitos ocorrem no ambiente da residência, o segundo indicador confirma que a maioria absoluta dos agressores pertence ao círculo íntimo da vítima. Ao somar as categorias 'Companheiro ou ex' (46,1%), 'Parente' (7,5%)

e 'Pais ou padrastos' (2,2%), observa-se que em 55,8% dos casos o autor possui um vínculo de confiança ou afeto direto com a mulher.

Após a etapa de limpeza e definição do período de análise, as variáveis selecionadas passaram por um processo de transformação e harmonização. Esse procedimento teve como finalidade converter os dados brutos em variáveis estruturadas, possibilitando sua utilização nas análises estatísticas sobre o impacto de eventos futebolísticos na violência contra a mulher.

3.3. Transformação

A Figura 3 apresenta a distribuição dos tipos de violência registrados na base de dados entre 17/06/2020 e 31/12/2021 no estado do Rio de Janeiro. As categorias de violência psicológica (33,3%), física (33,2%) e moral (23,8%) respondem conjuntamente por 90,3% dos registros, enquanto a violência patrimonial (5,0%) e a violência sexual (4,8%) somam menos de 10% do total.

Diante dessa distribuição, optou-se por restringir a análise às três categorias majoritárias. Além de representarem a maior parte dos casos, violência psicológica, física e moral são as formas mais diretamente associadas a conflitos interpessoais imediatos, dinâmica central para o objetivo deste estudo, que investiga o efeito de eventos esportivos sobre o comportamento da violência doméstica.

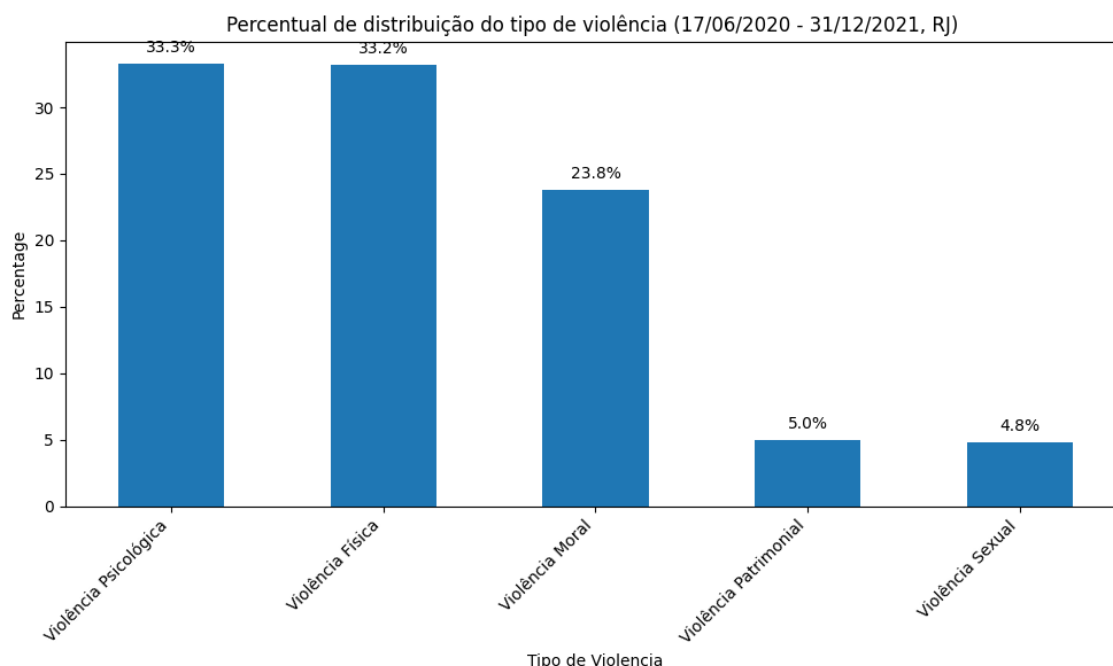


Figura 3. Distribuição percentual dos tipos de violência doméstica registrados entre 17/06/2020 e 31/12/2021, Rio de Janeiro.

A identificação dos casos de violência doméstica na base de dados seguiu os critérios estabelecidos pelo Art. 5º da Lei nº 11.340/2006, que define violência doméstica como aquela praticada no âmbito da unidade doméstica, no âmbito familiar ou em qualquer relação íntima de afeto. Em conformidade com esse dispositivo, foram selecionados os registros que atendessem a ao menos um dos seguintes critérios:

- (i) **Enquadramento legal:** casos em que o campo `lei_aplicada` continha o valor “Lei 11.340/06”, indicando que o próprio agente responsável pelo registro classificou a ocorrência sob a Lei Maria da Penha;
- (ii) **Relação entre vítima e autor:** casos em que o campo `relacao_vitima_autor` correspondia a um dos vínculos previstos na lei: cônjuge, companheiro, namorado, ex-parceiro ou familiar, caracterizando relação íntima de afeto ou vínculo familiar independentemente do enquadramento legal registrado;
- (iii) **Local do fato:** casos em que o campo `local_fato` indicava “Residência”, sugerindo convívio doméstico entre as partes mesmo quando a relação ou o enquadramento legal não tenham sido registrados de forma precisa.

A adoção de critérios alternativos é intencional, assim como o Art. 5º da lei, basta a presença de um dos critérios para que o caso seja classificado como violência doméstica. O uso de critérios cumulativos resultaria em filtragem mais restritiva do que o próprio texto legal determina, com risco de subnotificação dos casos relevantes para a análise.

A partir desses dados, realizou-se o cálculo da frequência diária de denúncias, contabilizando valores únicos em uma determinada coluna, neste caso, a coluna `'data_fato'`. O resultado foi estruturado em uma nova base de dados, na qual cada linha representa uma data e o número total de ocorrências no dia.

Segundo o Instituto de Segurança Pública [Instituto de Segurança Pública 2020], cerca de 39,7% das vítimas de lesão corporal foram agredidas aos sábados e domingos. Considerando que a maioria das partidas de futebol ocorre nesses dias, faz-se necessária a distinção entre os dias da semana, a fim de evitar uma relação causal indevida entre o aumento das ocorrências e a realização de jogos e foram excluídos da amostra os feriados nacionais para evitar viés de sazonalidade atípica no volume de registros policiais. Para isso, foi criado o atributo `dia_da_semana`, codificado de 0 (segunda-feira) a 6 (domingo).

Foram selecionados os feriados nacionais e estaduais do Rio de Janeiro mas sua abordagem diferiu entre as duas etapas da análise. Nos testes de hipótese, eles foram excluídos da amostra para isolar o efeito do calendário esportivo, mantê-los poderia distorcer as médias diárias, já que feriados nacionais alteram o padrão de denúncias por razões alheias ao futebol.

Na regressão binomial negativa, os feriados foram reincorporados à base e controlados por uma variável *dummy*. Essa escolha tem dupla finalidade. Primeiro, permite que o coeficiente de `dia_de_jogo` seja estimado já descontando o efeito dos feriados, ou seja, a comparação entre dias com e sem jogo passa a ocorrer em condições equivalentes, sem que datas comemorativas contaminem a estimativa. Segundo, o próprio coeficiente da variável `is_feriado` torna-se informativo: ele revela o quanto os feriados, por si sós, elevam ou reduzem o volume de denúncias em relação a um dia comum, independentemente da ocorrência de jogos. Dessa forma, a modelagem não apenas corrige um possível viés, mas também quantifica um efeito que, como se verá nos resultados, mostrou-se consistente e estatisticamente significativo em todos os cenários analisados.

Para uma análise sobre o impacto do resultado dos jogos foi criado o atributo `resultado_jogo`. O processamento envolveu o cruzamento dos placares oficiais com a condição do clube (mandante ou visitante), categorizando o desfecho em três níveis no-

minais: vitória, derrota ou empate. Após os procedimentos descritos as bases de dados se encontram da seguinte maneira:

Tabela 5. Atributos selecionados - Violência contra mulher

Coluna	Tipo de Dado	Descrição
data_fato	Data	Data
dia_da_semana	Numérico	Representação numérica do dia da semana.
Ano	Numérico	Ano de referência do registro.
num_denuncias	Numérico	Variável de contagem: número total de registros de violência doméstica por dia.

Tabela 6. Atributos selecionados - Jogos do Flamengo

Coluna	Tipo de Dado	Descrição
data_fato	Data	Data
resultado_jogo	Numérico	Variável categórica: Vitória, Derrota, Empate

A integração dos conjuntos de dados foi realizada por meio do método de junção à esquerda (*Left Join*), utilizando a base de ocorrências criminais como tabela primária e a variável data como chave de integração entre as bases. Esse procedimento assegurou a manutenção de todos os registros diários do período analisado, adicionando informações sobre os resultados dos jogos apenas nas datas em que ocorreram partidas. Nos dias em que não houve jogo, as células correspondentes permaneceram inicialmente vazias, sendo posteriormente preenchidas com a categoria “Sem jogo”.

Para viabilizar a modelagem estatística, foi criada uma variável indicadora (dummy) denominada dia_de_jogo, assumindo valor 1 para datas com partidas e 0 para as demais. Nas instâncias onde não houve correspondência entre as bases (dias sem jogo), a variável relativas ao desempenho esportivo foi codificadas com o rótulo 'Ausente', evitando o descarte de observações e preservando a variabilidade necessária para a aplicação dos testes estatísticos e Modelos Lineares Generalizados (MLG).

A estrutura final das variáveis após a integração e que serão utilizadas na análise é detalhada na Tabela 7

Tabela 7. Descrição das variáveis utilizadas

Variável	Tipo	Descrição
data_fato	Fato	Data do registro das denúncias (formato AAAA-MM-DD)
num_denuncias	Inteiro	Número total de denúncias de violência doméstica registradas no dia
dia_da_semana	Inteiro	Dia da semana codificado numericamente (0 = segunda-feira, ..., 6 = domingo)
ano	Inteiro	Ano de referência do registro
is_feriado	Binária	Indica se a data corresponde a um feriado nacional ou estadual (<i>True/False</i>)
Resultado_jogo	Catégorica	Desfecho da partida do Flamengo na data (Vitória, Derrota, Empate ou Não houve jogo)
dia_de_jogo	Binária	Indica se houve partida do Flamengo na data (<i>True/False</i>)

A partir desses dados, foi possível calcular as médias de ocorrências de violência doméstica por dia da semana, exceto feriados. A Figura 4 apresenta esse resultado, analisando o período durante a pandemia (2020–2021).

O impacto é heterogêneo ao longo da semana. Segunda (+12,9%), quinta (+8,5%), sábado (+1,6%) e domingo (+10,6%) mostram elevação em dias de jogo, enquanto terça (-2,7%), quarta (-5,8%) e sexta (-2,2%) registram leve redução. Dos 116 jogos analisados, a distribuição dos jogos ao longo da semana se deu da seguinte maneira: o domingo foi o dia com mais partidas (37), seguido pela quarta-feira (24), sábado (19), quinta-feira (12), terça-feira (11), segunda-feira (5), e sexta-feira (3), além de 5 jogos nos feriados. Em feriados, dias de jogo registraram média de 119 ocorrências contra 160 em dias sem jogo (-26,8%). Esse resultado aparentemente paradoxal pode refletir um efeito de composição: poucos jogos foram realizados em feriados na amostra.

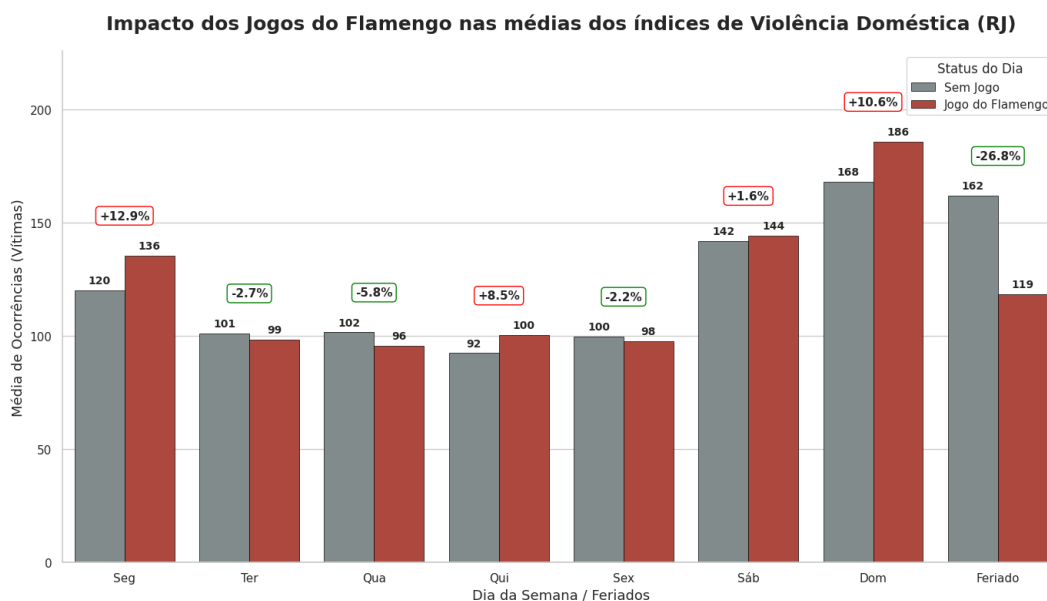


Figura 4. Médias de ocorrências por dia da semana no intervalo de 18/06/20 até 31/12/2021

Entretanto, a simples diferença entre médias não é suficiente para sugerir a existência de correlação ou relação casual entre o aumento nas denúncias e a ocorrência de jogos do Flamengo. Para conferir rigor científico aos achados e verificar a significância estatística dessas variações, os dados foram submetidos a testes estatísticos e à modelagem via Modelo Linear Generalizado (MLG), detalhados na subseção 3.4.

3.4. Mineração de Dados

A etapa de mineração de dados representa o núcleo do framework KDD, na qual métodos computacionais e estatísticos são aplicados para a extração de padrões e relações significativas nos dados previamente processados. Nesta pesquisa, a mineração foi delineada para investigar se a ocorrência de partidas de futebol exerce um efeito estatisticamente relevante sobre os registros de violência doméstica. Para assegurar a robustez dos achados, a análise foi estruturada em dois eixos metodológicos complementares:

1. **Testes de Hipóteses Estatísticas:** Foram aplicados testes de comparação de grupos (*t de Student* e *Mann-Whitney*) com o objetivo de verificar se existe uma diferença significativa entre as médias de denúncias nos dias de jogos e nos dias sem eventos esportivos.
2. **Modelagem Preditiva via Modelos Lineares Generalizados (MLG):** Em um segundo momento, visando uma análise multivariada mais profunda, empregaram-se Modelos Lineares Generalizados com distribuição Binomial Negativa. Esta técnica foi selecionada por sua adequação a dados de contagem que apresentam superdispersão (variância superior à média), permitindo isolar o impacto das variáveis esportivas enquanto se controla o efeito de variáveis de confusão, como dias da semana.

A seguir, detalham-se os parâmetros de implementação de cada um desses métodos.

3.4.1. Testes estatísticos

A definição do teste estatístico foi baseada no objetivo da análise que é a comparação entre dias com e sem jogos de futebol estratificados por dia da semana e nas características dos dados das amostras.

Os grupos foram considerados como amostras independentes pois as ocorrências de violência doméstica em dias de jogo e dias sem jogo de um determinado dia da semana e não possuem vínculo entre si. Cabe ressaltar que foram excluídos os feriados da base de análise, com o intuito de mitigar o viés de confusão provocado por datas comemorativas, que tradicionalmente apresentam variações atípicas no comportamento social e nos índices de criminalidade.

Como os registros foram feitos em datas diferentes ao longo de um ano e meio, não há uma relação de causa e efeito entre os dias comparados. As variáveis são classificadas como quantitativas discretas, pois são resultado da contagem de denúncias no dia, assumindo apenas valores inteiros e não negativo.

Segundo Vieira [Vieira 2018], para a comparação de dois grupos independentes, o teste t de Student e o teste de Mann–Whitney são definidos como as ferramentas de análise mais consolidadas. O teste t de Student é um procedimento paramétrico que exige que a distribuição das amostras seja normal (em formato de 'sino'). Já o teste de Mann–Whitney, por ser não paramétrico, não exige uma distribuição específica. Portanto, a verificação da normalidade dos dados constitui um passo essencial para a escolha entre esses testes; para isso, utilizou-se o teste de Shapiro–Wilk, seguindo a recomendação de [Razali and Wah 2011], que aponta esse procedimento como o mais robusto para amostras pequenas ($n < 50$).

Para a interpretação do teste de Shapiro–Wilk, adotou-se o nível de significância de 5% ($\alpha = 0,05$). Segundo [Vieira 2018], este é o limiar de probabilidade mais utilizado na literatura científica para a tomada de decisão estatística, caso o p -valor seja inferior a 0,05, rejeita-se a hipótese nula de normalidade, indicando que os dados não seguem uma distribuição normal. A estatística (W) representa o quanto os dados se aproximam de uma distribuição normal, o valor varia entre 0 e 1, quanto mais próximo de 1 mais os dados se parecem com uma distribuição normal. O resultado da aplicação do teste é ilustrado na tabela 8.

Tabela 8. Resultados do teste de normalidade de Shapiro–Wilk para os dias com e sem jogos

Dia da semana	Grupo	Estatística (W)	p-valor	Distribuição
Segunda-feira	Com jogos	0,9647	0,8401	Normal
	Sem jogos	0,9881	0,7571	Normal
Terça-feira	Com jogos	0,9318	0,4290	Normal
	Sem jogos	0,9458	0,0071	Não normal
Quarta-feira	Com jogos	0,9531	0,3164	Normal
	Sem jogos	0,9739	0,2859	Normal
Quinta-feira	Com jogos	0,9559	0,7238	Normal
	Sem jogos	0,8871	<0,0001	Não normal
Sexta-feira	Com jogos	0,9700	0,6678	Normal
	Sem jogos	0,9902	0,8503	Normal
Sábado	Com jogos	0,9166	0,0978	Normal
	Sem jogos	0,9600	0,0568	Normal
Domingo	Com jogos	0,9775	0,6441	Normal
	Sem jogos	0,9628	0,1965	Normal

A análise dos resultados do teste de normalidade indica que apenas os dados referentes à quinta-feira e terça-feira não apresentam distribuição normal. Dessa forma, para esse grupo será utilizado o teste não paramétrico Mann-Whitney. Para os demais dias da semana é necessário aplicar o teste de Levene para verificar a homogeneidade das variâncias entre os grupos analisados. Essa etapa é necessária porque o teste T de Student pressupõe que as amostras comparadas apresentem variâncias iguais. O resultado da aplicação do teste é ilustrado na tabela 9. Como todas as amostras possuem variâncias iguais, será aplicado o teste T Student nas mesmas.

Tabela 9. Resultados do teste de Levene para verificação da homogeneidade das variâncias.

Dia da semana	Estatística(F)	p-valor	Conclusão
Segunda-feira	1.6916	0.1975	Variâncias iguais
Quarta-feira	0.2475	0.6203	Variâncias iguais
Sexta-feira	1.2125	0.2744	Variâncias iguais
Sábado	1.2331	0.2704	Variâncias iguais
Domingo	0.2056	0.6515	Variâncias iguais

3.4.2. Modelagem via Modelos Lineares Generalizados (MLG)

A variável dependente `num_denuncias` representa contagens diárias de denúncias — valores inteiros e não negativos. Isso inviabiliza a regressão linear clássica, que pressupõe resíduos normais e admite previsões negativas. A alternativa adotada foi a dos **Modelos**

Lineares Generalizados (MLG), capazes de acomodar distribuições da família exponencial por meio de uma função de ligação apropriada.

A primeira distribuição considerada foi a de *Poisson*, padrão para dados de contagem. Contudo, dados criminais tendem a exibir superdispersão isto é, a variância supera a média ($\text{Var}(Y) > \mathbb{E}(Y)$), condição que invalida o modelo de Poisson, pois ele assume igualdade entre esses dois momentos e, na presença de superdispersão, subestima os erros padrão dos coeficientes [Fávero and Belfiore 2017]. A solução foi a Regressão Binomial Negativa, que introduz um parâmetro adicional de dispersão e , segundo [Cameron and Trivedi 2013], constitui uma extensão natural do modelo de Poisson para situações em que essa restrição não se sustenta.

Codificação das variáveis

Antes da aplicação do modelo, as variáveis explicativas foram codificadas conforme sua natureza. As variáveis `dia_jogo` e `tem_feriado` já estavam em formato binário (0 ou 1), não exigindo nenhuma transformação adicional. As variáveis `resultado_jogo`, `dia_da_semana` e `ano`, por serem categóricas com mais de duas categorias, foram convertidas em *dummies*. Em cada uma delas, uma categoria foi omitida como nível de referência procedimento necessário para evitar a multicolinearidade perfeita, conhecida como armadilha da variável *dummy*, que ocorre quando as categorias de um grupo somam um valor constante e tornam a matriz do modelo singular.

Cenários analíticos

A modelagem foi estruturada em quatro cenários para capturar a influência do futebol sob diferentes recortes. Em todos eles, adotou-se a função de ligação logarítmica:

$$\ln(\mu_i) = \beta_0 + \sum_k \beta_k X_{ki} + \epsilon_i \quad (1)$$

onde μ_i é o número esperado de denúncias no dia i , β_0 o intercepto e X_{ki} o conjunto de variáveis explicativas e de controle.

- **Cenário 1 — Apenas domingos (jogo vs. sem jogo):** compara domingos com e sem partida, controlando por ano e feriado.
- **Cenário 2 — Apenas domingos (desfecho do jogo):** substitui a variável de presença de jogo pelas *dummies* Vitória, Derrota e Empate, mantendo os mesmos controles, para avaliar se o resultado altera o volume de denúncias.
- **Cenário 3 — Série completa (jogo vs. sem jogo):** estende a análise a todos os dias da semana, acrescentando as *dummies* de dia da semana ao conjunto de controles.
- **Cenário 4 — Série completa (desfecho do jogo):** recorte mantendo os mesmos controles do Cenário 3 mas com as *dummies* de resultado, verificando se vitória, derrota ou empate produzem efeitos distinguíveis analisando todos os dia da semana.

Os coeficientes foram convertidos em **Razões de Taxas de Incidência (IRR)**, o que permite ler os resultados diretamente como variações percentuais no número de denúncias associadas a cada variável. Os resultados são discutidos na Seção 4

4. Análise dos resultados

4.0.1. Testes estatísticos

Os resultados dos testes de hipótese (Tabela 10) mostram que, apesar das variações percentuais nas médias observadas ao longo da semana, somente o domingo apresentou diferença estatisticamente significativa entre dias com e sem jogos do Flamengo ($p = 0,0080$).

Nesse dia, o teste t de Student rejeitou a hipótese nula ao nível de 5%, o que aponta uma associação entre a realização de partidas aos domingos e o aumento nas notificações de violência doméstica. Os demais dias não atingiram significância estatística, ainda que alguns exibam variações percentuais relevantes na análise descritiva.

O caso mais chamativo é o da segunda-feira: o aumento de 12,9% nas médias é expressivo, mas o p-valor de 0,1021 não permite rejeitar H_0 . Isso pode decorrer da maior dispersão dos dados nesse recorte ou simplesmente do tamanho da amostra, insuficiente para detectar um efeito que, a julgar pela magnitude observada, pode de fato existir.

Tabela 10. Resultados dos testes de hipótese para comparação entre dias com e sem jogos

Dia	Teste	Estatística	p-valor	Conclusão
Segunda-feira	t-Student	-1,6560	0,1021	Não rejeita H_0
Terça-feira	Mann-Whitney	389,5000	0,5793	Não rejeita H_0
Quarta-feira	t-Student	1,2977	0,1983	Não rejeita H_0
Quinta-feira	Mann-Whitney	312,5000	0,2239	Não rejeita H_0
Sexta-feira	t-Student	0,1945	0,8463	Não rejeita H_0
Sábado	t-Student	-0,3596	0,7202	Não rejeita H_0
Domingo	t-Student	-2,7224	0,0080	Rejeita H_0

Nível de significância adotado: $\alpha = 0,05$. Terça e Quinta-feira utilizaram Mann-Whitney por violação da normalidade. H_0 : não há diferença significativa entre as médias (ou medianas) dos grupos.

4.0.2. Modelagem

As Tabelas 11 e 12 apresentam os resultados dos quatro cenários de Regressão Binomial Negativa, organizados em dois grupos: Cenários 1 e 2, restritos aos domingos, e Cenários 3 e 4, que abrangem a série semanal completa. Para cada variável são reportados o coeficiente estimado, o p-valor, a razão de incidência (IRR) e o respectivo intervalo de confiança de 95%.

Os quatro modelos convergiram normalmente e mostraram-se globalmente significativos. A presença de superdispersão foi confirmada pelo parâmetro α significativo em

todos os cenários ($p < 0,001$), validando a escolha da Regressão Binomial Negativa em detrimento do modelo de Poisson.

Tabela 11. Regressão Binomial Negativa — Cenários 1 e 2 (apenas domingos)

Variável	Cenário 1 (jogo vs. sem jogo)				Cenário 2 (desfecho do jogo)			
	Coef.	p	IRR	IC 95%	Coef.	p	IRR	IC 95%
Constante	5,141	<0,001	—	—	5,142	<0,001	—	—
<i>Variáveis de interesse</i>								
Dia de jogo	0,098	0,006	1,103	[1,028; 1,184]	—	—	—	—
Derrota	—	—	—	—	0,175	0,001	1,191	[1,076; 1,318]
Vitória	—	—	—	—	0,071	0,096	1,074	[0,987; 1,168]
Empate	—	—	—	—	0,045	0,479	1,046	[0,924; 1,185]
<i>Controles</i>								
Feriado	0,340	0,002	1,404	[1,129; 1,749]	0,339	0,002	1,404	[1,135; 1,737]
Ano 2021	-0,024	0,525	0,977	[0,907; 1,051]	-0,025	0,486	0,975	[0,908; 1,048]
<i>Diagnóstico</i>								
α (dispersão)	0,019	<0,001	—	—	0,018	<0,001	—	—
Pseudo R^2			0,019				0,024	
N			80				80	

Categoria de referência para resultado do jogo: domingos sem jogo. LLR p-valor: 0,002 (Cenário I) e 0,002 (Cenário II).

Tabela 12. Regressão Binomial Negativa — Cenários 3 e 4 (série semanal completa)

Variável	Cenário 3 (jogo vs. sem jogo)				Cenário 4 (desfecho do jogo)			
	Coef.	<i>p</i>	IRR	IC 95%	Coef.	<i>p</i>	IRR	IC 95%
Constante	4,816	<0,001	—	—	4,816	<0,001	—	—
<i>Variáveis de interesse</i>								
Dia de jogo	0,016	0,438	1,016	[0,975; 1,059]	—	—	—	—
Derrota	—	—	—	—	0,066	0,105	1,069	[0,986; 1,157]
Vitória	—	—	—	—	0,014	0,586	1,014	[0,964; 1,066]
Empate	—	—	—	—	-0,022	0,575	0,978	[0,904; 1,058]
<i>Controles</i>								
Feriado	0,289	<0,001	1,335	[1,241; 1,436]	0,292	<0,001	1,338	[1,245; 1,439]
Ano 2021	-0,044	0,008	0,957	[0,926; 0,989]	-0,045	0,007	0,956	[0,923; 0,988]
<i>Efeitos do dia da semana (ref.: segunda-feira)</i>								
Terça-feira	-0,173	<0,001	0,842	[0,794; 0,892]	-0,170	<0,001	0,844	[0,795; 0,895]
Quarta-feira	-0,193	<0,001	0,824	[0,777; 0,876]	-0,193	<0,001	0,824	[0,776; 0,875]
Quinta-feira	-0,252	<0,001	0,777	[0,733; 0,824]	-0,253	<0,001	0,776	[0,731; 0,824]
Sexta-feira	-0,162	<0,001	0,851	[0,803; 0,902]	-0,162	<0,001	0,851	[0,802; 0,903]
Sábado	0,157	<0,001	1,170	[1,104; 1,241]	0,158	<0,001	1,171	[1,105; 1,242]
Domingo	0,379	<0,001	1,461	[1,377; 1,551]	0,376	<0,001	1,455	[1,372; 1,544]
<i>Diagnóstico</i>								
α (dispersão)	0,027	<0,001	—	—	0,027	<0,001	—	—
Pseudo R^2			0,092				0,092	
<i>N</i>			562				562	

Categoria de referência para dia da semana: segunda-feira. Categoria de referência para resultado do jogo: sem jogo. LLR *p*-valor < 0,001 em ambos os cenários.

Cenários I e II: Recorte aos Domingos

No Cenário 1, restrito aos domingos, a presença de jogo do Flamengo associou-se a um aumento médio de 10,3% nas denúncias em relação a domingos sem jogo ($p = 0,006$), resultado estatisticamente significativo. O Cenário 2 desagregou o efeito pelo desfecho da partida. Apenas a derrota atingiu significância estatística, associando-se a 19,1% mais denúncias do que domingos sem jogo ($p = 0,001$). Vitória ($p = 0,096$) e empate ($p = 0,479$) não alcançaram significância.

Cenários III e IV: Série Semanal Completa

Ao incorporar todos os dias da semana com controle por efeito fixo, o Cenário 3 revelou que o efeito isolado da presença de jogo desapareceu ($p = 0,438$), absorvido pela sazonalidade estrutural da semana. Tomando a segunda-feira como categoria de referência, os dias úteis registraram volumes consistentemente menores de denúncias (todos $p < 0,001$), enquanto sábado e domingo superaram a linha de base, confirmando a concentração estrutural de ocorrências nos fins de semana. O Cenário 4 replicou esse padrão substituindo a variável binária de jogo pelas categorias de desfecho da partida.

Nenhuma delas, derrota ($p = 0,105$), empate ($p = 0,575$) ou vitória ($p = 0,586$), atingiu significância estatística, indicando que o resultado da partida não acrescenta poder explicativo quando a análise abrange toda a semana.

Em conjunto, os resultados indicam que a associação entre futebol e violência doméstica é contextualmente condicionada: o efeito emerge com nitidez aos domingos e se intensifica na derrota, mas é absorvido pelo padrão estrutural da semana quando a análise é ampliada para todos os dias.

Diagnóstico de resíduos de Pearson

Para validar a adequação da Regressão Binomial Negativa, realizou-se o diagnóstico de resíduos de Pearson em cada cenário (Tabela 13). O critério central é a razão entre o qui-quadrado de Pearson e os graus de liberdade (Qui²/gl): valores próximos de 1 indicam que o modelo captura adequadamente a dispersão dos dados; valores muito superiores sinalizariam superdispersão residual não explicada.

Em todos os cenários, a razão Qui²/gl ficou entre 1,057 e 1,091, e a média dos resíduos foi igual a zero, indicando ausência de viés sistemático e bom ajuste global do modelo.

O comportamento dos resíduos diferiu ligeiramente entre os recortes. Nos Cenários 1 e 2, restritos aos domingos, o teste de Shapiro-Wilk não rejeitou a normalidade dos resíduos ($p = 0,161$ e $p = 0,258$, respectivamente), e os valores extremos mantiveram-se dentro de uma faixa estreita ($\pm 2,75$). Nos Cenários 3 e 4, que abrangem a série completa, o teste rejeitou a normalidade ($p < 0,001$), e o valor máximo dos resíduos atingiu 5,78, indicando a presença de algumas observações influentes. Esse resultado é esperado em séries temporais longas com dados de contagem e não compromete a validade das estimativas, dado que a Regressão Binomial Negativa não pressupõe normalidade dos resíduos.

Tabela 13. Diagnóstico de resíduos de Pearson por cenário

Cenário	Média	DP	Min / Max	Qui ² /gl	W	p (S-W)
Cenário 1	0,000	1,007	-2,172 / 2,750	1,068	0,977	0,161
Cenário 2	0,000	1,005	-2,234 / 2,402	1,091	0,980	0,258
Cenário 3	0,000	1,019	-2,854 / 5,785	1,057	0,966	<0,001
Cenário 4	0,000	1,019	-2,862 / 5,773	1,060	0,966	<0,001

DP = desvio padrão dos resíduos de Pearson. Qui²/gl = razão entre o qui-quadrado de Pearson e os graus de liberdade. W = estatística de Shapiro-Wilk.

A Figura 5 apresenta a inspeção visual dos resíduos para todos os cenários. Os resíduos se distribuem de forma aproximadamente aleatória em torno de zero, sem padrões sistemáticos que indiquem problemas de especificação ou heteroscedasticidade relevante.

Nos Cenários 1 e 2, os Q-Q plots mostram aderência próxima à diagonal teórica ao longo de toda a distribuição, com leve afastamento apenas nas caudas, comportamento

consistente com os p-valores do Shapiro-Wilk. Nos Cenários 3 e 4, o desvio na cauda superior é mais pronunciado, correspondendo às observações isoladas de valor máximo 5,78 identificadas na Tabela 13. Como a Regressão Binomial Negativa não pressupõe normalidade dos resíduos, esse achado não compromete a validade das estimativas.

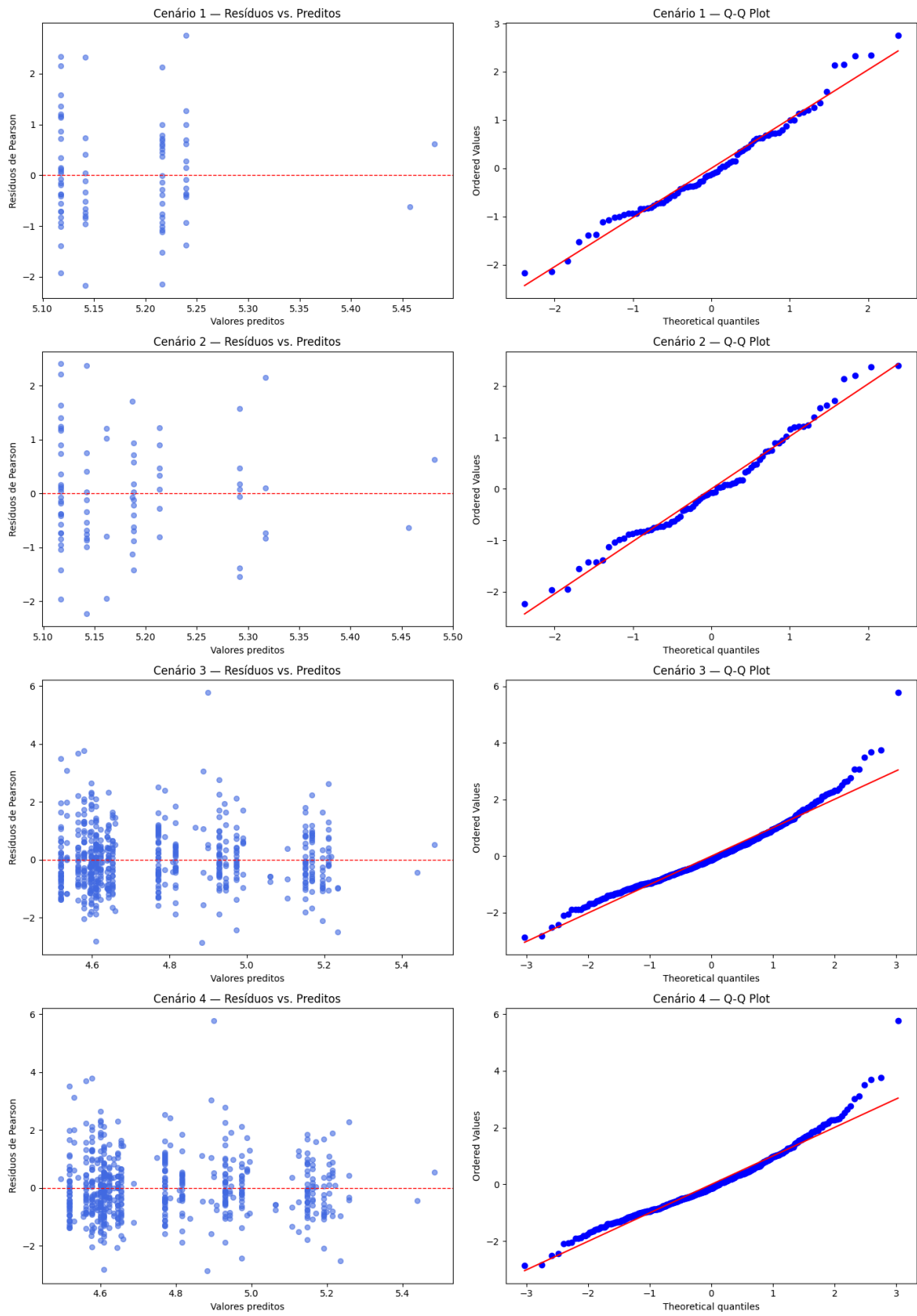


Figura 5. Diagnóstico de resíduos de Pearson por cenário. Coluna esquerda: resíduos versus valores preditos. Coluna direita: Q-Q plot de normalidade.

5. Discussão

Os resultados indicam que a associação entre futebol e violência doméstica não é uniforme, mas sim condicionada pelo contexto em que o jogo ocorre: o efeito emerge com nitidez aos domingos e se intensifica na derrota, mas perde expressão quando diluído ao longo da semana.

No recorte aos domingos, a presença de jogo associou-se a um aumento médio de 10% nas denúncias em relação a domingos sem partida, quando desagregado pelo resultado (Cenário 2), apenas a derrota atingiu significância estatística, associando-se a 19,1% mais denúncias do que domingos sem jogo. Esse resultado é coerente com o mecanismo de frustração-agressão: resultados negativos geram estados emocionais que, associados ao forte vínculo identitário entre torcedor e clube, podem transbordar para o espaço doméstico. O achado converge com [Card and Dahl 2011], que identificaram aumento nos registros de violência física de homens contra parceiras íntimas após derrotas inesperadas na NFL, e com [Árabe 2022], que encontrou associação semelhante para o futebol brasileiro entre 2015 e 2019.

O coeficiente de feriado estatisticamente significativo ao longo dos quatro cenários indicam que o modelo captura padrões reais e consistentes nos dados, e não artefatos estatísticos. Esse resultado aponta, ainda, para o maior tempo de convivência doméstica como um mecanismo de risco transversal, presente independentemente da realização de jogos. A diferença significativa entre 2020 e 2021, observada apenas nos Cenários 3 e 4, indica que o fenômeno apresentou trajetória declinante ao longo do período, mas que esse padrão só se torna identificável quando a estrutura semanal dos dados é levada em conta. Nos cenários restritos aos domingos que compõem os Cenários 1 e 2, a variação entre os anos não alcança significância estatística.

Os valores de pseudo R^2 obtidos entre 0,019 e 0,024 nos cenários restritos aos domingos e 0,092 na série semanal completa indicam que as variáveis esportivas explicam uma parcela limitada da variação total nas denúncias. Isso é esperado: a violência doméstica é um fenômeno multidimensional, influenciado por fatores socioeconômicos, culturais e individuais que extrapolam o escopo deste estudo. O objetivo não é explicar a totalidade das ocorrências, mas identificar se o futebol atua como um catalisador situacional.

Entre as limitações do estudo, destacam-se a subnotificação estrutural das denúncias, possivelmente agravada pelo isolamento pandêmico, a ausência de dados sobre consumo de álcool como variável mediadora, fator identificado pela literatura como agravante relevante [Ivandić et al. 2021] e o recorte a um único clube e estado, o que ressoa recorte a um único clube e estado, o que restringe a generalização dos achados. No entanto, a escolha de um único clube é metodologicamente justificada: ao evitar a sobreposição de jogos de times diferentes no mesmo dia, garante-se que o efeito observado seja atribuível a uma partida específica, o que seria inviável com dois ou mais clubes, dado que seus jogos frequentemente ocorrem nas mesmas datas, dificultando o isolamento do resultado que gerou o efeito. O desenho observacional impede inferências causais.

6. Considerações Finais

Este estudo trouxe evidências que partidas do Flamengo estão associadas a elevações nos registros de violência doméstica contra a mulher no Rio de Janeiro, com efeito concen-

trado nos domingos e potencializado pela derrota. A persistência dessa associação em um contexto de isolamento social, sem público nos estádios e com restrições ao convívio coletivo, sugere que o futebol atua como catalisador emocional da violência intrafamiliar independentemente de mediadores externos.

Os achados reforçam a pertinência de políticas de proteção sensíveis ao calendário esportivo, como o reforço de linhas de atendimento e ações preventivas em domingos de jogos. Para pesquisas futuras, recomenda-se a replicação do estudo em outros estados e clubes, a incorporação de variáveis mediadoras como consumo de álcool, e a comparação entre os períodos com e sem público nos estádios, oportunidade metodológica valiosa para isolar o papel do ambiente coletivo na relação entre futebol e violência doméstica.

Referências

- Agência Brasil (2020a). Campeonato carioca volta quinta-feira, com flamengo x bangu no maracanã.
- Agência Brasil (2020b). Clubes decidem pela continuação do campeonato carioca.
- Agência Brasil (2020c). Mp-rj recomenda a suspensão de possível retorno do campeonato carioca.
- Agência Brasil (2021a). Conselho do brasileirão rejeita retorno de público a estádios.
- Agência Brasil (2021b). Flamengo x defesa y justicia terá quase 20 mil no estádio em Brasília.
- Brasil (2006). Lei nº 11.340, de 7 de agosto de 2006 (Lei Maria da Penha).
- Brasil (2020). Lei nº 14.022, de 7 de julho de 2020.
- BRASIL (2022). Portaria GM/MS nº 913, de 22 de abril de 2022. Declara o encerramento da Emergência em Saúde Pública de Importância Nacional (ESPIN) em decorrência da infecção humana pelo novo coronavírus (2019-nCoV). Seção 1, p. 1.
- Cameron, A. C. and Trivedi, P. K. (2013). *Regression Analysis of Count Data*. Cambridge University Press, Cambridge, 2 edition.
- Card, D. and Dahl, G. B. (2011). Family violence and football: The effect of unexpected emotional cues on violent behavior. *The Quarterly Journal of Economics*, 126(1):103–143.
- CNN Brasil (2021). Maracanã volta a receber torcedores em primeiro evento teste na pandemia.
- Dias, M. B. (2010). *A Lei Maria da Penha na Justiça*. Revista dos Tribunais.
- ESPN Brasil (2020a). Cbf reafirma que brasileiro começa em 9 de agosto e diz que times de SP podem jogar em outro estado.
- ESPN Brasil (2020b). Covid-19: Cbf anuncia mudanças e deixa protocolo de testagem mais rígido; veja novas regras.
- ESPN Brasil (2020c). Federação do rio suspende campeonato carioca por 15 dias por causa do coronavírus.
- Evans, M. L., Lindauer, M., and Farrell, M. E. (2020). A pandemic within a pandemic — intimate partner violence during Covid-19. *New England Journal of Medicine*, 383(24):2302–2304.
- Exame (2021). Conselho do brasileirão rejeita volta de público a estádios.
- Fávero, L. P. and Belfiore, P. (2017). *Manual de Análise de Dados: Estatística e Modelagem Multivariada*. Elsevier, Rio de Janeiro, 1 edition.
- Fayyad, U., Piatetsky-Shapiro, G., and Smyth, P. (1996). From data mining to knowledge discovery in databases. *AI Magazine*, 17(3):37–54.
- Fórum Brasileiro de Segurança Pública (2020). Violência doméstica durante a pandemia de Covid-19. Technical report, Fórum Brasileiro de Segurança Pública.

- Garcia, L. P. and Santana, L. R. (2020). Evolução das desigualdades sociais em saúde relacionadas à violência contra a mulher durante a pandemia de COVID-19 no Brasil. *Ciência & Saúde Coletiva*, 25(9):3303–3312.
- GloboEsporte.com (2020). Cbf suspende todas as competições a partir de segunda-feira.
- Instituto Avon and Fórum Brasileiro de Segurança Pública (2022). Violência contra mulheres e o futebol. Technical report, Instituto Avon / FBSP.
- Instituto de Segurança Pública (2020). Dossiê mulher 2020. Acesso em: 7 jan. 2026.
- Instituto de Segurança Pública (2026). Dossiê mulher - isp visualização: Base de dados (2014-2021). Aba Notas Metodológicas. Dados extraídos em formato .csv.
- INSTITUTO PREFAB FUTURE (2025). Pesquisa de torcida no rio de janeiro: Flamengo lidera na capital. Publicado por: Lance!. Acesso em: 30 jan. 2026.
- Ivandić, R., Kirchmaier, T., and Torres-Blas, N. (2021). Football, alcohol and domestic abuse. Discussion Paper 1781, Centre for Economic Performance, London School of Economics and Political Science, London, UK.
- Prefeitura do Rio de Janeiro (2020a). Decreto Rio n.º 47.282, de 21 de março de 2020. Diário Oficial do Município do Rio de Janeiro.
- Prefeitura do Rio de Janeiro (2020b). Prefeitura anuncia volta de mais atividades na cidade com restrições: academias, estúdios de tatuagem e serviços de bares, restaurantes e lanchonetes à mesa.
- Prefeitura do Rio de Janeiro (2021a). Decreto Rio n.º 48.641, de 19 de março de 2021. Diário Oficial do Município do Rio de Janeiro.
- Prefeitura do Rio de Janeiro (2021b). Decreto Rio n.º 48.644, de 22 de março de 2021. Diário Oficial do Município do Rio de Janeiro.
- Prefeitura do Rio de Janeiro (2021c). Decreto Rio n.º 49.411, de 16 de setembro de 2021. Diário Oficial do Município do Rio de Janeiro.
- Prefeitura do Rio de Janeiro (2021d). Decreto Rio n.º 49.692, de 26 de outubro de 2021. Diário Oficial do Município do Rio de Janeiro.
- Prefeitura do Rio de Janeiro (2021e). Prefeitura anuncia plano de retomada e lança o calendário de eventos Rio de Novo.
- Prefeitura do Rio de Janeiro (2021f). Prefeitura libera eventos em locais abertos a partir da próxima terça-feira.
- Prefeitura do Rio de Janeiro (2021g). Prefeitura libera público em três jogos do Flamengo que funcionarão como eventos-teste.
- Razali, N. M. and Wah, Y. B. (2011). Power comparisons of shapiro-wilk, kolmogorov-smirnov, lilliefors and anderson-darling tests. *Journal of Statistical Modeling and Analytics*, 2(1):21–33.
- Vieira, S. (2018). *Introdução à Bioestatística*. Elsevier, Rio de Janeiro, 5 edition.
- World Health Organization (2020). Covid-19 and violence against women: What the health sector and partners can do.

World Health Organization (2021). Violence against women prevalence estimates, 2018: Global, regional and national prevalence estimates for intimate partner violence against women and global and regional prevalence estimates for non-partner sexual violence against women. Technical report, WHO, Geneva.

Árabe, I. (2022). Gol contra: Impacto das partidas de futebol na violência contra a mulher. Master's thesis, Universidade de São Paulo.