

# Um Dashboard para Análise de Cancelamentos em Aplicativos de Mobilidade Urbana na Cidade do Rio de Janeiro

Raquel Torres de Paiva<sup>1</sup>, Ana Cristina Bicharra Garcia<sup>1</sup>, Fernanda Bertini Viégas<sup>2</sup>

<sup>1</sup> Programa de Pós-Graduação em Informática (PPGI)  
Universidade Federal do Estado do Rio de Janeiro (UNIRIO)

Rio de Janeiro – RJ – Brasil  
raquel.paiva@edu.unirio.br  
cristina.bicharra@uniriotec.br

<sup>2</sup> Harvard University  
Cambridge – MA – EUA  
fviegas@harvard.edu

**Resumo.** *Cancelamentos em aplicativos de mobilidade urbana permanecem frequentes e, para passageiros, muitas vezes são percebidos como um problema sem explicação clara, reduzindo a confiabilidade do serviço. Esta pesquisa investiga esse cenário na cidade do Rio de Janeiro a partir de uma abordagem orientada a data fusion e visualização, propondo um dashboard que integra dados de corridas (previamente tratados e anonimizados) com múltiplas fontes públicas e georreferenciadas. O artefato organiza e correlaciona informações sobre o território, dinâmicas urbanas e manifestações de insatisfação em plataformas de reclamação, permitindo explorar padrões de cancelamento e levantar hipóteses sobre possíveis motivações, incluindo indícios de redlining, efeitos de eventos urbanos e aspectos associados à percepção de segurança. O método adota Design Science Research e estrutura um pipeline técnico que cobre aquisição, pré-processamento, modelagem geoespacial/temporal, construção de indicadores e implementação full stack (back-end e front-end). As análises do dashboard incluem (i) um módulo de Comentários que avalia comentários dentro da plataforma e em uma plataforma externa de Reclamações, (ii) um módulo de redlining que contrapõe taxas de cancelamento com recortes socioespaciais e (iii) um módulo de tempo de resposta institucional a ocorrências urbanas, explorando cancelamentos como possível indicador indireto de diferenças regionais na duração e resolução de eventos. A avaliação empírica com participantes indicou percepção predominantemente positiva quanto à clareza das visualizações, navegação e utilidade do artefato para apoiar a identificação de padrões e a formulação de hipóteses investigáveis.*

**Abstract.** *Ride cancellations in urban mobility apps remain frequent and, from the passengers' perspective, are often experienced as a poorly explained issue, undermining service reliability. This study investigates this scenario in Rio de Janeiro through a data fusion and visualization-oriented approach, proposing a dashboard that integrates ride data (previously cleaned and anonymized) with multiple public, georeferenced sources. The artifact organizes and correlates*

*information about the territory, urban dynamics, and expressions of dissatisfaction found in complaint platforms, enabling the exploration of cancellation patterns and the formulation of hypotheses about possible drivers, including signals consistent with redlining, the effects of urban events, and aspects related to perceived safety. The research follows Design Science Research and defines a technical pipeline covering data acquisition, preprocessing, spatiotemporal modeling, indicator construction, and a full stack implementation (back-end and front-end). Dashboard analyses include (i) a Comments module that examines in-app reviews and complaints posted on an external consumer-complaint platform, (ii) a redlining module that contrasts cancellation rates across socio-spatial strata, and (iii) an institutional response-time module for urban incidents, using cancellations as a potential indirect indicator of regional differences in event duration and resolution. An empirical evaluation with participants reported predominantly positive perceptions regarding visualization clarity, navigation, and the artifact's usefulness to support pattern identification and the formulation of testable hypotheses.*

## **1. Introdução**

Aplicativos de mobilidade urbana transformaram a forma de deslocamento em grandes cidades, mas ainda convivem com altos índices de cancelamento e impactos diretos na experiência do passageiro, incluindo frustração, imprevisibilidade e perda de confiança no serviço. O problema central desta dissertação é a falta de clareza sobre as motivações dos cancelamentos, em um contexto onde fatores operacionais, territoriais e sociotécnicos se misturam, e onde também surgem discussões sobre vieses e discriminação em serviços *peer-to-peer*. Trabalhos prévios sugerem que decisões de aceitação e cancelamento podem refletir vieses individuais e condições do ambiente urbano, e que o tema tem relevância social e espaço para intervenções por meio de Sistemas de Informação [Brown 2023], [Monachou and Ashlagi 2019], [Pandey and Caliskan 2021], [Rosenblat et al. 2017].

No caso do Rio de Janeiro, eventos urbanos (por exemplo, acidentes, alagamentos, ocorrências de segurança) e características territoriais podem afetar disponibilidade, rotas e risco percebido, influenciando o comportamento no serviço. Isso motiva a adoção de uma abordagem baseada em integração de dados e visualização, capaz de apoiar análises explicativas e exploração de padrões em diferentes escalas espaciais e temporais.

### **1.1. Questões de pesquisa e objetivos**

A pesquisa busca responder às seguintes questões, alinhadas ao artefato proposto:

- **RQ1:** O que os dados de aplicativos de corrida podem revelar a respeito do impacto de eventos urbanos?
- **RQ2:** Como os índices de cancelamento podem funcionar como um indicador do tempo de resposta aos eventos que ocorrem na cidade?
- **RQ3:** Os dados de aplicativos de mobilidade, quando correlacionados com informações públicas geoprocessadas, podem revelar indícios de práticas de *redlining* (recusa ou restrição sistemática de serviços com base em atributos territoriais/socioeconômicos) ou de segregação socioeconômica?

O objetivo geral é propor e avaliar um dashboard que consolide e visualize dados heterogêneos, permitindo explorar padrões de cancelamento e apoiar a formulação de hipóteses sobre possíveis motivações. Como objetivos específicos, destacam-se: (i) integrar e modelar dados de fontes distintas sob uma perspectiva geoespacial e temporal; (ii) definir indicadores e filtros para investigação do fenômeno; (iii) implementar e disponibilizar o artefato; (iv) realizar avaliação empírica de utilidade, clareza e navegação.

## 2. Trabalhos relacionados

A literatura aponta que cancelamentos podem refletir tanto fatores operacionais quanto aspectos sociotécnicos associados a vieses e discriminação. Estudos discutem discriminação direta, estatística/proxy e vieses *taste-based* [Brown 2023], [Monachou and Ashlagi 2019], além de evidências de efeitos discriminatórios em avaliações e decisões em plataformas [Rosenblat et al. 2017]. Outras abordagens analisam vieses em algoritmos e precificação dinâmica, indicando relações entre localização, perfil demográfico e condições de oferta [Pandey and Caliskan 2021]. Em paralelo, há interesse crescente em métodos e sistemas que apoiem transparência, explicação e auditoria de decisões algorítmicas, incluindo propostas orientadas a explicabilidade e monitoramento [Hevner and Chatterjee 2010].

Para organizar essas contribuições e evidenciar o espaço em que esta dissertação se insere, a Tabela 1 sintetiza os principais trabalhos empíricos e teóricos que relacionam cancelamentos e viés em plataformas de mobilidade e outros serviços digitais similares. A coluna de limitações resume, em cada caso, restrições de dados, escopo ou método que motivaram as decisões de foco e abordagem desta pesquisa.

O diferencial desta dissertação é concentrar-se em um artefato de visualização e integração de dados para o contexto do Rio de Janeiro, tratando cancelamentos como um fenômeno observável no território e no tempo. A proposta organiza múltiplas fontes e oferece um mecanismo prático para exploração, comunicação de evidências e geração de conjecturas investigáveis.

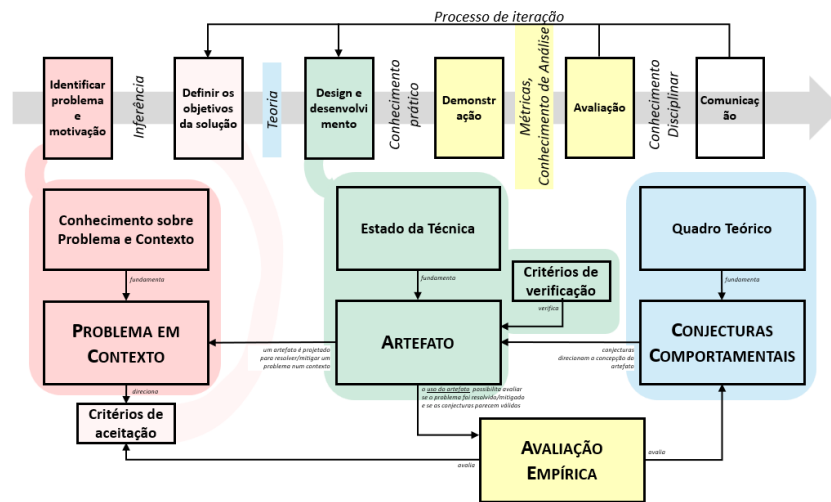
## 3. Metodologia de pesquisa

A pesquisa segue *Design Science Research* (DSR), articulando construção do artefato, critérios de verificação e avaliação empírica [Dresch et al. 2015], [Hevner and Chatterjee 2010]. O trabalho estrutura:

- **Conhecimento do problema e contexto:** entendimento do domínio e das limitações de explicação do cancelamento para passageiros.
- **Projeto e desenvolvimento do artefato:** definição de pipeline, modelagem e implementação do dashboard.
- **Avaliação:** estudo empírico com tarefas guiadas e questionário Likert para utilidade, clareza e navegação.

A Figura 1 representa, de forma simplificada, essa relação entre o processo iterativo da DSR e os elementos centrais do Modelo-DSR. Na parte superior, destacam-se as etapas clássicas do ciclo de *design* (identificar o problema e a motivação, definir os objetivos da solução, projetar e desenvolver, demonstrar, avaliar e comunicar). Na parte inferior, aparece o arranjo conceitual do Modelo-DSR: o conhecimento sobre problema e contexto fundamenta a formulação do problema em contexto; o estado da técnica e o

quadro teórico subsidiam o *design* do artefato e a formulação de conjecturas comportamentais; critérios de verificação e de aceitação são definidos para orientar a construção e a avaliação; e a avaliação empírica realimenta o conhecimento sobre o fenômeno e o próprio quadro teórico [Pimentel et al. 2019], [Pimentel et al. 2020].



**Figura 1. Metodologia e Modelo-DSR simplificado adotados nesta dissertação. (Figura 2.1 da Dissertação)**

Fonte: Adaptado de [Pimentel et al. 2019], [Pimentel et al. 2020].

## 4. Base de dados, fusão e modelagem

O estudo consolida dados de corridas previamente tratados e anonimizados, integrando fontes públicas e geoprocessadas (por exemplo, recortes territoriais oficiais, dados socioeconômicos, ocorrências urbanas e eventos de violência, além de reclamações de usuários). A combinação é orientada por (i) Modelagem geoespacial, (ii) Modelagem temporal e (iii) Tratamentos e qualidade.

### 4.1. Pipeline técnico e implementação

O pipeline técnico organiza aquisição, pré-processamento, modelagem geoespacial/temporal, exploração e desenho de indicadores, implementação de *back-end* (Flask + PyMongo), *front-end* (HTML/CSS/JS com mapas e gráficos) e publicação com foco em desempenho e estabilidade. A operacionalização *full stack* busca viabilizar consultas agregadas e interações com filtros, fornecendo uma experiência exploratória sem exigir ferramentas especializadas do usuário final.

## 5. Dashboard: páginas, técnicas e esquema de dados

A integração dos dados e as visualizações deste estudo foram estruturadas à luz de propostas consolidadas da literatura de visualização que relacionam tipos de variáveis, esquemas de dados e famílias de técnicas. Em particular, adotou-se a definição de *slots* de dados apresentada por [Viégas et al. 2007], que descreve os papéis desempenhados por colunas textuais, numéricas e hierárquicas na construção de gráficos e mapas. Essa formalização permitiu explicitar, para cada página do *dashboard*, quais combinações de variáveis (textuais, numéricas, temporais e regionais) compõem cada visualização.

Neste trabalho, de maneira análoga, adotamos cinco tipos de *slots*, que definem as demandas mínimas de dados por técnica [Viégas et al. 2007]:

- **U** (*Desestruturado*, do inglês *Unstructured*): texto livre (comentários, títulos, descrições);
- **T** (*Textual*): rótulos/identificadores categóricos (por exemplo, bairro, dia da semana, gênero);
- **N** (*Numérico*): medidas numéricas (contagem, taxa, distância, duração);
- **T+** (*Textual múltiplo*): hierarquias ou múltiplas categorias;
- **N+** (*Numérico múltiplo*): múltiplas medidas (por exemplo, séries empilhadas).

O dashboard foi concebido como um artefato modular, com páginas dedicadas a diferentes perguntas analíticas e técnicas de visualização. A Tabela 2 sintetiza páginas, técnicas e coleções/esquemas de dados associados.

**Tabela 2. Páginas do dashboard, técnicas de visualização e esquemas de dados (síntese).**

Página	Técnicas/visuais principais	Coleções/esquemas de dados
Info geral	Séries temporais e distribuição; indicadores agregados; filtros	<code>rides_original</code> , <code>ride_features</code>
Favelas e <i>redlining</i>	Mapas coropléticos e camadas; comparação por recorte territorial	<code>favelas_rij</code> , <code>rides_original</code> , <code>ride_features</code> , <code>censo_renda</code>
Impacto da violência	Mapas e indicadores em janelas temporais; correlações exploratórias	<code>events (violência)</code> , <code>rides_original</code> , <code>ride_features</code>
Mapa de ocorrências	Mapas por tipo de ocorrência; filtros por período e local	<code>ocorrencias</code> , <code>ride_features</code>
Análise de comentários	Distribuições por categoria; evidências textuais sumarizadas	<code>app_comments</code>
Análise Reclame Aqui	Distribuições por classe e tema; recortes por tipo de reclamante	<code>reclame_aqui_complaints</code>

A tabela reforça o núcleo do trabalho: fusão de fontes e visualizações orientadas a perguntas, conectando dados de corridas, território, eventos urbanos e relatos textuais em um único ambiente exploratório.

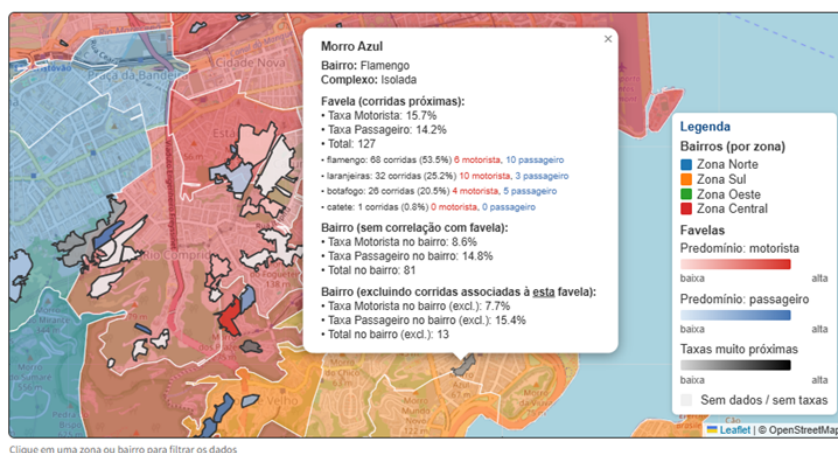
## 6. Análises e resultados

As análises do dashboard foram organizadas para sustentar as questões de pesquisa, com ênfase em recortes territoriais e temporais e em uma leitura sociotécnica do fenômeno.

### 6.1. Indícios de *redlining* e recortes socioespaciais (RQ1/RQ3)

O módulo de *redlining* contrasta taxas de cancelamento com recortes territoriais e indicadores socioeconômicos, permitindo observar padrões espaciais persistentes e diferenças

### Mapa de Redlining



Clique em uma zona ou bairro para filtrar os dados

#### Fontes de dados

• Fonte dos bairros: IBGE — "Censo 2022: População e domicílios por bairros (dados preliminares)".

• Fonte identificação das Favelas: DATA.RIO — "Limite Favelas 2019" — Prefeitura da Cidade do Rio de Janeiro.

**Figura 2. Visão final de redlining: tooltips com percentuais por favela e bairros adjacentes (Figura 6.14 da dissertação)**

por zona/bairro, conforme Figura 2 . A abordagem não afirma causalidade, mas oferece evidências exploratórias para discussão e formulação de hipóteses. Para fortalecer a interpretação, o dashboard também incorpora sinais qualitativos a partir de reclamações e comentários, alinhando achados quantitativos a percepções relatadas em plataformas externas, conforme módulos de avaliação de comentários na Figura 3. A literatura sobre discriminação em serviços e *ridesharing* orienta a leitura dessas evidências como indícios dependentes de contexto [Brown 2023], [Monachou and Ashlagi 2019], [Pandey and Caliskan 2021], [Rosenblat et al. 2017].

## 6.2. Tempo de resposta a ocorrências urbanas (RQ2)

Além de identificar áreas críticas, o dashboard inclui uma análise sobre diferenças regionais no tempo de resposta e na duração de eventos, utilizando cancelamentos como indicador indireto para investigar possíveis efeitos sobre o serviço. A lógica é observar se regiões com maior duração/complexidade de ocorrências tendem a apresentar padrões distintos de cancelamento em janelas temporais relacionadas aos eventos. Essa análise é útil tanto para leitura de impacto urbano quanto para discutir a confiabilidade percebida do serviço em situações de instabilidade, conforme observado na Figura 4.

## 6.3. Avaliação empírica do artefato

A avaliação empírica foi conduzida com participação assíncrona e anônima, combinando roteiro de tarefas e questionário Likert (1 a 5) cobrindo clareza das visualizações, navegação e utilidade para identificação de padrões. Os resultados agregados indicam médias e medianas elevadas, com proporção expressiva de respostas em 4 e 5, sugerindo percepção predominantemente positiva do dashboard como instrumento de exploração e apoio à análise. Também emergiram sugestões de melhoria, como reforço de sumários estatísticos e destaques para apoiar leitura comparativa em algumas visualizações. A Tabela 3 resume as avaliações do questionário de testes.



**Figura 3. Exemplo com indícios de discriminação e redlining (Figura 6.10 da Dissertação)**



**Figura 4. Comportamento dos cancelamentos em relação aos eventos urbanos por zona (Figura 6.17 da Dissertação)**

**Tabela 3. Avaliação do piloto – resumo por item (escala Likert) (Tabela 6.1 da Dissertação)**

N	Média	Mediana	%4-5
14.0	4.14	5.0	78.57
14.0	4.43	5.0	85.71
14.0	3.86	4.0	78.57
14.0	4.07	4.5	85.71
14.0	4.21	5.0	78.57

Fonte: Elaboração própria, 2025.

## 7. Discussão e contribuições

A contribuição principal é um artefato de Sistema de Informação que materializa um pipeline de integração e visualização de dados para investigar cancelamentos no território e no tempo, com potencial de apoiar tanto pesquisa quanto tomada de decisão. Em termos científicos e tecnológicos, o trabalho contribui com:

- **Integração de fontes heterogêneas** (corridas, território, eventos e relatos textuais) sob modelagem geoespacial/temporal coerente, viabilizando análises antes fragmentadas.
- **Mecanismo exploratório para indícios de *redlining*** e para leitura de impactos urbanos, sem extrapolar causalidade, mas ampliando transparência e capacidade de auditoria informal.
- **Avaliação empírica** que evidencia utilidade percebida do artefato para identificar padrões e formular hipóteses investigáveis.

Como limitações, destacam-se a natureza observacional dos dados e a indisponibilidade de variáveis internas da plataforma (por exemplo, sinais completos de decisão do motorista, detalhes finos sobre o momento do cancelamento e informações sensíveis), o que restringe inferências causais e reforça a necessidade de triangulação com novos dados e estudos complementares.

## 8. Conclusão

Esta dissertação endereça a falta de clareza sobre motivações de cancelamentos em aplicativos de mobilidade urbana por meio de um dashboard que integra e visualiza dados de corridas, território, eventos urbanos e relatos de usuários. O artefato permite observar padrões espaço-temporais, levantar indícios de *redlining* e explorar cancelamentos como indicador indireto para diferenças regionais no tempo de resposta a ocorrências, contribuindo para a compreensão sociotécnica do fenômeno e para a discussão de confiabilidade do serviço. A avaliação empírica sugere que o dashboard é percebido como claro e útil para exploração e geração de insights. Como trabalhos futuros, destacam-se ampliação de bases, refinamento de métricas temporais e integração de evidências adicionais que permitam análises mais explicativas e comparáveis entre cidades e plataformas.

## Referências

- Abramova, O. (2022). No matter what the name, we're all the same? examining ethnic online discrimination in ridesharing marketplaces. *Electronic Markets*, 32(3):1419–1446.

- Brown, A. E. (2023). Prevalence and mechanisms of discrimination: Evidence from the ride-hail and taxi industries. *Journal of Planning Education and Research*, 43(2):268–280.
- Dresch, A., Lacerda, D. P., and Antunes Júnior, J. A. V. (2015). *Design Science Research: método de pesquisa para avanço da ciência e tecnologia*. Bookman, Porto Alegre.
- Ge, Y., Knittel, C. R., MacKenzie, D., and Zoepf, S. (2020). Racial discrimination in transportation network companies. *Journal of Public Economics*, 190:104205.
- Hevner, A. and Chatterjee, S. (2010). *Design Research in Information Systems*, volume 22 of *Integrated Series in Information Systems*. Springer, New York.
- Mejia, J. and Parker, C. (2021). When transparency fails: Bias and financial incentives in ridesharing platforms. *Management Science*, 67(1):166–184.
- Monachou, F. G. and Ashlagi, I. (2019). Discrimination in online markets: Effects of social bias on learning from reviews and policy design. In *Advances in Neural Information Processing Systems*, volume 32.
- Moody, J., Middleton, S., and Zhao, J. (2019). Rider-to-rider discriminatory attitudes and ridesharing behavior. *Transportation Research Part F: Traffic Psychology and Behaviour*, 62:258–273.
- Pandey, A. and Caliskan, A. (2021). Disparate impact of artificial intelligence bias in ridehailing economy’s price discrimination algorithms. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society (AIES ’21)*, pages 822–833, New York, NY, USA. Association for Computing Machinery.
- Pimentel, M., Filippo, D., and Marcondes dos Santos, T. (2020). Design science research: pesquisa científica atrelada ao design de artefatos. *RE@D - Revista de Educação a Distância e eLearning*, 3(1).
- Pimentel, M., Filippo, D., and Santoro, F. M. (2019). Design science research: fazendo pesquisas científicas rigorosas atreladas ao desenvolvimento de artefatos computacionais projetados para a educação. *Metodologia de Pesquisa em Informática na Educação: Concepção da Pesquisa*, pages 5–29.
- Rosenblat, A., Levy, K. E. C., Barocas, S., and Hwang, T. (2017). Discriminating tastes: Uber’s customer ratings as vehicles for workplace discrimination. *Policy & Internet*, 9(3):256–279.
- Salvat, C. R. (2021). Discrimination in artificial intelligence. Trabalho de conclusão de curso, Universitat Pompeu Fabra, Barcelona.
- Tushev, M., Ebrahimi, F., and Mahmoud, A. (2022). A systematic literature review of anti-discrimination design strategies in the digital sharing economy. *IEEE Transactions on Software Engineering*, 48(12):5148–5157.
- Viégas, F. B., Wattenberg, M., van Ham, F., Kriss, J., and McKeon, M. (2007). ManyEyes: a site for visualization at internet scale. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1121–1128.

**Tabela 1. Síntese de trabalhos relacionados a cancelamentos associados a vieses (Tabela 3.1 da Dissertação)**

<b>Referência</b>	<b>Trabalho abordado</b>	<b>Limitações encontradas</b>
[Ge et al. 2020]	Auditoria randomizada em TNCs; efeito de nomes em aceite/cancelamento e tempos de espera.	Foco em duas cidades/duas plataformas; visibilidade de atributos difere por plataforma.
[Brown 2023]	Comparação ride-hail versus táxis (LA); redução de hiatos de serviço no <i>ride-hail</i> .	Resultados de uma região; mecanismos subjacentes pouco identificados.
[Mejia and Parker 2021]	Experimento de campo sobre transparência versus viés pós-aceite.	Efeito heterogêneo por grupo; plataforma/mercado únicos.
[Rosenblat et al. 2017]	Avaliações de clientes como veículos de discriminação; implicações legais.	Dados proprietários inacessíveis; inferências indiretas.
[Moody et al. 2019]	Atitudes discriminatórias entre passageiros e uso de corridas compartilhadas.	Medidas auto-declaradas; relação causal com cancelamentos não direta.
[Abramova 2022]	Experimento de <i>Discrete choice</i> evidenciando que sobrenomes e sinais de qualidade podem influenciar os cancelamentos; desvantagens para certos grupos.	Contexto específico para o mercado analisado.
[Pandey and Caliskan 2021]	Medição em larga escala de <i>disparate impact</i> em tarifas usando dados de Chicago.	Identificação não experimental; agregação espacial.
[Salvat 2021]	Discussão sobre testes obrigatórios de viés, <i>trade-offs</i> de transparência e governança.	Caráter conceitual; requer validação em contexto local.
[Tushev et al. 2022]	Síntese de vieses em plataformas e técnicas de <i>fairness</i> (SLR).	Escopo amplo; pouca evidência aplicada a cancelamentos.
[Monachou and Ashlagi 2019]	Modelo de mercado com vieses de crença; política de <i>Directed Matching</i> reduz <i>gap</i> .	Base teórica; não testado em mobilidade.

Fonte: Elaboração própria, 2025.