

logo/logo_nova.pdf

Boletim de Projetos 2024.1

Conteúdo

| | | |
|-----|---|----|
| I | Close Elections e Canais de Combate à Violência Contra a Mulher | 2 |
| II | O mercado brasileiro de corridas de montanha | 15 |
| III | Análise da acessibilidade após possível intervenção na Rua Uberabinha | 25 |
| IV | Consultoria FRST Falconi - Aplicações de simulação de Monte Carlo | 44 |

Close Elections e Canais de Combate à Violência Contra a Mulher

Políticas Públicas

Pesquisadores: Ana Beatriz Parra Ferreira, Antônio Vicente Fernandes de Andrade, Maria Fernanda Pizarro Carvalho

Orientador: Adriano Dutra Teixeira

Resumo

Estudo sobre o efeito da representação política feminina em políticas de combate à violência doméstica. Análise do papel da Delegacia Especializada de Atendimento à Mulher (DEAM) como um canal de combate à violência contra a mulher. Como metodologia, foi utilizada uma Regressão Descontínua (RDD) aplicada em Close Elections.

Palavras-chave: Close Elections, Prefeitas, Violência Doméstica

1 Introdução

A violência contra a mulher é um grave problema social e de saúde pública no Brasil. As manifestações de violência contra a mulher não se limitam apenas às agressões físicas, abrangendo também diferentes formas de violência psicológica, emocional, patrimonial e financeira. Tais violências podem ser analisadas em relação ao espaço de ocorrência, manifestando-se tanto no âmbito privado quanto no público. Em todos os casos, o gênero é o fator comum que as caracteriza, seja em contextos familiares ou em relações de natureza pública, uma vez que as questões subjacentes a essas violências são, fundamentalmente, questões de gênero, envolvendo mulheres ou problemas considerados “femininos” pela cultura vigente (Schraiber, 2005).

Nesse contexto, a promoção do bem-estar feminino em face da violência exige a avaliação dos recursos destinados à mitigação desse fenômeno. A análise da efetividade dos canais de combate à violência contra a mulher, e, em particular, o conhecimento específico sobre esses canais, torna-se de extrema importância e motiva a presente pesquisa. Assim, como será detalhadamente abordado na Seção 2, a participação feminina na política será utilizada para embasar a avaliação de um canal de combate à violência contra a mulher: as Delegacias Especializadas de Atendimento à Mulher.

1.1 Representação feminina e seu papel no Combate a Violência de Gênero

É importante trazer à tona a representatividade nos espaços públicos e de liderança no Brasil. A representatividade feminina já é explorada na academia. O paper “Future electoral impacts of having a female mayor” explora como o Brasil elegeu, ao longo das últimas eleições, cada vez mais mulheres para cargos

de prefeitas e como o desempenho positivo das mesmas contribuiu para a eleição de mais mulheres se tornando deputadas e governadoras em eleições federais (Arvate; Firpo; Pieri, 2017). Os resultados do estudo têm implicações para a compreensão de como a eleição de líderes femininas afeta a representação política das mulheres e o potencial de ameaça de estereótipo e voto com viés de gênero.

Consoante à isso, existem evidências sobre o efeito das mulheres na política em índices de violência. O paper “Fighting violence against women: The role of female political representation”, apresenta o feito da gestão de uma prefeita nos casos de violência doméstica em municípios brasileiros (Bochenkova; Buonanno; Galletta, 2023). Conclui-se que em municípios em que a pessoa eleita é mulher, os casos de violência contra as mulheres diminuem, em média, em 50%. Este resultado foi obtido através de uma metodologia que faz uso de uma regressão descontínua, que será utilizada nesta pesquisa e aprofundada na seção 3.2.

A partir dos estudos sobre a participação de mulheres e sua influência na redução de violência feminina, alguns questionamentos são alavancados, como “Quais são os principais canais de combate à violência doméstica?” ou “Quais destes canais são utilizados pelas prefeitas para a redução da violência?”. Responder tais perguntas não é uma tarefa trivial e envolve avaliar diversos possíveis canais de combate, sendo um procedimento que extrapola o escopo da seguinte pesquisa. Dessa forma, o proposta deste estudo é avaliar a efetividade de um provável canal de combate à violência contra a mulher, a Delegacia Especializada de Atendimento à Mulher, detalhada na seção 1.2.

1.2 Delegacia Especializada de Atendimento à Mulher (DEAMs)

As Delegacias Especializadas de Atendimento à Mulher (DEAMs) são unidades brasileiras da Polícia Civil criadas com o objetivo específico de oferecer atendimento especializado e acolhedor para mulheres vítimas de violência doméstica, física, psicológica, sexual, patrimonial e outras formas de abusos.

As DEAMs buscam não apenas a resolução de crimes, mas também a promoção da autonomia das mulheres, oferecendo apoio jurídico, psicológico e social, com o intuito de romper o ciclo de violência e garantir a proteção das vítimas. Criadas pelos governos estaduais a partir de 1985, as DEAMs constituíram-se, até recentemente, na principal, senão muitas vezes a única, política pública de prevenção e combate à violência contra as mulheres (Silva, 2012).

Evidências na literatura apresentam o impacto da implementação de Delegacias Especializadas de Atendimento à Mulher nas taxas de abusos e agressões que acontecem em relações românticas, chamadas formalmente de Intimate Partner Violence (IPV). O paper “Women’s police stations and intimate partner violence: Evidence from Brazil” realiza essa avaliação fazendo uso de um *staggered Diff-in-Diff*, comparando os municípios que implementaram e que não implementaram as delegacias (Perva; Reynolds, 2017).

O estudo conclui que embora não se encontre uma forte associação média, as DEAMs demonstram ser altamente eficazes entre mulheres jovens que vivem em áreas metropolitanas. Estabelecer uma DEAM em um município metropolitano está associado a uma redução na taxa de homicídios femininos de 1,23 mortes por 100.000 mulheres de 15 a 49 anos - aproximadamente uma redução de 17% na taxa de homicídios femininos grandes cidades. Paralelamente, a redução na taxa de homicídios de mulheres de 15 a 24 anos é ainda maior: 5,57 mortes por 100.000 mulheres.

Concatenando a conclusões dos estudos apresentados em 1.1 e 1.2, as ponderações construídas até o momento embasam a hipótese deste estudo: **“A implementação de Delegacias Especializadas em**

Atendimento a Mulher (DEAMs) é um canal de combate à violência doméstica no mandato de prefeitas?”

2 Modelagem Teórica

Na literatura econômica, a unidade familiar é comumente vista como uma unidade cooperativa, na qual os agentes buscam maximizar a utilidade do núcleo familiar de forma “altruísta”. Entretanto, famílias acometidas pela violência doméstica se encaixam mais adequadamente em um jogo estratégico, no qual cada parceiro busca maximizar a própria utilidade, sujeito à restrição de que o cônjuge permaneça na relação. (Farmer; Tiefenthaler, 1997).

É importante frisar que o modelo em questão aborda um contexto presente em unidades familiares tradicionais e que seguem um raciocínio de papéis de gênero que, na contemporaneidade, pode ser visto como obsoleto em outras dinâmicas familiares.

Para entender quais fatores impactam na tomada de decisão dos sujeitos que compõem esta unidade familiar é necessário olhar para as funções de utilidade e restrições orçamentárias de cada cônjuge. Nesta análise econômica da violência doméstica, duas importantes tomadas de decisão têm grande importância: a decisão de quanta violência será aplicada pelo cônjuge e a decisão de terminar o casamento por parte da mulher. Essas tomadas de decisão são reflexões das utilidades de cada cônjuge dentro de um casamento.

No contexto desse modelo, a utilidade refere-se à satisfação ou bem-estar que cada indivíduo obtém das decisões e ações dentro do relacionamento. Cada pessoa tenta maximizar sua utilidade, tomando decisões sobre como agir em relação ao outro, levando em consideração tanto os custos quanto os benefícios de suas escolhas.

2.1 Utilidade do Homem

No modelo, o homem busca maximizar sua utilidade ao interagir com a mulher de uma forma que beneficie seus próprios interesses. Isso pode incluir fatores como controle, poder, ou acesso a recursos dentro do relacionamento. Se o homem acredita que pode aumentar sua utilidade ao exercer controle sobre a mulher — incluindo o uso da violência — ele pode ser incentivado a praticá-la, especialmente se o custo da violência for percebido como baixo (ou seja, se ele acredita que não haverá consequências significativas para suas ações).

A utilidade do homem também está associada à manutenção da relação, à obtenção de benefícios materiais, emocionais ou sociais, e ao controle sobre a mulher, que pode ser interpretado como uma forma de garantir a estabilidade de suas próprias necessidades e interesses. No entanto, é importante notar que, no modelo, a utilidade do homem pode ser afetada por diversas variáveis, como os custos da violência (se houver punições ou problemas sociais associados a ela) e as alternativas disponíveis fora do relacionamento.

Primeiramente, a função utilidade do homem é definida como:

$$U^M = U^M(S(V), C^M, \eta) \quad (1)$$

Em que $S(V)$ representa uma função autoestima que agrega utilidade quando o sujeito reforça uma relação de poder sobre sua parceira através, por exemplo, da violência (V) empregada sobre ela. O η representa o benefício matrimonial, isto é, o benefício de se manter no casamento, e C^M representa o

consumo desse indivíduo que deriva, naturalmente, de sua restrição orçamentária que é definida pelas equações 2 e 3.

$$I^M - t^M = P_c * C^M \quad (2)$$

$$C^M = \frac{I^M - t^M}{P_c} \quad (3)$$

A igualdade na restrição orçamentaria (2) indica que o sujeito só gastará aquilo que ele é capaz de gastar. Ou seja, seu gasto total ($P_c * C^M$) será igual sua renda líquida ($I^M - t^M$). Aqui, I^M representa sua renda bruta, t^M representa as transferências que o sujeito faz para a esposa, como arcar com as contas da casa, por exemplo, P_c é o preço da cesta de consumo e C^M é o seu consumo em si.

Isolando o consumo (C^M) em (3), e o substituindo-o em (1), temos a função utilidade completa desse indivíduo:

$$U^M = U^M(S(V), \frac{I^M - t^M}{P_c}, \eta) \quad (4)$$

2.2 Utilidade da Mulher

Para a mulher, a utilidade está ligada à satisfação e segurança que ela obtém dentro do relacionamento, que pode incluir aspectos como afeto, apoio emocional e segurança financeira. No entanto, o modelo também considera os custos associados à violência doméstica, que afetam negativamente sua utilidade. A mulher pode ser levada a aceitar ou até mesmo a justificar a violência em certos casos, se acreditar que os benefícios da relação superam os custos da agressão, ou se não perceber alternativas viáveis para escapar do abuso (por exemplo, devido à dependência financeira ou medo de retaliações).

Além disso, a utilidade da mulher pode ser moldada pela percepção do risco e pelas normas sociais em torno da violência. Se a mulher acredita que a violência é uma parte aceitável de um relacionamento ou se as instituições não oferecem um suporte adequado, sua utilidade pode ser negativamente impactada, mas ela pode não ter as condições para mudar essa dinâmica.

Analogamente, a utilidade da mulher segue uma distribuição semelhante à do homem,

$$U^W = U^W(V, C^W, \eta, U^M) \quad (5)$$

com o seu consumo (C^W) e o benefício matrimonial (η) presentes. As principais diferenças são que agora, ao invés de uma função autoestima, a violência empregada entra diretamente no cálculo da utilidade de forma negativa (V) e que a utilidade do homem entra no bem estar da mulher.

$$I^W + t^W = P_c * C^W \quad (6)$$

$$C^W = \frac{I^W + t^W}{P_c} \quad (7)$$

$$U^W = U^W(V, \frac{I^W + t^W}{P_c}, \eta, U^M(S(V), \frac{I^M - t^M}{P_c}, \eta)) \quad (8)$$

Outra diferença está na restrição orçamentária da mulher (6), com as transferências do homem agora adicionando à sua renda líquida. Isolando o consumo em (7) e substituindo (7) e (4) em (5) chegamos na equação completa da utilidade da mulher no casamento (8).

2.3 Ponto de Ameaça

O principal ponto da análise é que a mulher, além da utilidade no casamento, pondera a sua utilidade enquanto solteira para tomar a decisão de se manter, ou não, no casamento. O ponto de ameaça da mulher, ou seja, a sua utilidade fora do casamento ($\bar{U}^W(C_S^W)$) é uma função do consumo da mulher fora do casamento, que segue, naturalmente, uma outra restrição orçamentária:

$$I^W + X = P_C * C_S^W \quad (9)$$

$$C_S^W = \frac{I^W + X}{P_C} \quad (10)$$

$$\bar{U}^W\left(\frac{I^W + X}{P_C}\right) \quad (11)$$

Observa-se, em 9, que agora as transferências de renda do marido já não entram na restrição orçamentária da mulher. No lugar delas, tem-se o (X), que representa a Rede de Apoio da mulher fora do casamento, ou seja, familiares ou serviços estatais que dão suporte à mulher caso ela se divorcie, por exemplo.

2.4 O Problema de Maximização

Formalmente, o problema pode ser expresso da seguinte maneira:

$$\max_{V, C^M} U^M(S(V), C^M, \eta) \quad \text{s.a.} \quad U^W = U_S^W \quad (12)$$

A restrição imposta ao problema indica que a mulher só permanecerá no casamento caso sua utilidade dentro da relação seja pelo menos igual à sua utilidade de reserva.


A resolução do problema de maximização, utilizando a técnica do Lagrangeano, conduz a dois resultados fundamentais:

$$\frac{U_V^M}{U_C^M} = -\frac{U_V^W}{U_C^W} \quad (13)$$

$$U^W = \bar{U}^W \quad (14)$$


O primeiro resultado, equação (13), evidencia que, no equilíbrio, as taxas marginais de substituição entre violência e consumo para ambos os cônjuges são iguais em módulo, mas de sinais opostos. Isso significa que um aumento na violência imposta pelo homem gera uma necessidade proporcionalmente maior de compensação em consumo para a mulher, de modo que ela aceite continuar na relação. Esse equilíbrio pode ser visualizado graficamente na Figura 1.

O segundo resultado, equação (14), indica que a utilidade da mulher se mantém exatamente no limite da sua utilidade de reserva. Caso ocorra qualquer alteração que aumente essa utilidade externa



relatorios/close_elections/graficos/micro1.pdf

Figura 1: Representação gráfica do equilíbrio inicial



relatorios/close_elections/graficos/micro2.pdf

Figura 2: Deslocamento do equilíbrio após o aumento da rede de apoio

ao casamento, como o fortalecimento de redes de apoio, o ponto de tangência na curva de equilíbrio se desloca.

2.4.1 Impacto do Aumento da Rede de Apoio

A presença de uma rede de apoio mais ampla e eficaz para a mulher altera a equação de equilíbrio. Na prática, isso reduz a quantidade de violência que ela está disposta a tolerar dentro da relação, tornando o divórcio uma opção mais viável. Esse efeito pode ser observado na Figura 2, onde o novo ponto de equilíbrio se desloca para uma situação em que a mulher aceita níveis menores de violência antes de optar por sair do casamento.

3 Modelagem Empírica

3.1 Close Elections - Eleições Acirradas

A análise de Close Elections é fundamental para esta pesquisa, pois permite avaliar os impactos causais da competição eleitoral sobre políticas públicas e respostas institucionais, como a atuação das Delegacias Especializadas de Atendimento à Mulher (DEAMs). Em conjunto com a metodologia de Regressão Descontínua (RDD), discutida na Seção 3.2, eleições municipais em que a margem de vitória entre dois candidatos foi estreita foram utilizadas, de modo a explorar os efeitos da competitividade eleitoral.

O conceito de Close Elections baseia-se na intuição de que, à medida que a diferença de votos entre o vencedor e o perdedor se reduz, a eleição torna-se mais competitiva e seu resultado mais probabilístico (Grimmer *et al.*, 2011). Isso significa que, em disputas muito acirradas, pequenas variações no comportamento do eleitorado podem determinar o vencedor, tornando o desfecho eleitoral quase aleatório. Esse fenômeno confere ao estudo de Close Elections um caráter de experimento natural, pois permite comparar contextos eleitorais semelhantes em que o único fator determinante da vitória foi uma pequena variação na votação.

A aleatoriedade inerente às eleições marginais é frequentemente utilizada como suposição estatística útil em inferência causal. Em um cenário com apenas dois candidatos, uma eleição altamente competitiva é aquela em que ambos possuem chances equivalentes de vencer. No limite, quando a margem eleitoral se aproxima de zero, o resultado pode ser interpretado como se fosse determinado pelo lançamento de uma moeda justa. Esse princípio embasa a metodologia de RDD, uma vez que torna as características individuais dos candidatos, partidos e distritos ortogonais ao desfecho eleitoral. Assim, é possível estimar efeitos causais de forma mais confiável, sem as preocupações usuais com endogeneidade que afetam estudos eleitorais convencionais (Grimmer *et al.*, 2011).

Neste estudo, a seleção das eleições marginais foi realizada a partir dos microdados do Tribunal Superior Eleitoral (TSE), permitindo identificar disputas com pequenas margens de vitória. Adicionalmente, os dados foram combinados com registros administrativos de violência de gênero, provenientes do Data-SUS e de órgãos governamentais, para investigar se mudanças políticas decorrentes de eleições acirradas afetam as taxas de violência doméstica e a atuação das DEAMs. Também foram construídas estatísticas descritivas para examinar o perfil dos candidatos e das eleições analisadas, incluindo a proporção de homens e mulheres concorrendo e a distribuição partidária dos candidatos.

A abordagem baseada em Close Elections e RDD permite, assim, avaliar de maneira robusta os efeitos da competitividade eleitoral sobre políticas de enfrentamento à violência de gênero, minimizando

problemas de viés de seleção e confundimento. Os resultados obtidos podem contribuir para um melhor entendimento sobre como disputas eleitorais influenciam decisões institucionais e políticas públicas voltadas para a proteção das mulheres.

3.2 Regressão Descontínua - RDD

A Regressão Descontínua (RDD) é uma metodologia econométrica amplamente utilizada para inferência causal em cenários onde existe um ponto de corte bem definido separando unidades tratadas e não tratadas. No contexto eleitoral, essa abordagem explora a descontinuidade gerada pela margem de vitória mínima em eleições acirradas para identificar efeitos causais da escolha de um candidato sobre diferentes desfechos políticos e sociais.

O principal pressuposto do RDD é que, ao redor do limiar de vitória, os municípios que elegeram um candidato e aqueles que elegeram seu adversário são comparáveis, pois pequenas variações na votação são essencialmente aleatórias. Isso permite que a eleição marginal funcione como um experimento natural, no qual se pode inferir o impacto da vitória de determinado grupo político sem os vieses normalmente presentes em estudos observacionais (Grimmer *et al.*, 2011).

Para garantir a validade da inferência causal, são adotadas técnicas como a escolha otimizada do *bandwidth*, testes de balanceamento das covariáveis e análise de sensibilidade para verificar a robustez dos resultados. Além disso, gráficos de descontinuidade são utilizados para inspecionar visualmente os efeitos estimados, complementando a análise estatística.

Dessa forma, a aplicação do RDD nesta pesquisa permite isolar os efeitos da competição política sobre a violência de gênero e avaliar como mudanças no governo local impactam na efetividade das DEAMs em combater a violência doméstica.

A especificação da regressão é escrita da seguinte forma:

$$y_{t,\mu} = \beta_0 + \gamma D_g + f(Mv_\mu) + x_{t,\mu}^\mu + x_{t,\mu}^p + \varepsilon_{t,\mu} \quad (15)$$

Onde $y_{t,\mu}$ representa o número de homicídios de mulheres no município μ no período t . O termo β_0 denota o intercepto da equação, a variável D_g indica a eleição de uma mulher como prefeita, sendo um fator de interesse para o estudo, e γ representa justamente o efeito causal dessa variável sobre os homicídios de mulheres. A função $f(Mv_\mu)$ captura a influência da margem de vitória entre prefeitas e prefeitos sobre os homicídios. Os termos $x_{t,\mu}^\mu$ e $x_{t,\mu}^p$ representam, respectivamente, conjuntos de características municipais e dos candidatos. $\varepsilon_{t,\mu}$ é o termo de erro, que captura a variação nos homicídios de mulheres não explicada pelas variáveis do modelo.

3.3 Dados

A pesquisa utiliza um conjunto de dados detalhado sobre eleições municipais no Brasil, combinando informações eleitorais com registros administrativos de violência de gênero. Os dados utilizados foram obtidos de diferentes fontes.

As informações sobre eleições municipais foram extraídas da base de dados do Tribunal Superior Eleitoral (TSE), acessadas por meio da biblioteca **electionsBR**. Esses dados incluem o total de eleitores, a votação por candidato e as margens de vitória em diferentes pleitos. Para complementar a análise, foram incorporadas características socioeconômicas municipais obtidas a partir de bases do IBGE, permitindo

controlar fatores estruturais que possam influenciar tanto os resultados eleitorais quanto os índices de violência.

No que diz respeito aos dados de violência, a pesquisa utiliza registros de homicídios como proxy para a violência de gênero. Essa escolha se justifica pelo fato de que homicídios, especialmente feminicídios, são menos sujeitos à subnotificação em comparação a outros tipos de violência doméstica, como agressões e ameaças. Os dados de homicídios foram extraídos do sistema de informações do DataSUS, garantindo maior confiabilidade na mensuração do fenômeno analisado.

Além disso, os dados sobre a atuação das Delegacias Especializadas de Atendimento à Mulher (DEAMs) foram obtidos por meio do Ministério da Mulher, a partir de um conjunto de informações disponibilizado por pesquisadoras contatadas durante a realização deste projeto. Esses dados permitem avaliar a presença das DEAMs nos municípios ao longo do tempo.

Os dados eleitorais foram processados para identificar eleições acirradas, definidas como aquelas em que a diferença de votos entre o vencedor e o segundo colocado está dentro de um intervalo específico. A variável de interesse, que mede a intensidade da competição eleitoral, será posteriormente utilizada na análise de regressão descontínua. Além disso, os registros de violência foram organizados de maneira a permitir a identificação de tendências antes e depois das eleições, viabilizando a avaliação de possíveis impactos das mudanças políticas na resposta institucional e na incidência de violência doméstica.

Foram realizadas diversas etapas de tratamento e limpeza. Primeiramente, foram selecionadas apenas eleições municipais com margem de vitória estreita, uma vez que esse critério permite uma análise mais robusta dos efeitos causais. Em seguida, os indicadores de violência foram normalizados, possibilitando a comparação entre diferentes municípios e ao longo do tempo. Por fim, as diversas fontes de dados foram integradas em uma única base consolidada, permitindo a realização de análises econométricas que associam os resultados eleitorais a mudanças nos padrões de violência e na resposta institucional das DEAMs.

A partir dessa base estruturada, o método de RDD foi aplicado para avaliar a relação entre a característica eleitoral e a atuação das DEAMs, bem como os impactos nas estatísticas de violência doméstica.

4 Resultados

Aplicando as metodologias descritas na modelagem empírica, os resultados obtidos evidenciaram um impacto positivo da eleição de prefeitas na redução dos homicídios em ambientes domésticos. A estimação utilizando RDD permite identificar um efeito causal direto da representatividade feminina na política sobre a violência contra as mulheres.

A Tabela 1 apresenta os coeficientes estimados utilizando as variáveis de interesse homicídio doméstico e margem de vitória feminina nas eleições, mostrando que há uma redução estatisticamente significativa de aproximadamente 1,2 homicídios por 100 mil habitantes no mandato quando uma mulher é eleita prefeita.

A Figura 3 ilustra essa descontinuidade na margem eleitoral, reforçando a hipótese de que a presença de prefeitas reduz a violência (homicídios) contra as mulheres. Observa-se uma quebra na tendência dos homicídios domésticos ao longo da margem de vitória feminina, sugerindo que a estimativa reflete um efeito causal válido, interpretado como um Local Average Treatment Effect (LATE) para os municípios próximos ao limiar de vitória eleitoral.

Para validar a robustez dessa relação causal, foi conduzido um teste placebo analisando o impacto da

| Método | Coef. | Erro Padrão | z | P-valor | [95% C.I.] |
|----------------|--------|-------------|--------|---------|-------------------|
| Convencional | -1.114 | 0.502 | -2.221 | 0.026 | [-2.097 , -0.131] |
| Viés-Corrigido | -1.246 | 0.502 | -2.483 | 0.013 | [-2.229 , -0.262] |
| Robusto | -1.246 | 0.566 | -2.199 | 0.028 | [-2.356 , -0.136] |

Tabela 1: Resultados da estimação por RDD para homicídios domésticos.




Figura 3: Visualização do RDD - Efeito da eleição de uma prefeita sobre homicídios domésticos.

eleição em um período anterior, representado na Figura 4. Como esperado, o efeito não é significativo, indicando que a redução nos homicídios ocorre apenas após a eleição da prefeita, e não antes.

Além disso, a Figura 5 apresenta um teste de robustez, demonstrando que o Local Average Treatment Effect (LATE) se dilui a medida que a largura de banda aumenta. O que, novamente, reforça a validade da suposta causalidade encontrada.


Em relação às Delegacias Especializadas de Atendimento à Mulher (DEAMs), não foi possível rodar a regressão descontínua para avaliar seu impacto devido ao número reduzido de municípios que implementaram essas delegacias. Como mostrado nas Tabelas 2 e 3, do total de municípios, 577 não possuíam DEAMs em funcionamento durante os três mandatos analisados. Além disso, um número muito pequeno de municípios que apresentaram close elections implementaram as DEAMs durante as eleições de 2004, 2008 e 2012, o que dificultou uma análise estatística robusta. Os detalhes sobre os prefeitos participantes de eleições próximas e que implementaram as DEAMs estão apresentados na Tabela 3.

Portanto, a relação entre a presença de DEAMs e a redução dos homicídios domésticos permanece inconclusiva, uma vez que não foi possível estimar esse efeito de maneira estatisticamente válida como planejado no início desta pesquisa.



relatorios/close_elections/graficos/rdd_plot_t_1.pdf

Figura 4: Teste Placebo: Efeito da eleição anterior sobre homicídios domésticos (t-1).



relatorios/close_elections/graficos/robustez.pdf

Figura 5: Teste de robustez.

| Número de DEAMs | Frequência |
|-----------------|------------|
| 0 | 577 |
| 1 | 34 |
| 2 | 4 |

Tabela 2: Quantidade de DEAMs por município brasileiro

| Ano | Mulher eleita | Homem eleito |
|------|---------------|--------------|
| 2004 | 0 | NA |
| 2008 | 2 | 1 |
| 2012 | 6 | 6 |

Tabela 3: DEAMs implementadas por prefeitos e prefeitas nos anos analisados

5 Conclusão

A replicação da metodologia utilizada no estudo original confirmou que a eleição de prefeitas leva a uma redução significativa na taxa de homicídios, tanto no âmbito geral quanto no contexto da violência doméstica. Os resultados robustos obtidos pela Regressão Descontínua corroboram a tese de que a representatividade feminina no executivo municipal pode influenciar políticas públicas que reduzem a violência contra as mulheres.

No entanto, a análise das Delegacias Especializadas de Atendimento à Mulher (DEAMs) não permitiu uma conclusão definitiva sobre sua efetividade como canal de combate à violência doméstica. A quantidade insuficiente de municípios com delegacias implementadas impossibilitou a aplicação do método de RDD para estimar o impacto dessas unidades sobre os homicídios. Esse resultado sugere que as DEAMs não são, até o momento, um canal amplamente utilizado para combater a violência doméstica em nível municipal.

Para pesquisas futuras, recomenda-se explorar quais mecanismos específicos são empregados por prefeitas para reduzir a violência de gênero, bem como investigar outras políticas públicas que possam desempenhar esse papel.

Referências

- ARVATE, Paulo; FIRPO, Sergio; PIERI, Renan. Future electoral impacts of having a female mayor. **Brazilian Political Science Review**, SciELO Brasil, v. 11, e0001, 2017.
- BOCHENKOVA, Alena; BUONANNO, Paolo; GALLETTA, Sergio. Fighting violence against women: The role of female political representation. **Journal of Development Economics**, Elsevier, v. 164, p. 103140, 2023.
- FARMER, Amy; TIEFENTHALER, Jill. An economic analysis of domestic violence. **Review of social Economy**, Taylor & Francis, v. 55, n. 3, p. 337–358, 1997.

GRIMMER, Justin *et al.* Are close elections random? **Unpublished manuscript**, p. 1–38, 2011.

PEROVA, Elizaveta; REYNOLDS, Sarah Anne. Women’s police stations and intimate partner violence: Evidence from Brazil. **Social Science & Medicine**, Elsevier, v. 174, p. 188–196, 2017.

SCHRAIBER, Lilia Blima. **Violência dói e não é direito: a violência contra a mulher, a saúde e os direitos humanos**. [S. l.]: Unesp, 2005.

SILVA, Kelly. As DEAMs, as corporações policiais e a violência contra as mulheres: representações, dilemas e desafios. **Revista Brasileira de Segurança Pública**, Fórum Brasileiro de Segurança Pública, v. 6, n. 1, p. 132–154, 2012.

O mercado brasileiro de corridas de montanha

Pesquisadores: Iara Vivian e Henrique Puppi

Orientador: Neto Concon

Resumo

O projeto tem como objetivo realizar um estudo sobre o mercado de *trail running* brasileiro, através da construção de um banco de dados com informações relevantes para o mercado de corrida de montanha no Brasil. Dessa forma, será possível entender as proporções desse mercado atualmente e traçar o perfil do corredor de montanha no Brasil.

Palavras-chave: Construção de base, corrida de montanha

1 Introdução

A corrida de montanha (ou trail running) é uma vertente da corrida tradicional praticada fora dos ambientes urbanos, como: serras, montanhas e trilhas. Essa modalidade tem crescido no mundo, conforme mostrado na Figura 1.

Figura 1: Número de corridas ao longo dos anos no mundo.

relatorios/paraty/figuras/image1.png

Figura 2: Distribuição geográfica das corridas pelo mundo.



Já no Brasil, a corrida de montanha ainda é pouco difundida se comparada a outros países que já tradicionalmente a praticam, como evidenciado pela Figura 2.

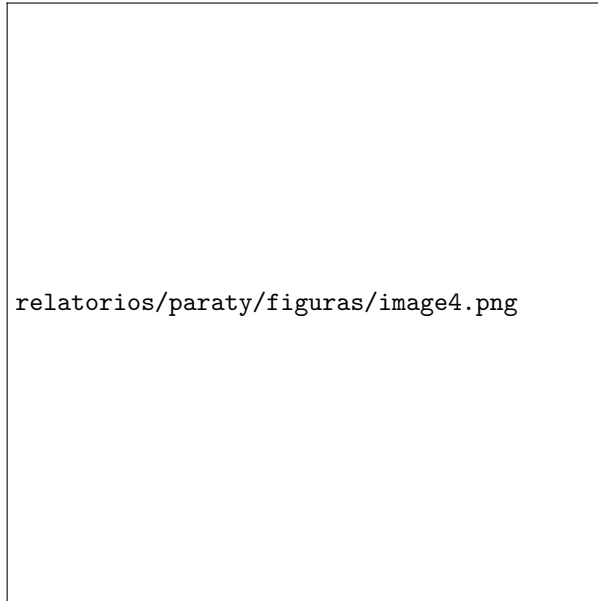
Em outubro de 2016, o Brasil participou pela primeira vez do Campeonato Mundial de Trail Running (Trail World Championship – TWC) que foi realizado em Portugal, a delegação brasileira contou com 6 atletas que competiram contra as outras 39 delegações presentes, somando um total de 302 corredores na competição (ULTRA TRAIL DU MONT BLANC, 2016). Ademais, também em 2016, 90 atletas brasileiros participaram do maior evento de trail running do mundo, o Ultra Trail du Mont Blanc (UTMB), que ocorre anualmente na França, Itália e Suíça, que contava na época com mais de 8000 corredores por ano (Oliveira Junior, 2017).

Apesar do crescimento da modalidade no Brasil, ainda são escassas as pesquisas realizadas sobre o assunto, dificultando, assim, a exploração desse mercado novo que se expande pelo país. Segundo Daniel Ramos, o seu artigo foi o primeiro relacionado à trail running escrito em língua portuguesa e, essa carência de materiais a respeito do mercado, acarreta uma dificuldade para criar competições e circuitos ao redor do país pois organizadores de corridas brasileiros precisam de uma estimativa de participantes para ter rentabilidade econômica, uma vez que é necessário alugar local, comprar garrafas de água, equipamentos, faixas etc. Além disso, a falta de informação prejudica os próprios atletas a encontrarem provas futuras para participarem (Daniel JUNIOR, 2017).

O UTMB apresentou em seu portfólio a Paraty Brazil, primeiro circuito oficial do esporte no Brasil, que ocorreu entre os dias 22-24 de Setembro de 2023 (ULTRA TRAIL DU MONT BLANC, 2023). Hoje a UTMB conta com 4 corridas no Brasil, todas promovidas pela Paraty, porém com diferentes quilometragens (duas de 20 km, uma de 50 km e uma de 100 km), e com aproximadamente 36,6 mil corredores brasileiros registrados no site.

De acordo com Neto Concon, CSO e corredor da Paraty Brazil, uma das maiores dificuldades de

Figura 3: Corrida Paraty Brazil 2023



organizar um circuito de trail running no Brasil é a expectativa numérica de participantes. O CSO também mencionou situações em que as vagas disponibilizadas e esperadas para uma corrida foram preenchidas em menos de 1 semana, mostrando um exemplo da falta de pesquisa de mercado.

Assim, o presente artigo tem como objetivo realizar um estudo sobre o mercado de trail running brasileiro, através da construção de um banco de dados com informações relevantes para o mercado de corrida de montanha no Brasil. Dessa forma, será possível entender as proporções desse mercado atualmente e traçar o perfil do corredor de montanha no Brasil.

2 Desenvolvimento

2.1 Revisão Literária

A construção de uma base de dados para auxiliar a análise do mercado de corridas de montanha no Brasil é um empreendimento novo, e que exige um entendimento das práticas de pesquisa de mercado. Por ser um mercado novo e crescente na cultura brasileira, não há indícios de uma central de informações estruturada, e essa é a lacuna a ser preenchida.

Segundo o livro "Market Research in Practice" de Paul Hague, Matthew Harrison, Julia Cupman e Oliver Truman, a pesquisa de mercado é fundamental para coletar dados relevantes que ajudam a entender melhor o comportamento dos consumidores, identificar tendências e oportunidades, e tomar decisões estratégicas informadas. Este livro destaca a importância de uma abordagem metódica e sistemática na coleta e análise de dados, enfatizando que a precisão e a relevância das informações coletadas são cruciais para o sucesso de qualquer análise de mercado (Hague *et al.*, 2021).

O trabalho de Hague *et al.* (Hague *et al.*, 2021) ressalta que uma pesquisa de mercado eficaz deve começar com uma definição clara dos objetivos de pesquisa e uma compreensão detalhada do mercado-alvo. A coleta de dados pode envolver uma combinação de métodos quantitativos e qualitativos, como

pesquisas, entrevistas, e análise de dados secundários disponíveis. A metodologia recomendada no livro, pode ser aplicada diretamente à iniciativa de construção de uma base de dados, garantindo que a coleta de informações seja abrangente e precisa.

Complementando essa perspectiva, o artigo "Pesquisa de Mercado" de Hélio de Avellar Teixeira (UFMG) enfatiza a importância da pesquisa de mercado na identificação de oportunidades e ameaças no ambiente competitivo. Teixeira (Avellar Teixeira, 2012) argumenta que a pesquisa de mercado é uma ferramenta estratégica essencial, auxiliar à qualquer organização que deseja se posicionar de forma eficaz no mercado. No caso do mercado de corridas de montanha, a pesquisa pode revelar insights valiosos sobre as preferências dos corredores, os fatores que influenciam suas escolhas de eventos, e as lacunas no mercado que podem ser exploradas para desenvolver novos serviços ou melhorar os existentes.

Teixeira (Avellar Teixeira, 2012) também destaca a necessidade de uma análise cuidadosa dos dados coletados, para garantir que as conclusões tiradas sejam válidas e aplicáveis. Ele sugere que, o uso de técnicas estatísticas e ferramentas de análise de dados, pode melhorar significativamente a qualidade das inferências feitas. Este ponto é particularmente relevante para o projeto, pois a construção de uma base de dados completa, e a aplicação de análises, permitirão a identificação de padrões e tendências que podem não ser imediatamente evidentes.

Para a construção de uma base de dados para análise do mercado de corridas de montanha no Brasil, deve-se seguir uma rota metodológica detalhada, conforme recomendado pela literatura especializada em pesquisa de mercado. Segundo o livro de Hague et al. (Hague *et al.*, 2021), a estruturação deve começar com o dimensionamento do mercado, coletando dados imprescindíveis para a exploração. Hélio de Avellar Teixeira, no artigo "Pesquisa de Mercado", enfatiza a importância de identificar as necessidades dos envolvidos, para aumentar a eficiência do processo (Avellar Teixeira, 2012). Questões de pesquisa específicas para a coleta de dados também são essenciais, uma coleta sem alvo não fornece ajuda. Após a coleta, utilizam-se técnicas estatísticas para análise, garantindo que as conclusões sejam válidas e aplicáveis. A etapa final envolve a integração dos dados coletados na construção de uma base de dados completa e acessível, que será mantida e atualizada regularmente, conforme recomendado por Hague et al. (Hague *et al.*, 2021).

2.2 Análise teórica

A organização de um evento de corrida de montanha envolve uma série de fatores econômicos que são essenciais para o sucesso do evento. Entre esses fatores, destaca-se a necessidade de uma infraestrutura adequada, que inclui a instalação de aid stations, que são pontos de apoio ao longo do percurso onde os corredores podem reabastecer-se com água, alimentos e receber assistência médica, caso necessário.

Além das aid stations, a organização de uma corrida de montanha exige a aquisição de equipamentos diversos e a implementação de uma infraestrutura que garanta a segurança e a experiência dos participantes. Devido à natureza desses eventos, realizados em trilhas e ambientes montanhosos, é imprescindível uma sinalização clara e adequada, além de medidas de segurança adicionais para mitigar os riscos associados ao terreno.

Para garantir a viabilidade econômica do evento, é fundamental um planejamento detalhado, incluindo a estimativa do consumo de água por participante, por exemplo. Com base no número esperado de competidores, a organização pode calcular a quantidade necessária de garrafas de água a serem adquiridas, evitando assim desperdícios e prejuízos financeiros.

Nesse contexto, a análise das variáveis envolvidas torna-se crucial para a compreensão do mercado de corridas de montanha. As variáveis estudadas neste trabalho serão utilizadas em análises posteriores e podem ser divididas em duas categorias principais: as relacionadas aos corredores e as relacionadas às corridas. As variáveis relacionadas aos corredores incluem faixa etária, nome completo, URI (Identificador Uniforme de Recurso), nacionalidade, sexo, time, patrocinador, descrição e categoria da corrida em que participam (como 20K, 50K, 100K e 100m). Já as variáveis relacionadas às corridas incluem nome da corrida, URI, título, cidade ou país, dia, mês, ano, categoria da corrida, distância e ganho de elevação.

2.3 Análise descritiva e Dados

Dado que o mercado de corridas de montanha é ainda pouco explorado e estudado, este trabalho busca consolidar dados relevantes sobre essas competições em uma base unificada, a fim de possibilitar estudos que contribuam para o entendimento desse segmento. A fonte dos dados utilizados na criação desta base é o site da UTMB. Os dados sobre os corredores e as corridas são coletados a partir dos eventos realizados pela UTMB, sendo incluídos na base sempre que um corredor participa de um desses eventos.

A presente análise não adota uma abordagem holística no que se refere à utilização de múltiplas fontes de dados disponíveis online. Embora tenham sido identificadas outras bases de dados relevantes, como a da International Trail Running Association (ITRA - International Trail Running Association, 2024), optou-se por realizar uma análise unilateral, limitando-se a uma única fonte de dados.

É fundamental discutir a precisão e a completude dos dados utilizados nesta análise. Apesar de a base de dados selecionada fornecer informações valiosas sobre os corredores de montanha, há lacunas importantes que podem influenciar a interpretação dos resultados. Por exemplo, a ausência de variáveis como a transmissão televisiva dos eventos, que poderia impactar significativamente a visibilidade e o interesse do público, limita a abrangência das conclusões. A exclusão de dados relevantes, seja por indisponibilidade ou por limitações da base selecionada, exige cautela na extrapolação dos resultados. Essa limitação deve ser reconhecida e levada em consideração na interpretação dos achados e na formulação de recomendações subsequentes.

Os dados disponíveis no site da UTMB abrangem competições realizadas desde o ano 2000 até o momento em que esta pesquisa foi conduzida. Para a análise, foram extraídos dados de eventos que ocorreram entre 2000 e meados de junho de 2024. A coleta abrangeu todas as corridas registradas no site, incluindo competições de diversas partes do mundo. Especificamente, todas as informações disponíveis sobre corredores brasileiros foram incluídas, enquanto para os corredores de outras nacionalidades, foram extraídos dados suficientes apenas para permitir comparações quantitativas.

2.4 Metodologia

Para combater o problema de falta de dados estruturados sobre corridas de montanha no Brasil, foi criado um script em Python utilizando a técnica de web scraping. O web scraping é um método automatizado de coleta de dados da web. Esse processo envolve a utilização de scripts que acessam páginas da rede, extraem informações específicas e as armazenam de forma organizada. No caso deste projeto, o script foi programado para visitar websites de eventos e plataformas de inscrição de corridas de montanha, coletando dados como: informações sobre os participantes, localização das corridas, datas dos eventos, e outras variáveis relevantes. O web scraping é uma ferramenta utilizada para coletar grandes volumes de dados que não estão organizados em arquivos, ou apenas, não disponíveis para o público geral.

Figura 4: Diagrama Entidade-Relacionamento do projeto.



Os dados extraídos pelo script de webscraping foram então armazenados em uma base de dados SQL. SQL, que significa *Structured Query Language*, é uma linguagem padrão para gerenciamento de bases de dados relacionais. Utilizando uma base de dados SQL, é possível organizar os dados coletados de maneira estruturada, facilitando consultas, atualizações e análises.

A estrutura da base de dados inclui três tabelas principais: "Corredores", "Corridas" e "Corredor-Corrída" como apresentado na Figura 4. A tabela "Corredores" contém informações detalhadas sobre os participantes, incluindo identificador único, faixa de idade, nome completo, entre outras. Já a tabela "Corridas", armazena exclusivamente detalhes sobre os eventos de corrida, como identificador único, nome do evento, local, entre outros.

A tabela "Corredor-Corrída" estabelece o relacionamento entre corredores e corridas, registrando a participação de cada corredor em cada evento. Esta tabela contém o identificador da corrida (id-corrida), o identificador do corredor (id-corredor), o tempo do corredor na corrida (time) e o ranking do corredor na corrida (rank). Este relacionamento se organiza na forma de 1 para N dos dois lados da tabela, uma vez que um corredor corre múltiplas corridas, assim como uma corrida contém diversos corredores. Assim, a junção desses dados permite numerosas análises cruzando os dados de corredores e corridas, facilitando a extração de insights valiosos sobre o comportamento dos participantes e as características das corridas (Figura 4).

3 Resultados

Após a construção da base de dados por meio da técnica de Webscraping, foram realizadas diversas análises a partir dos dados extraídos.

A análise principal focou no crescimento do mercado de corridas de montanha. Conforme apresentado na Figura 5, observa-se um crescimento exponencial no número de corredores participantes ao longo dos

Figura 5: Quantidade de corredores que participaram de corridas de corridas em cada ano

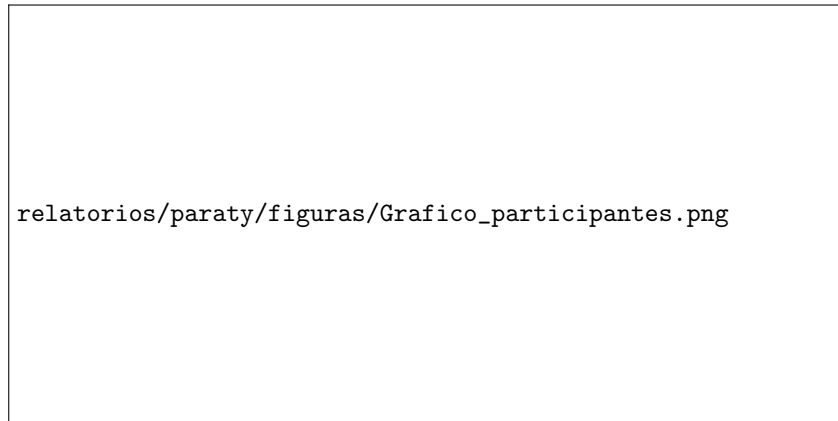
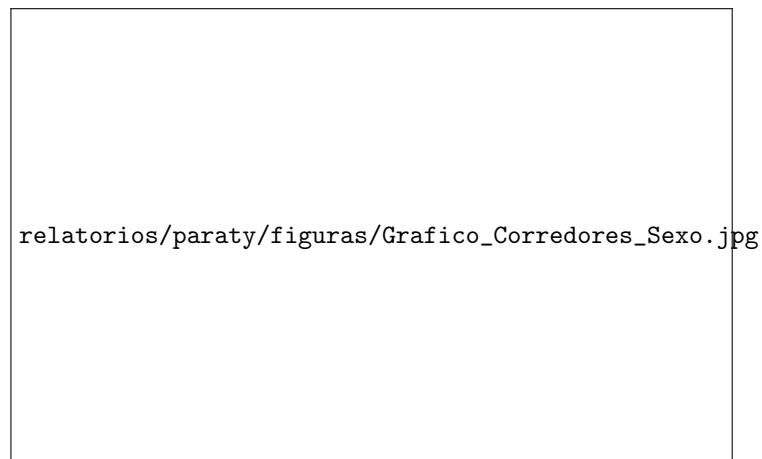


Figura 6: Distribuição de corredores de montanha brasileiros em atividade por sexo



anos, com exceção do período marcado pela pandemia de COVID-19. Esses resultados indicam que o mercado de corridas de montanha está em expansão, reforçando a relevância e o potencial deste estudo.

3.1 Distribuição de Corredores de Montanha por Sexo

O estudo da distribuição de corredores de montanha por sexo, representada na Figura 6, revela uma predominância masculina no cenário atual. Dos corredores analisados, 24.090 são homens, enquanto 12.212 são mulheres, mostrando uma significativa diferença na participação entre os gêneros.

Embora os homens constituam a maioria dos participantes, a presença expressiva de mulheres indica um nível considerável de diversidade nas corridas de montanha. Esses dados são valiosos para entender o perfil atual do mercado e podem ser utilizados para desenvolver estratégias específicas de inclusão e promoção voltadas para ambos os públicos, tanto masculino quanto feminino.

Figura 7: Faixa de Idade dos Corredores Brasileiros



3.2 Faixa de Idade dos Corredores Brasileiros

A distribuição etária dos corredores de montanha brasileiros, ilustrada na Figura 7, apresenta uma tendência clara que segue um padrão semelhante ao de uma distribuição normal, porém ligeiramente deslocada para a esquerda.

Conforme o gráfico, observa-se uma baixa expressividade de corredores nas faixas etárias de 0 a 20 anos e acima dos 75 anos (particularmente nas faixas de 75-79 anos e 80+). Essas faixas indicam uma menor participação de jovens e idosos no cenário das corridas de montanha no Brasil.

Já nas faixas intermediárias, especialmente entre 55 e 74 anos, há uma participação de média expressividade. Esse grupo representa uma parcela considerável de corredores, mas sem o mesmo impacto numérico das faixas mais jovens.

A maior concentração de corredores está nas faixas etárias mais jovens e de meia-idade, variando entre 21 e 54 anos, que apresentam alta expressividade. Esses dados indicam que o perfil típico dos corredores de montanha no Brasil está centrado nas idades produtivas, com um foco particular na faixa dos 30 e 40 anos.

Essa análise permite uma melhor compreensão do público-alvo das corridas de montanha no país, destacando a necessidade de estratégias direcionadas a diferentes faixas etárias para incentivar a participação e retenção em cada grupo.

4 Conclusão

O estudo realizado sobre o mercado de corridas de montanha no Brasil mostrou que, apesar do crescimento dessa modalidade, ainda há uma falta de dados estruturados que possa sustentar o desenvolvimento de novas competições e o aprimoramento de estratégias de marketing e organização. Através da criação de uma base de dados estruturada, foi possível observar o crescente interesse por essa modalidade, bem como identificar as características demográficas dos participantes.

A análise dos dados revelou que o mercado está em expansão, com um número crescente de corredores

participando de eventos ao longo dos anos. Além disso, a distribuição por sexo e faixa etária dos corredores fornece informações valiosas sobre o perfil dos participantes, o que pode ajudar na criação de estratégias mais eficientes de promoção e desenvolvimento do esporte.

Em termos de recomendações futuras, é crucial que mais pesquisas sejam realizadas para complementar os dados existentes, especialmente no que diz respeito a variáveis como os fatores que influenciam a escolha dos corredores por determinadas corridas, além da expansão de fontes de dados e a criação de uma plataforma centralizada para agregar informações sobre as corridas de montanha no Brasil.

Por fim, a construção da base de dados e a análise realizada fornecem uma base sólida para futuras investigações sobre o mercado de trail running no Brasil, permitindo uma compreensão mais profunda do comportamento dos corredores e as dinâmicas desse mercado emergente.

4.1 Limitações

As análises realizadas no presente projeto possuem algumas limitações e vieses que devem ser levados em consideração ao interpretar os resultados. Em primeiro lugar, a unilateralidade das fontes de dados utilizadas é um fator crítico. Todas as informações foram extraídas do site da UTMB, por meio de técnicas de web scraping, o que pode enviesar os resultados e limitar a representatividade dos dados. Dependendo do site, as informações podem não refletir completamente o universo dos corredores de montanha, deixando de lado dados de outros segmentos ou regiões não capturados por essa fonte.

Além disso, há uma questão de precisão e completude dos dados. A fonte utilizada pode não conter todas as informações relevantes para a análise, o que compromete a robustez dos resultados. Dados cruciais, como condições ambientais, variações regionais e detalhes sobre o perfil socioeconômico dos corredores, podem não estar disponíveis, o que limita a capacidade de realizar análises mais profundas ou identificar correlações importantes que poderiam surgir com fontes de dados mais amplas e diversas.

Outro ponto importante a ser considerado é a estrutura da página web. As páginas das quais os dados são extraídos podem sofrer alterações frequentes em seu layout ou na organização das informações. Isso pode, inevitavelmente, causar interrupções no processo de coleta de dados via web scraping, exigindo ajustes recorrentes no script e, em casos extremos, inviabilizando a extração dos dados. A manutenção contínua do código é um desafio que precisa ser monitorado ao longo do tempo.

Reconhecer essas limitações é fundamental para a interpretação cuidadosa dos resultados e para a busca de melhorias futuras no estudo.

4.2 Passos futuros

Para aprimorar a análise e aumentar a precisão dos dados, duas estratégias podem ser implementadas: a padronização da localidade das corridas e a expansão da base de dados.

Atualmente, os dados referentes à localidade das corridas e à nacionalidade dos corredores estão em formatos variados, como países, cidades, estados, coordenadas geográficas e outras designações. Para garantir a consistência e facilitar a análise, será implementada uma padronização dessas informações. Utilizaremos o Google Maps para converter e padronizar todas as localizações para o formato de cidades.

Essa padronização não só melhorará a qualidade dos dados, mas também permitirá uma análise mais coesa e comparativa das informações. Com a localidade padronizada, será possível identificar padrões e tendências de forma mais eficaz, além de comparar dados de diferentes regiões com maior precisão.

Ainda, para enriquecer ainda mais o estudo, planejamos expandir a base de dados para incluir dados de corredores de todas as nacionalidades cadastrados na Ultra-Trail du Mont-Blanc (UTMB), e não apenas corredores brasileiros.

A coleta desses dados globais pode levar aproximadamente 30 dias e envolve a obtenção de informações de cerca de 2,75 milhões de corredores. Esta expansão permitirá analisar tendências e padrões não apenas no Brasil, mas também em uma escala internacional, proporcionando insights valiosos sobre o perfil dos corredores de montanha ao redor do mundo. A inclusão desses dados ampliará a compreensão das características demográficas, preferências e comportamentos dos corredores em diferentes regiões, oferecendo uma análise mais completa e representativa do mercado global de corridas de montanha.

Referências

AVELLAR TEIXEIRA, Hélio de. Pesquisa de Mercado. **Universidade Federal de Minas Gerais**, 2012.

HAGUE, Paul *et al.* **Market Research in Practice**. 4th. [S. l.]: Kogan Page, 2021.

ITRA - INTERNATIONAL TRAIL RUNNING ASSOCIATION. **Find a Runner**. [S. l.: s. n.], 2024. Acessado em: 24 dez. 2024. Disponível em: <https://itra.run/Runners/FindARunner>.

OLIVEIRA JUNIOR, Daniel Ramos. **Mercado do trail running no Paraná: análise dos agentes estruturantes responsáveis pelo desenvolvimento da modalidade**. [S. l.: s. n.], 2017.

Análise da acessibilidade após possível intervenção na Rua Uberabinha

Políticas Públicas

Pesquisadores: Isabella Daccache, João Vitor Heinke, João Vitor Rocha e Lucca Kao

Orientadoras: Laura Janka, Natascha Vital e Heloisa Escudeiro

Resumo

Com o objetivo de transformar os espaços públicos mais verdes e próprios para pedestres e ciclistas, o projeto "Repensando Espaços Públicos – Rua Uberabinha" estuda transformar uma rua local em uma praça pública. Dessa forma, o presente estudo busca investigar as consequências para a acessibilidade das modalidades de transporte (carro, bicicleta e pedestres) após uma possível implementação do projeto localizado no bairro Itaim em São Paulo. A partir desta pesquisa é possível concluir que desconsiderando o trânsito, a área de alcance dos tipos de transporte não seriam afetadas, além de que 30,78% das rotas de pedestres e 61,19% de ciclistas seriam beneficiados ou não sofreriam alterações, enquanto 69,10% das rotas de carro seriam prejudicadas ou não sofreriam alterações.

Palavras-chave: Uberabinha, acessibilidade, isócrona, oportunidade

1 Introdução

A palavra "acessibilidade" em meio a discussão do dia a dia provavelmente acarretará em um ponto de vista sobre acessibilidade de pessoas detentoras de alguma deficiência motora ou cerebral, e de modo visual, a lembrança do quadrado azul com o desenho de um cadeirante. No entanto, a palavra possui cunho mais genérico do que usualmente é utilizada, com outro significado para a palavra, Silva Vieira (2006) apresenta o seguinte conceito: "A acessibilidade é o acesso fácil, qualidade do que é acessível. A falta de acessibilidade no transporte coletivo está associada as grandes distâncias e longas viagens". Dessa forma, o presente estudo concentrará esforços para compreender a acessibilidade viária nos arredores do bairro Itaim Bibi, São Paulo, mais especificamente com ponto central o Insper – Instituto de Pesquisa e Ensino localizado na rua Quatá.

Em 2014, a cidade de São Paulo recebeu um novo direcionamento na sua organização viária e urbana, em outras palavras, atualizou seu Plano Diretor, o qual não era alterado desde 2002. Neste ano, portanto, São Paulo propôs ideias primárias sobre a construção de uma cidade mais inclusiva e acessível para modos de transporte que ultrapassassem o veículo particular, ou seja, a busca por uma maior acessibilidade para pedestres e ciclistas. Propostas essas que enfatizavam a construção de mais áreas verdes, transformação de logradouros em praças públicas, aumentando assim, os espaços de convivência e a segurança viária do público alvo.

Esses aspectos enfatizados no Plano Diretor de 2014, foram matrizes da construção do projeto “Repensando Espaços Públicos – Rua Uberabinha” da Gestão Urbana da cidade de São Paulo. Projeto este que reuniu diversas ideias de transformação no logradouro, buscando acrescentar características as quais são citadas no Plano Diretor. Algumas das ideias apresentadas foram: criação de um espaço público de encontro e convívio; intervenções que garantam segurança na caminhabilidade do local e oferta de comércio e serviços, ações essas que bloqueariam a rua para a passagens de carros, focalizando tal espaço para pedestres e ciclistas.

De tal modo, tendo em vista o fluxo diário na rua local do bairro do Itaim por conta de instituições de ensino nos arredores do logradouro, como, Instituto de Ensino e Pesquisa e Universidade Anhembi Morumbi, este estudo compreende a relevância do impacto de tal decisão e busca esclarecer as consequências para cada uma das modalidades de transporte, tendo como localidade principal de análise o Insper.

2 Revisão Literária

Em busca de exemplos similares a tal decisão, a literatura apresenta duas importantes questões: dificuldade sobre a mensuração do conceito de acessibilidade; compreensão sobre o modo de interação e envolvimento entre o espaço analisado e seus usuários. Isto é, a literatura sobre estruturas viárias possui poucas metodologias elaboradas e aprovadas pela academia sobre a mensuração e comparação entre a acessibilidade de um local e outro, ou uma possível estrutura viária e sua alternativa. Ademais, sob o ponto de vista do urbanismo, decisões como da Rua Uberabinha ultrapassam barreiras de análises absolutas, permeando um âmbito mais humano e que envolve um certo conhecimento empírico.

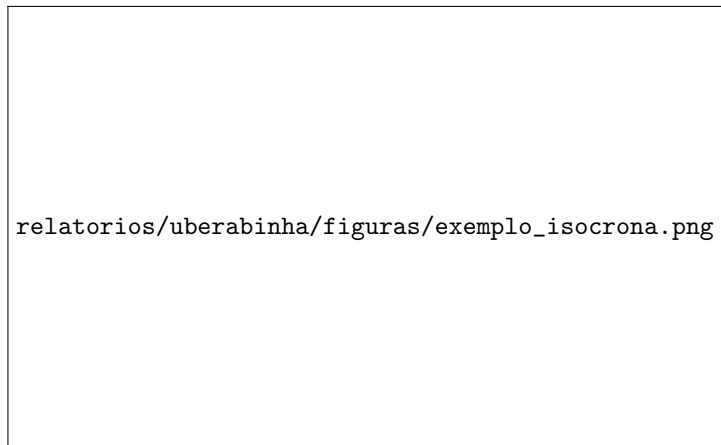
Primeiramente com relação ao entendimento do vínculo entre o espaço e os usuários, o urbanismo atesta que de modo descritivo é necessária a concepção sobre o uso do espaço no momento presente. Em outras palavras, a construção e planejamento de lugares deve ser sob a perspectiva da colaboração, a partir da expertise da comunidade local, que busca seu próprio bem-estar, ponto de vista este dominado a partir da metodologia do *Place Making* Melo Brito Rocha *et al.* (2024).

Tendo em vista que tal método se baseia por completo em competências humanas, para a contagem e compreensão sobre o uso do espaço pelos indivíduos, a validade dos resultados gerados é duvidosa. Isto pois, o foco e assertividade de uma pessoa na contagem de carros, pedestres, ciclistas e motociclistas não é alta, o que cria uma incerteza sobre tal. Dessa forma, é discutível a robustez de estudos baseados no Place Making, com uma possibilidade de melhora de tais a partir do uso de diferentes técnicas de contagem. Jantsch (2020) apresenta de modo conciso e claro essa dificuldade da contagem manual em comparação ao uso de tecnologias, analisando o erro de cada método e seus benefícios e malefícios. Não só ele, como Osinski e Luz França (2020) demonstram outro método mais complexo de contagem, porém mais prático, a partir do uso de algoritmos de visualização de imagens de trânsito.

Por outro lado, com um ponto de vista microeconômico Koenig (1980), apresenta um método de mensuração da acessibilidade de um local a partir de duas maneiras diferentes: isócronas e somatório de oportunidades. O primeiro, isócronas, define de forma cartográfica a área de alcance de uma pessoa, partindo de um ponto de origem utilizando uma modalidade de transporte específica, apresentando uma comparação da área de alcance entre as modalidades. Ou seja, um método que metrifica a acessibilidade como sinônimo do alcance obtido a depender da origem e o transporte utilizado, um exemplo disso é retratado na Figura 1. No exemplo fictício citado, a origem é representada pelo ponto preto, que é rodeado por duas áreas, azul clara e escura, as quais representam o alcance que uma pessoa teria em 10

e 20 minutos de deslocamento, respectivamente, utilizando uma bicicleta.

Figura 1: Exemplo isócrona



Uma definição pragmática para o conceito de isócrona é apresentado pelo Esri (n.d.): “Uma linha em um mapa conectando pontos de tempo percorrido igual, especialmente, tempo de viagem para ou de uma localização fornecida.”

Outro método de análise sobre a acessibilidade levantado por Koenig é a partir da somatória das oportunidades entre uma origem i e seus possíveis destinos j . Essa metodologia é melhor representada pela expressão matemática 16:

$$A_i = \sum_j O_j f(C_{ij}) \quad (16)$$

Com relação aos elementos da expressão o primeiro O_j se refere ao número de oportunidades existentes no destino j . Já a função $f(C_{ij})$, traduz a função de impedância, elemento este que pondera as oportunidades de acordo com a rota ij , utilizada. Desse modo, o somatório de todas as oportunidades ponderadas mensuram a acessibilidade do local i , representado por A_i .

O raciocínio contido em tal expressão é compreendido da seguinte maneira. O ponto de origem i é o local de estudo de interesse, o qual está sendo analisado sua acessibilidade. Após o recorte da área de estudo desejada, seja essa um raio de 3Km ao redor do local de origem, ou uma região de logradouros nos arredores da origem, é necessário o levantamento de todas as oportunidades presentes nessa área. O conceito de oportunidade não é especificado por Koenig em seu estudo, e portanto, para a atual pesquisa este será utilizado em referência a localidades onde uma pessoa poderia interagir e obter um bem-estar/felicidade. Em outras palavras, supermercados, restaurantes, academias, shoppings foram alguns dos estabelecimentos reconhecidos pelos estudos como oportunidades a serem localizadas na área de estudo.

Dessa forma, a análise se aplica sobre cada caminho realizado entre a origem i , o destino j , a quantidade de oportunidades que o indivíduo teria de acesso no destino, sendo toda essa expressão ponderada por uma função de impedância do caminho. A ideia central é de que, ao se deslocar de i para j o indivíduo tenha uma felicidade, e essa felicidade é interferida pelo caminho utilizado, sofrendo prejuízos caso a impedância seja maior. Impedância neste contexto pode ser entendida como resistência, sendo, portanto, um fator inversamente proporcional a acessibilidade de um local, de modo que quanto maior

a impedância em tais rotas, menos acessível seria o local. Os fatores contidos na função de impedância são inúmeros, porém para melhor compreensão da metodologia, exemplos desses fatores são, distância percorrida, tempo do trajeto e segurança da rota. Logo, quanto maior a distância percorrida, ou quanto maior o tempo do trajeto, ou até quanto maior a taxa de assaltos e furtos na rota utilizada, maior a impedância do específico trajeto ij , e assim, menos acessível é o local.

Tendo isso em vista, a hipótese que os modelos de Koenig propõe para a situação de análise sobre a possível intervenção na Rua Uberabinha podem ser divididas em dois grupos. Para o grupo de indivíduos que fazem uso da região baseando-se em transportes de carro, a hipótese é de que a alteração afetaria negativamente ou não teria nenhum efeito sobre a acessibilidade desse tipo de transporte, isto pois, o fechamento da rua mudaria todas as rotas que passavam por tal rua, o que aumenta a distância de deslocamento, e assim, aumentaria a impedância e provavelmente diminuiria a área das isócronas. Contudo, o presente estudo não buscou quantificar essa impedância, mas é claro a partir da metodologia que esse efeito seria negativo ou de magnitude irrisória, sendo impossível um benefício para os carros.

Já por outro lado, para o grupo de pedestres e ciclistas o fechamento da rua para carros traria um efeito positivo na acessibilidade, uma vez que sem a passagem de carros o espaço se torna mais seguro para pedestres e ciclistas que não necessitarão se atentar a possíveis acidentes. Dessa forma, essa alteração teria efeito positivo ou irrisório, não sendo possível prejudicar os pedestres e ciclistas.

Assim sendo, a hipótese a ser testada empiricamente é se essa alteração viária afeta de alguma forma as isócronas de cada uma das modalidades de transporte. Para além dessa hipótese, o mapeamento das oportunidades que estão ao redor do ponto de origem (Insper), quais são acessíveis para cada tipo de transporte em determinado tempo, e quais utilizam a Rua Uberabinha em suas rotas são objetivos de estudo também.

3 Metodologia

Assim como apresentado pela revisão da literatura, o estudo descritivo sobre o local foi feito com o intuito de entender como a população local utiliza o espaço e qual o fluxo de carros, pedestres e ciclistas durante diferentes períodos do dia. Para isso, foram feitas contagens durante períodos de trinta minutos espaçados durante o dia, documentando a quantidade de carros/pedestres/ciclistas que desciam a Rua Quatá ou faziam a conversão na Rua Uberabinha. Os dias de coletas variaram entre duas semanas e diferentes dias da semana, com coletas em quatro dias no total, duas terças-feiras e uma segunda-feira e uma quarta-feira. O intuito de modo geral era buscar entender qual a magnitude do fluxo total diário no local, como também, qual modalidade de transporte mais utilizava o logradouro possivelmente transformado. Esses resultados serão discutidos posteriormente no trabalho, no entanto é interessante ressaltar as incertezas sobre a robustez no modo em que a análise é empregada.

A realização das contagens para compreensão do fluxo atual no local, de forma manual, é questionável, uma vez que a contagem por um ser humano é repleta de falhas, desatenção o que pode tornar os dados viesados. Como já citado, os períodos de coleta de dados possuíam duração de 30 minutos e esses eram subdivididos entre outros três conjuntos de contagem, buscando reduzir os riscos de erro de contagem. Contudo, para o olho humano, o foco durante dez minutos, analisando e contabilizando quatro diferentes tipos de modalidade de transporte: carro; pedestre; motocicleta e bicicleta, é duvidoso.

A contagem em si foi feita com auxílio do aplicativo Multi-counter, e dividida entre duas pessoas, cada um analisando um fluxo específico, seja este o de continuação na Rua Quatá (Ponto 4) ou conversão

na Rua Uberabinha (Ponto 5). As Figuras 2 e 3 apresentam respectivamente o ponto de coleta e a interface do aplicativo utilizado.

Figura 2: Local de contagem do estudo

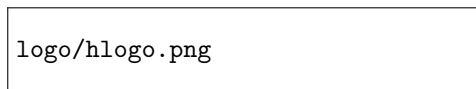
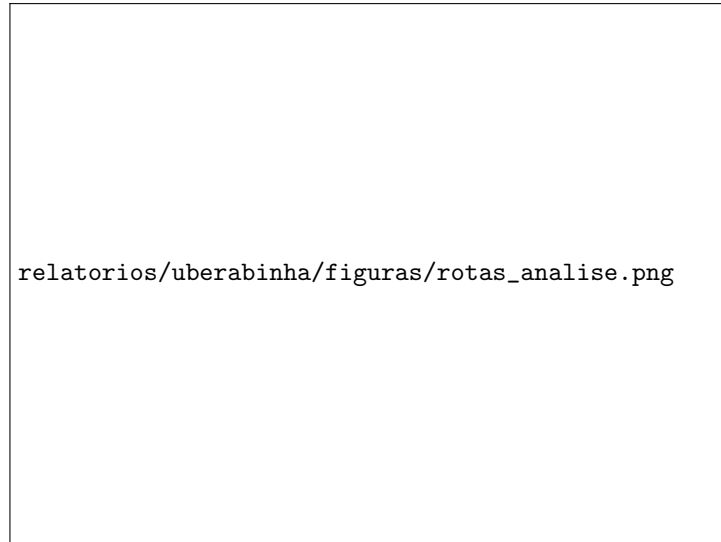


Figura 3: Exemplo das contagens no aplicativo - Multi-counter

Para a elaboração das isócronas e análise de acessibilidade sobre as oportunidades que cada tipo de transporte tem acesso em certo espaço de tempo, a ferramenta do `r5r` foi utilizada. Uma biblioteca da linguagem de programação R, construída por Pereira *et al.* (2021). Essa auxiliou na construção das isócronas e cálculos sobre distância e tempo de deslocamento do ponto de origem até cada uma das oportunidades mapeadas. Com relação a alteração da estrutura viária e construção dos mapas, essas ações foram feitas utilizando o software de informação geográfica QGIS Geographic Information System, a partir dos dados retirados do GeoSampa.

Com relação ao mapeamento de todas as oportunidades ao redor do Insper, foi utilizada a ferramenta Google Places API. Esta ferramenta retorna uma lista de lugares correspondentes às especificações dentro de um raio estipulado. Para maximizar a identificação de oportunidades, foi criado um filtro específico para os tipos de lugares, dada a ampla variedade disponível na API.

Foram selecionados os tipos mais relevantes que englobam: *car_dealer; car_rental; car_repair; car_wash; gas_station; parking; farm; school; university; accounting; bank; restaurant; cafe; bar; coffee_shop; administrative_area_level_1; administrative_area_level_2; fire_station; police; dentist; doctor; drugstore; hospital; pharmacy; hotel; barber_shop; beauty_salon; insurance_agency; laundry; lawyer; real_estate_agency; veterinary_care; market; pet_store; shoe_store; shopping_mall; sporting_goods_store; store; supermarket; wholesaler; fitness_center e gym.*

A pesquisa tomou como premissa um raio de 3 quilômetros como a área de interesse para análise, tendo como centro as coordenadas do Insper. Durante o mapeamento, algumas dificuldades foram enfrentadas,

entre elas o número limitado de requisições permitidas. Para contornar essa limitação, o filtro de tipos de lugares foi crucial, além da criação de múltiplas contas do Google para aumentar o número de requisições. Outro desafio foi a quantidade limitada de oportunidades retornadas pela API para cada tipo de lugar dentro do círculo de 3 quilômetros, sendo apenas 20 lugares por tipo. A solução encontrada foi realizar requisições com círculos de raio menor, preenchendo o todo de 3km de raio.

Dessa maneira, foram utilizados círculos com raios de 200m para mapear toda a área, tendo em alguns espaços a sobreposição de requisições para justamente não existirem áreas não mapeadas. Para evitar duplicidade na lista de oportunidades, foram removidas as entradas repetidas, garantindo que cada oportunidade aparecesse apenas uma vez. Após possuir todas as oportunidades em um raio de 3Km ao redor do Insper, utilizou-se de outra API Google Maps Directions API, que nela é possível retornar o caminho que o Google Maps utilizaria de sugestão para sair de um local X para um destino Y . Realizou-se através de um código Python as requisições necessárias para as 3 diferentes modalidades de transporte presentes na pesquisa: carros, pedestres e bicicletas.

Foi também mapeado, em uma região menos abrangente, a qualidade dos restaurantes da região, buscando compreender a necessidade dos indivíduos de se deslocar mais de 1Km no horário do almoço. Tal detalhamento foi elaborado com auxílio da Google Maps Geocoding API para obter o *place id* de cada local, com base em sua latitude e longitude. Com essa informação, foi possível obter a avaliação de cada restaurante, permitindo assim uma análise de como os melhores restaurantes se distribuem na região.

4 Resultados

Os dados coletados durante as contagens ao final do estudo não foram considerados, isto pois como foi discutido na seção de metodologia, a validade das coletas era questionável. Por conta de uma escolha errônea na distribuição das coletas, a interpolação dos dados, para melhor visualização do fluxo diário foi impossibilitada, tornando a base de dados não robusta.

Por outro lado, a construção das isócronas trouxe resultados interessantes. Nas Figuras 4, 5 e 6, é possível analisar as isócronas de cada modalidade de transporte. De modo descritivo, os resultados são próximos do óbvio, a área de alcance do carro e da bicicleta são muito maiores que a área de alcance do pedestre, simplesmente por conta da velocidade de deslocamento que estes conseguem alcançar. Um ponto necessário a ser enfatizado sobre tais resultados é de que essas isócronas não estão considerando o trânsito como um elemento de análise.

Figura 4: Isócrona Carro sem intervenção na Rua Uberabinha



Figura 5: Isócrona Bicicleta sem intervenção na Rua Uberabinha



Figura 6: Isócrona Pedestre sem intervenção na Rua Uberabinha



Já nas Figuras 7, 8, 9 estão apresentadas as isócronas de cada modalidade de transporte com a possível alteração na Rua Uberabinha, isto é, impossibilidade de carros circularem em tal espaço. Com base na comparação entre as isócronas, é possível compreender que a diferença sobre o alcance dos tipos de transporte quando é feita a alteração na estrutura viária é irrisória, não representando nenhum benefício ou malefício de magnitude considerável.

Figura 7: Isócrona Carro com intervenção na Rua Uberabinha



Figura 8: Isócrona Bicicleta com intervenção na Rua Uberabinha



Figura 9: Isócrona Pedestre sem intervenção na Rua Uberabinha



Sobre a análise descritiva das oportunidades ao redor do Insper, foram utilizados os resultados das rotas oferecidas pelo Maps, para conseguir notar a importância da Rua Uberabinha no deslocamento de cada modal ao redor da área analisada. Dessa forma, foram construídos mapas para cada modalidade de transporte, onde estes separam as oportunidades entre dependentes da Rua Uberabinha ou não, no melhor caminho fornecido pelo Google Maps.

A partir das Figuras 10, 11, 12, pode-se notar a presença de regiões específicas que seriam mais afetadas pela mudança na Rua Uberabinha, com grande diferença entre quais oportunidades seriam afetadas dependendo da modalidade de transporte. Não só isso, como de forma visual, é evidente que os pedestres possuem menos oportunidades dependendo da Rua Uberabinha em seu melhor caminho.

Figura 10: Mapa de oportunidades dependentes da Rua Uberabinha para Carro



Figura 11: Mapa de oportunidades dependentes da Rua Uberabinha para Bicicleta

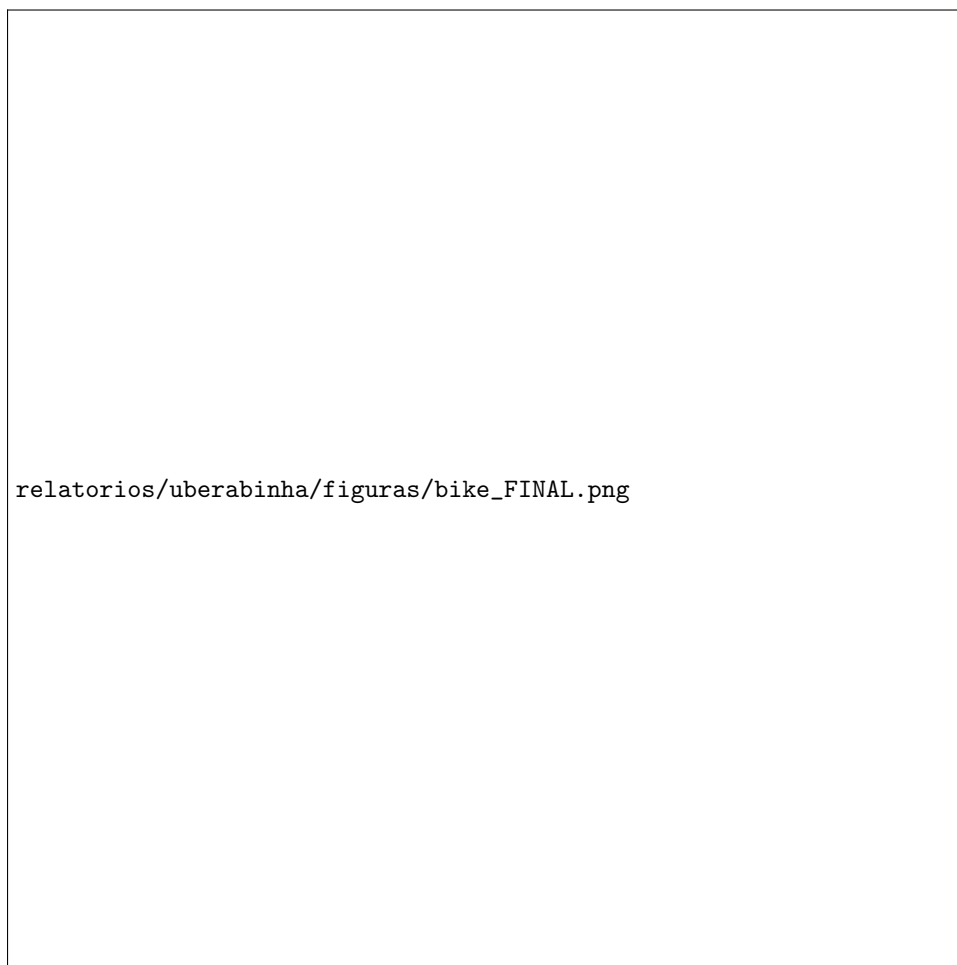


Figura 12: Mapa de oportunidades dependentes da Rua Uberabinha para Pedestre



Partindo para a análise detalhada dos restaurantes em um raio de 1Km do Insper, estes apresentam boa repercussão de acordo com as avaliações do Google Maps. Alguns deles não possuíam notas, no entanto, dentre os detentores de avaliação, esses possuíam sua maior concentração entre o intervalo de 3,5 - 5. Assim sendo, é possível entender que os indivíduos que frequentam diariamente a região não possuem necessidade de grande deslocamentos para realizar uma refeição, pelo menos não pela qualidade do estabelecimentos.

Figura 13: Mapa de restaurantes no raio de 1Km com avaliações do Google Maps

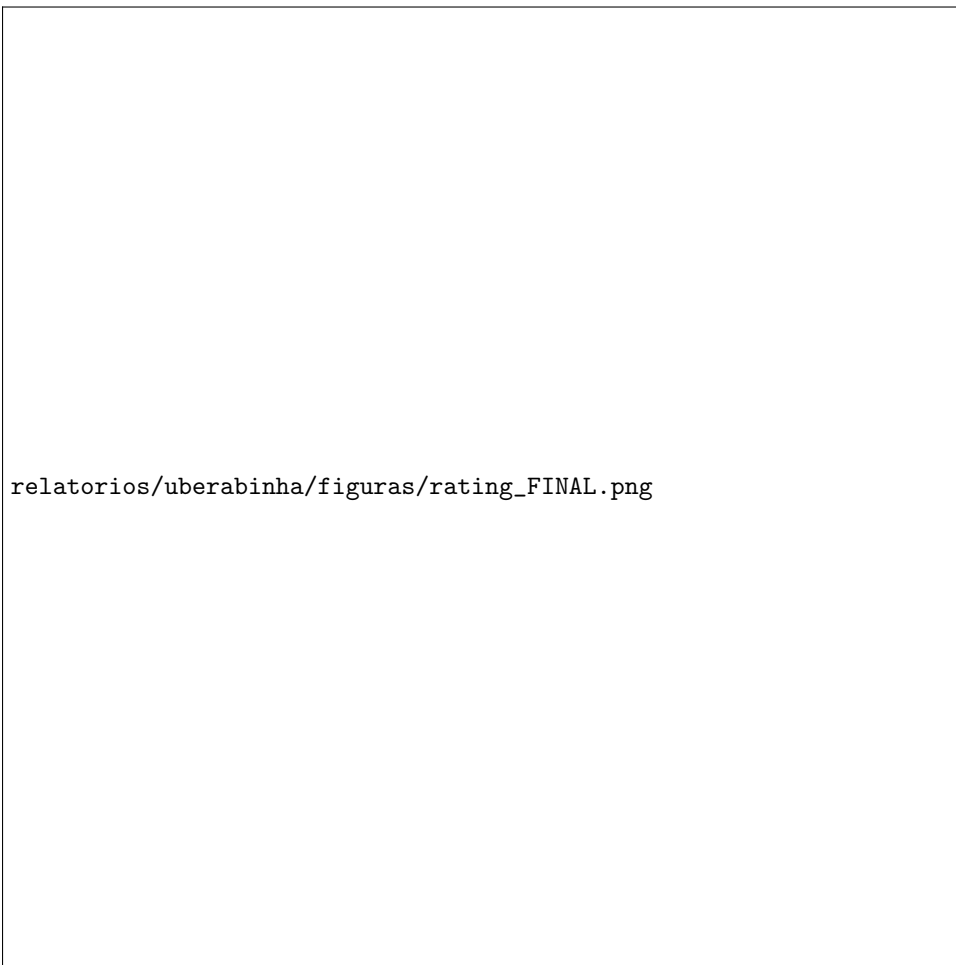


Figura 14: Acessibilidade pedestres para oportunidades selecionadas



Com base no mapeamento de todas as oportunidades, foram selecionados quatro tipos de estabelecimentos - Café, Academia, Supermercado e Restaurante - para uma análise mais aprofundada sobre o acesso de pedestres aos tipos de oportunidade citados. O que foi observado em tal análise é que o acesso dos pedestres as oportunidades selecionadas é considerável, sem necessitar de um deslocamento de mais de 10 minutos, uma vez que neste ponto o indivíduo já teria alcançado: 51 restaurantes; 12 academias; 9 cafés e 6 supermercados.

5 Conclusão

Após análise da hipótese sobre como uma intervenção na Rua Uberabinha afetaria a acessibilidade das três modalidades de transporte: carro, pedestres e ciclistas, é possível concluir alguns aspectos.

Primeiramente, sobre a análise das isócronas antes e depois de uma alteração viária, sem considerar o trânsito na modelagem, é possível concluir que tal alteração não teria impacto significativo na área de alcance de nenhum dos transportes.

Além disso, a partir do mapeamento de todas as oportunidades em um raio de 3Km do Insper, e análise sobre quais oportunidades utilizam a Rua Uberabinha no seu melhor caminho - de acordo com o Google Maps - juntamente com o modelo teórico de Koenig, é possível concluir que: Caso fosse feita uma alteração viária, 30,78% dos caminhos de pedestres e 61,19% de ciclistas teriam a possibilidade de serem beneficiados, enquanto colocaria 69,10% dos melhores caminhos para carros em risco de prejuízos.

Como citado anteriormente, uma limitação de tal estudo é a não implementação da análise do trânsito em nenhum dos resultados, justamente pela extensão da complexidade de tal elemento, o qual não era o foco da pesquisa.

A partir dos resultados encontrados e limitações do estudo, possíveis próximos passos para continuidade de tal análise seriam: simulação do fluxo viário aos arredores do Insper após a intervenção, contabilizando as mudanças de escolha dos indivíduos e o trânsito; simulação de alterações nos sentidos

viários dos logradouros da região em busca do melhor arranjo caso a intervenção ocorresse e compreensão da importância das oportunidades beneficiadas e prejudicadas para cada tipo de transporte, buscando mensurar de forma mais completa quantos indivíduos estariam sendo beneficiados e prejudicados.

Referências

ESRI. **Isochrone**. [S. l.: s. n.], n.d. Accessed: 2024-07-19. Disponível em: <https://support.esri.com/pt-br/gis-dictionary/isochrone>.

GOOGLE. **Google Maps Directions API**. [S. l.: s. n.], n.d. Accessed: 2024-07-19. Disponível em: https://developers.google.com/maps/documentation/directions/?hl=pt_BR.

GOOGLE. **Google Maps Geocoding API**. [S. l.: s. n.], n.d. Accessed: 2024-07-19. Disponível em: https://developers.google.com/maps/documentation/geocoding/?hl=pt_BR.

GOOGLE. **Google Places API**. [S. l.: s. n.], n.d. Accessed: 2024-07-19. Disponível em: https://developers.google.com/maps/documentation/places/web-service?hl=pt_BR.

JANTSCH, Nilson. **Desenvolvimento de um sistema de medição e controle de temperatura utilizando Arduino**. 2020. Trabalho de Conclusão de Curso (Engenharia de Controle e Automação) – Universidade de Caxias do Sul, Caxias do Sul. Accessed: 2024-07-19. Disponível em: <https://repositorio.ucs.br/xmlui/bitstream/handle/11338/5267/TCC%20Nilson%20Jantsch.pdf?sequence=1&isAllowed=y>.

KOENIG, J. G. Indicators of urban accessibility: theory and application. **Transportation**, v. 9, p. 145–172, 1980. Accessed: 2024-07-19. Disponível em: /mnt/data/Indicators%20of%20Urban%20Accessibility_J.G.Koenig.pdf.

MELO BRITO ROCHA, Luís Cláudio de *et al.* Perspectivas para placemaking no Brasil no contexto do design e da arquitetura: revisão da literatura do conceito e análise de casos de referência. **Cuaderno 213, Centro de Estudios en Diseño y Comunicación**, Centro de Estudios en Diseño y Comunicación, p. 173–189, 2024. ISSN 1668-0227.

OSINSKI, Henrique Domiciano; LUZ FRANÇA, Sibilla B. da. Estudo e implementação de algoritmos para contagem de veículos. *In*: SEMINÁRIOS de Microeletrônica do Paraná. Curitiba: UFPR, 2020. Accessed: 2024-07-19. Disponível em: http://eletrica.ufpr.br/microeletronica/pdfs/SeMicro-PR_-_2020_paper_8.pdf.

PEREIRA, Rafael H. M. *et al.* r5r: Rapid Realistic Routing on Multimodal Transport Networks with R⁵ in R. en. **Transport Findings**, mar. 2021. Publisher: Network Design Lab. DOI: 10.32866/001c.21262. Disponível em: <https://findingspress.org/article/21262-r5r-rapid-realistic-routing-on-multimodal-transport-networks-with-r-5-in-r>. Acesso em: 4 mar. 2021.

SÃO PAULO, Prefeitura de. **GeoSampa**. [S. l.: s. n.], 2024. Acessado em: 19-jul-2024. Disponível em: https://geosampa.prefeitura.sp.gov.br/PaginasPublicas/_SBC.aspx.

SILVA VIEIRA, Sérgio Roberto da. Agroindústria canavieira e políticas públicas: análise dos impactos sociais e ambientais. **Sociedade & Natureza**, v. 18, n. 35, p. 55–64, jul. 2006. Accessed: 2024-07-19.

Disponível em:

<https://seer.ufu.br/index.php/sociedadennatureza/article/download/9206/5668/35255>.

TEAM, QGIS Development. **QGIS Geographic Information System**. [S. l.: s. n.], 2024. Open Source Geospatial Foundation Project. Acessado em: 19-jul-2024. Disponível em:

<https://qgis.org/download/>.

6 Apêndice

Figura 15: Mapeamento oportunidades no raio de 3Km



Tabela 4: Mapeamento Oportunidades no raio de 3Km

| | Tipo | Contagem |
|----|---------------|----------|
| 1 | store | 3755 |
| 2 | restaurant | 1817 |
| 3 | beauty_salon | 1035 |
| 4 | school | 692 |
| 5 | car_repair | 619 |
| 6 | bar | 560 |
| 7 | gym | 473 |
| 8 | hospital | 425 |
| 9 | café | 340 |
| 10 | bank | 282 |
| 11 | car_dealer | 238 |
| 12 | market | 142 |
| 13 | university | 120 |
| 14 | supermarket | 94 |
| 15 | gas_station | 70 |
| 16 | shopping_mall | 56 |
| 17 | police | 16 |
| 18 | fire_station | 6 |

Figura 16: Acessibilidade das modalidades de transporte

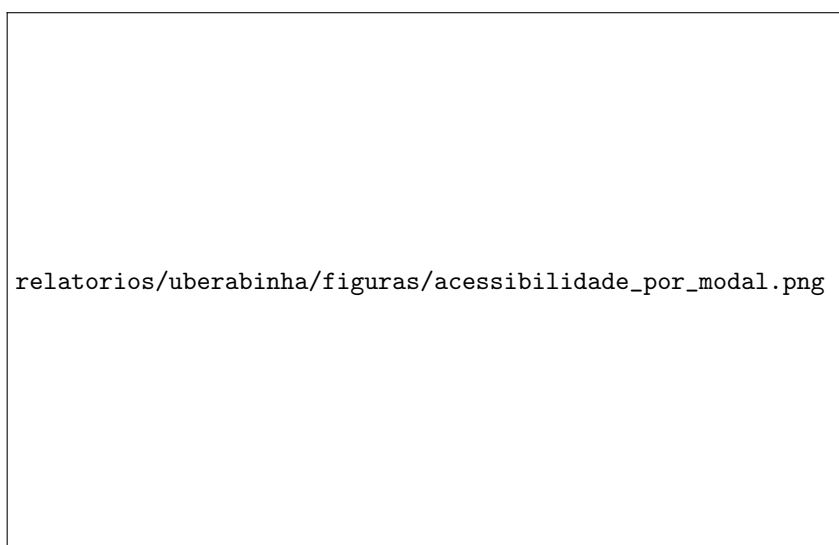


Figura 17: Dados da coleta manual no local de estudo



Consultoria FRST Falconi - Aplicações de simulação de Monte Carlo

Pesquisadores: Arthur Pioli, Felipe Bakowski, Thomas Lipsky, Ykaro Andrade

Orientador: José Renato Garcia Braga

Resumo

Este artigo utiliza métodos estatísticos para investigar frequência de acesso de usuários em uma plataforma para a empresa FRST – Falconi. O objetivo desta foi utilizar as teorias de Cadeia de Markov e Monte Carlo (MCMC) a fim de simular e obter a probabilidade de um usuário transicionar de uma página para um estado de abandono na página, por meio de uma análise de comportamento realizada sabendo, previamente, as probabilidades estacionárias de cada estado. Este estudo é relevante para a FRST- Falconi pois viabiliza um entendimento de quais áreas do site podem ser redesenhadas, mantidas ou excluídas.


Palavras-chave:

1 Introdução

Durante o semestre 2023.2, o Insper Data, em parceria com a FRST Falconi, realizou um projeto cuja finalidade era obter uma análise aprofundada do comportamento de usuários na plataforma própria da FRST utilizando a metodologia das Cadeias de Markov, que fornece um amparo teórico para modelar processos estocásticos. Como resultado deste projeto, os resultados fornecem um modo mais simples de identificar o comportamento dos usuários na plataforma para entender quais etapas no processo de navegação do usuário poderiam se beneficiar de uma mudança ou um redesenho da usabilidade da interface. Além disso, foi possível realizar cálculos dos modelos estacionários das Cadeias de Markov da plataforma e entender onde há maior índice de abandono dos usuários na página.

Ainda, a FRST enfrenta o problema de que não existe uma maneira objetiva de diagnosticar onde pode haver melhorias no âmbito de User Experience - UX. Assim, essa análise visa criar um sistema em que é possível identificar as variações de engajamento dos usuários, a fim de garantir uma melhor experiência para a equipe de desenvolvedores da FRST identificarem possíveis problemas na plataforma e para os usuários.

Com o intuito de realizar uma estimativa mais precisa com os resultados obtidos anteriormente, sem uma limitação de somente uma Cadeia de Markov, mas uma simulação de Monte Carlo aplicada à Cadeia (MCMC) consegue obter resultados mais precisos e aplicáveis ao cotidiano da FRST. Ao aplicar a simulação às Cadeias de Markov, obtém-se uma função distribuição de probabilidade (fdp), justamente da possibilidade da MCMC de prever o comportamento de diversos usuários, a partir de probabilidades estacionárias obtidas por modelos inferenciais.



relatorios/frst/figuras/markov.png

Figura 1: Exemplo de Cadeia de Markov

O objetivo do projeto é dado a partir do momento em que a simulação consegue obter a fdp do sistema da FRST, sendo possível, assim, obter uma avaliação mais precisa da sensibilidade do sistema para mudanças nos parâmetros que o compõem. Com este resultado, consegue-se prever o comportamento futuro deste sistema, utilizando as variações de sensibilidade, e quantificar as probabilidades dos novos estados possíveis.

Na FRST, hodiernamente, todas estas análises são realizadas manualmente e, como não há uma base de dados grande o suficiente para quantificar todas as variações em estados estacionários sem um MCMC, o algoritmo para análise se faz uma ferramenta imprescindível para a realização de mudanças em uma plataforma que necessita de um fluxo constante de novos conteúdos e features. A realização de uma análise de usuários mais consistente consegue resolver esta necessidade da FRST provendo cálculos mais precisos por meio de simulações.

2 Cadeias de Markov

A cadeia de Markov é um processo estocástico que evolui ao longo do tempo, passando por diversos estados. A sequência de estados é denotada pela coleção de valores, ou seja, é uma sequência de variáveis aleatórias dependentes. Em uma cadeia de Markov o único estado que importa é o anterior, dessa forma para determinar a sequência de valores que a cadeia pode assumir, podemos determinar a distribuição do próximo valor conhecendo apenas o valor anterior. A coleção de estados que uma cadeia de Markov pode visitar é chamada de espaço de estados. A distribuição de probabilidade condicional, que determina se a cadeia se move de um estado para o outro, é chamada de kernel de transição ou matriz de transição. A seguir tem-se uma representação gráfica de uma Cadeia de Markov:

3 Monte Carlo

A simulação de monte carlo é aplicada para problemas em que tentamos prever o comportamento de processos aleatórios. Diferente de um modelo de previsão normal, a Simulação de Monte Carlo prevê um conjunto de resultados com base em um intervalo de valores estimados em relação a um conjunto de valores de entrada fixos. Em outras palavras, uma simulação de Monte Carlo cria um modelo de resultados possíveis, usando uma distribuição de probabilidade, como uma distribuição uniforme ou normal, para qualquer variável que tenha incerteza inerente. Ela, então, irá recalculer os resultados sucessivamente, cada vez usando um conjunto diferente de números aleatórios entre os valores mínimo e máximo. Em um teste típico de Monte Carlo, este exercício pode ser repetido milhares de vezes para produzir um grande número de resultados prováveis. A seguir tem-se uma fórmula de uma simulação de Monte Carlo e Cadeia de Markov

$$\frac{P_m}{P_n} = \frac{\exp(\frac{-U_m}{K_b * T})}{\exp(\frac{-U_n}{K_b * T})} = \exp(-\frac{U_m - U_n}{K_b * T}) \quad (17)$$

A equação acima é usada em mecânica estatística ou termodinâmica para descrever a relação entre as probabilidades de um sistema estar em diferentes estados de energia. P_m e P_n são probabilidades de encontrar o sistema nos estados m e n U_m e U_n são as energias correspondentes a esses estados. k_B é a constante de Boltzmann. T é a temperatura absoluta. Essa equação é fundamental para entender como a probabilidade de um sistema estar em um certo estado de energia varia com a temperatura e as diferenças de energia. No contexto de MCMC, essa equação pode ser usada para calcular as probabilidades de transição entre estados em um sistema físico.

4 Aplicação do MCMC

Métodos de Monte Carlo via Cadeias de Markov (MCMC) consistem em construir uma cadeia de Markov com uma distribuição estacionária desejada e rodar essa cadeia por um longo período de tempo, até que ela se converta (aproximadamente) para essa distribuição estacionária. Os métodos de MCMC são usados basicamente para gerar amostras de uma distribuição que são correlacionadas, especialmente quando a distribuição em questão é complexa ou de alta dimensão.

O ponto principal da MCMC está na formulação de propriedades de transição apropriadas, ou seja, na definição de uma regra de transição que permita à cadeia de Markov explorar eficientemente o espaço de estados e, eventualmente, atingir a distribuição estacionária. Este raciocínio faz muito sentido dentro de uma lógica empresarial e, especialmente, da finalidade que a FRST procura, sendo esta uma busca por prever quais são os estados com maior probabilidade de abandono na plataforma, coisa que a obtenção dos fluxos associados a probabilidades estacionárias conseguem auxiliar a prever, e entender quais são os fluxos dos usuários dentro da plataforma.

5 Análise teórica

Para o escopo do projeto, faz-se necessário compreender o fluxo comportamental do usuário no site, ou seja, qual o caminho que um indivíduo percorre dentro da plataforma. Nessa medida, deve-se criar estados e atribuir um número a cada um deles, por exemplo uma página pode equivaler ao estado 3

e a ação de sair do site pode ser o estado 20, e então, pode-se montar uma matriz de transição, a qual descreve as probabilidades de transição em uma cadeia de Markov. Ainda, essa matriz é de suma importância para a análise, pois com ela é possível rodar as simulações de Monte Carlo e obter a função de distribuição de probabilidade para cada cenário. Por último, todas essas variáveis são latentes e não há necessidade de uma variável proxy (variável utilizada para substituir outra de difícil mensuração), pois todas as informações são fáceis de se obter, uma vez que esses dados já foram coletados.

6 Análise descritiva

Para a elaboração e sucesso deste projeto, é fundamental que os dados referentes ao tráfego de usuários em nossa plataforma sejam coletados, armazenados e organizados para a análise. Esses dados originam-se diretamente das interações que os usuários estabelecem com a plataforma - especificamente, através dos cliques efetuados e do tempo despendido em cada página. Tais interações são interpretadas como ações explícitas dos usuários, sendo registradas em detalhe a cada sessão de login. Este processo de coleta permite um mapeamento preciso e individualizado do comportamento de cada usuário, estabelecendo a base para nossa análise.

Cada ação do usuário é catalogada e vinculada ao seu respectivo login, garantindo a fonte de dados para análise. Essas informações são essenciais para definir os "estados" no contexto das Cadeias de Markov, onde cada página visitada na plataforma é considerada um estado possível. A transição de um estado para outro - ou de uma página para outra - é influenciada pela probabilidade baseada no comportamento histórico dos usuários, informações essas que são cruciais para a modelagem através das Cadeias de Markov.

7 Coleta e Estruturação dos Dados

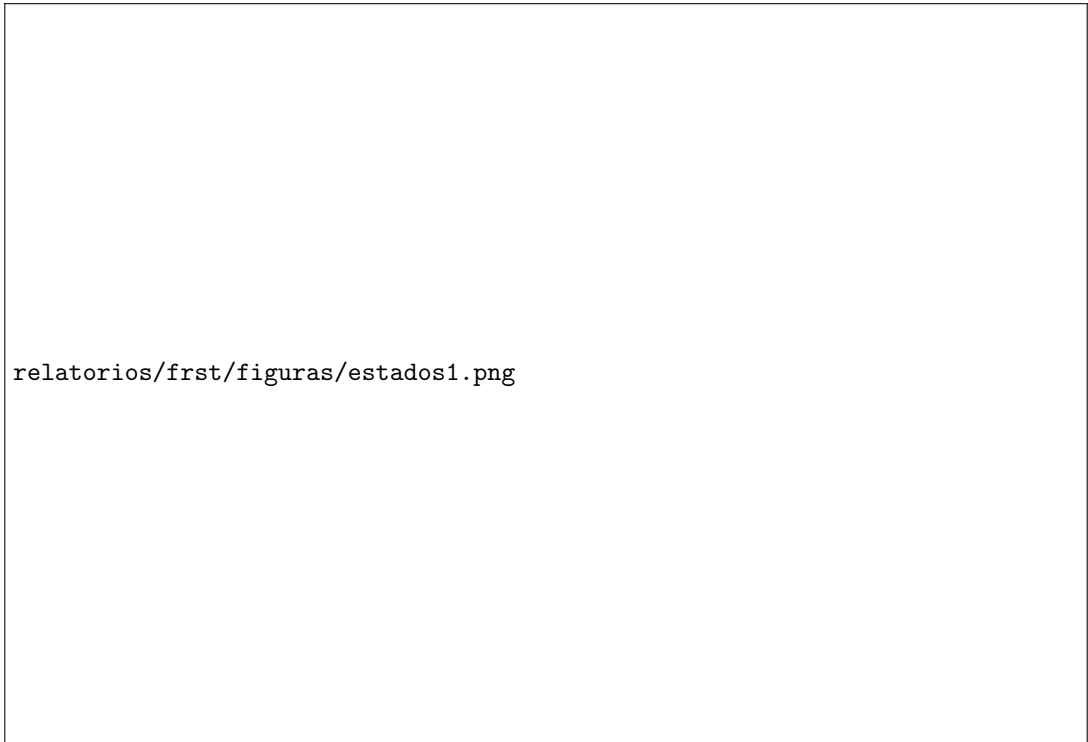
A coleta de dados é realizada por meio de ferramentas de tracking embutidas na plataforma, que registram cada clique e o tempo gasto em cada página por usuário. Esses dados são então estruturados em um formato que facilita a aplicação das técnicas de análise mencionadas. A estruturação adequada inclui a categorização das páginas (estados), o mapeamento das transições entre páginas e a quantificação do tempo despendido. Além disso, aspectos como a ordem das visitas às páginas e as ações realizadas em cada uma são detalhadamente registrados, permitindo uma análise granular do fluxo de navegação do usuário.

8 Visualização dos estados da FRST

Após a coleta de dados é possível visualizar os estados de transição (páginas do site) que compõem a cadeia de Markov através de um gráfico de rede, como pode ser visto a seguir:


Este gráfico representa todos os estados do site da FRST com todas as possibilidades de transição entre as páginas e a frequência que os usuários estão em cada página do site, mas como a visualização fica impossibilitada pela quantidade dos estados na cadeia.

Cabe pontuar que, esse gráfico é uma fração da base de dados inteira, pois utilizar uma parte menor ajuda na clareza visual, ainda mais que, os estados apresentados nesta visualização simplificada compõem os principais da plataforma da FRST.



relatorios/frst/figuras/estados1.png

Figura 2: Estados



relatorios/frst/figuras/estados2.png

Figura 3: Estados

9 Conclusão

A partir da implementação do método de Monte Carlo, a análise permitirá realizar múltiplas simulações de caminhos de navegação dos usuários pela plataforma, baseando-se nas probabilidades derivadas das Cadeias de Markov. Ao simular inúmeros cenários possíveis, podemos identificar padrões e tendências no comportamento do usuário, fornecendo insights valiosos para otimizações na usabilidade e no design da interface da plataforma FRST.