

Análise da Percepção Pública e Eficácia dos Investimentos na Saúde de Belo Horizonte: Uma Abordagem Baseada em Dados

Cleidimar Lacerda dos Passos¹

¹Universidade Federal de Viçosa
cleidimar.passos@ufv.br

Resumo

Resumo. Este trabalho apresenta uma análise da percepção pública sobre as Unidades de Pronto Atendimento (UPAs) de Belo Horizonte, utilizando técnicas de Mineração de Dados e Processamento de Linguagem Natural (PLN). Foram coletadas e analisadas mais de 6.000 avaliações da plataforma Google Maps entre 2020 e 2025. O estudo investiga a correlação entre os sentimentos expressos pelos cidadãos e os dados oficiais de despesas orçamentárias do município. Os resultados refutam a hipótese de que o aumento do investimento financeiro resulta em melhora imediata na percepção de qualidade, evidenciando uma correlação negativa em unidades críticas e sugerindo um cenário de gastos reativos frente ao colapso do sistema.

Resumo

Abstract. This work presents an analysis of public perception regarding Emergency Care Units (UPAs) in Belo Horizonte, using Data Mining and Natural Language Processing (NLP) techniques. Over 6,000 reviews were collected from the Google Maps platform between 2020 and 2025. The study investigates the correlation between sentiments expressed by citizens and official municipal budget expenditure data. The results refute the hypothesis that increased financial investment results in immediate improvement in quality perception, evidencing a negative correlation in critical units and suggesting a scenario of reactive spending in the face of system collapse.

1. Introdução

A gestão da saúde pública em grandes metrópoles enfrenta um desafio persistente: o hiato temporal entre a ocorrência de problemas estruturais e a tomada de decisão pelos gestores. Tradicionalmente, o feedback do cidadão é coletado através de canais burocráticos, como ouvidorias e censos anuais, que possuem baixa amostragem e alto tempo de processamento.

Com a popularização das plataformas digitais, o cidadão passou a atuar como um “sensor humano”, relatando experiências em tempo real nas redes sociais e plataformas de geolocalização. O objetivo deste trabalho é coletar, processar e analisar dados de avaliações públicas das 10 principais UPAs de Belo Horizonte, cruzando essas informações qualitativas com dados quantitativos de investimento financeiro.

2. Trabalhos Relacionados

A utilização de mídias sociais para monitoramento de saúde pública é uma área crescente na Ciência de Dados. Estudos anteriores demonstraram a eficácia do Twitter para detecção precoce de surtos de gripe e Dengue Dredze (2012). No contexto de serviços públicos, Silva et al. (2019) utilizaram dados do Facebook para avaliar a satisfação com o transporte público.

3. Conjunto de Dados e Metodologia

A metodologia adotada divide-se em quatro etapas: Coleta, Processamento, Enriquecimento (NLP) e Análise Estatística.

3.1. Coleta de Dados (Web Scraping)

Para a obtenção das opiniões, foi desenvolvido um coletor automático (*scraper*) utilizando a biblioteca **Selenium** em Python. O alvo foram as páginas de perfil no Google Maps das 9 UPAs regionais de Belo Horizonte e do Hospital Metropolitano Odilon Behrens.

- **Volume:** Foram extraídas 6.146 avaliações.
- **Atributos:** Usuário (anonimizado), Data (relativa convertida para absoluta), Nota (1 a 5 estrelas) e Texto do comentário.

3.2. Limitação dos Dados

A amostra de dados apresenta limitações inerentes ao método de coleta e à natureza da plataforma, as quais devem ser consideradas na interpretação dos resultados. Primeiramente, os dados coletados sofrem de **viés de seleção** (*self-selection bias*), pois os usuários que se dão ao trabalho de postar reviews tendem a estar em extremos de satisfação ou, mais comumente, de insatisfação. Isso super-representa a Nota 1, indicando um cenário mais crítico do que a média da população geral. Em segundo lugar, a conversão de datas relativas (e.g., “há 3 meses”) para datas absolutas introduz uma **imprecisão temporal** marginal. Por fim, o modelo é limitado à percepção em plataformas de geolocalização, não abrangendo usuários sem acesso a smartphones ou internet.) e Texto muito longo no comentário da avaliação.

3.3. Disponibilidade de Dados e Código

Visando a transparência e reprodutibilidade científica, todo o pipeline de dados desenvolvido neste trabalho, incluindo os scripts de coleta, limpeza, análise estatística e geração de gráficos, encontra-se disponível em repositório público no GitHub:

<https://github.com/cleidimar-passos/tpfinal-analise-mineracao-midias.git>

3.4. Dados Financeiros Oficiais

Para a análise de correlação, foram extraídos dados do **Portal de Dados Abertos da PBH**, especificamente os datasets de “Despesas Orçamentárias” (2020 a 2025).

3.4.1 Limitações do Isolamento Financeiro

Embora a filtragem por “Fundo Municipal de Saúde” e “Hospital Odilon Behrens” permita distinguir o orçamento geral do autárquico, é importante notar que a base de dados oficial **não permite isolar a execução orçamentária para UPAs individuais**. Com exceção do HOB, que possui orçamento próprio, as despesas do Fundo Municipal são agregadas e não detalham o repasse final para cada uma das outras nove unidades. Portanto, a correlação para o Cenário 1 (Geral) deve ser interpretada como a relação entre o investimento macro em saúde e a média de percepção da rede, e não como um impacto financeiro direcionado a cada UPA.

4. Resultados: Caracterização e Análise

4.1. Caracterização dos Dados

A análise exploratória inicial revela um cenário de insatisfação crônica. A nota média histórica do sistema é de **2.49** (em uma escala de 5). A distribuição das notas apresenta forte assimetria negativa, onde a maioria absoluta das avaliações é de 1 estrela.

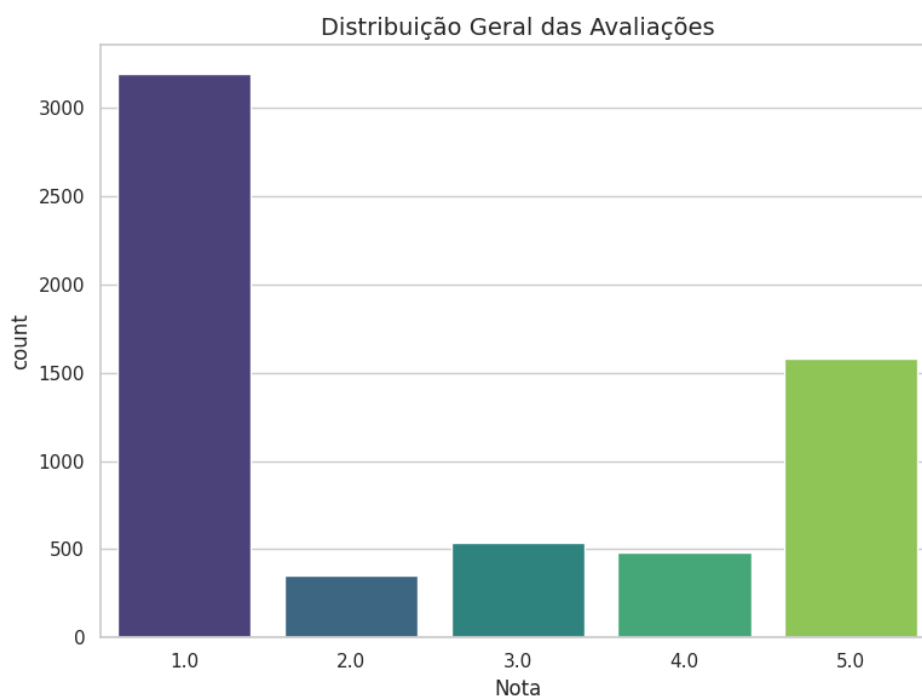


Figura 1: Distribuição de frequência das notas.

A Figura 2 demonstra o crescimento exponencial na participação dos usuários. Nos últimos cinco anos (2020–2024), o volume de avaliações cresceu mais de 300%, atingindo o pico de 1.222 reviews em 2024. Este aumento não apenas eleva a relevância estatística do dataset, mas também indica que o cidadão está recorrendo cada vez mais a plataformas digitais para expressar descontentamento, solidificando a utilidade do Google Maps como um sensor de crise.

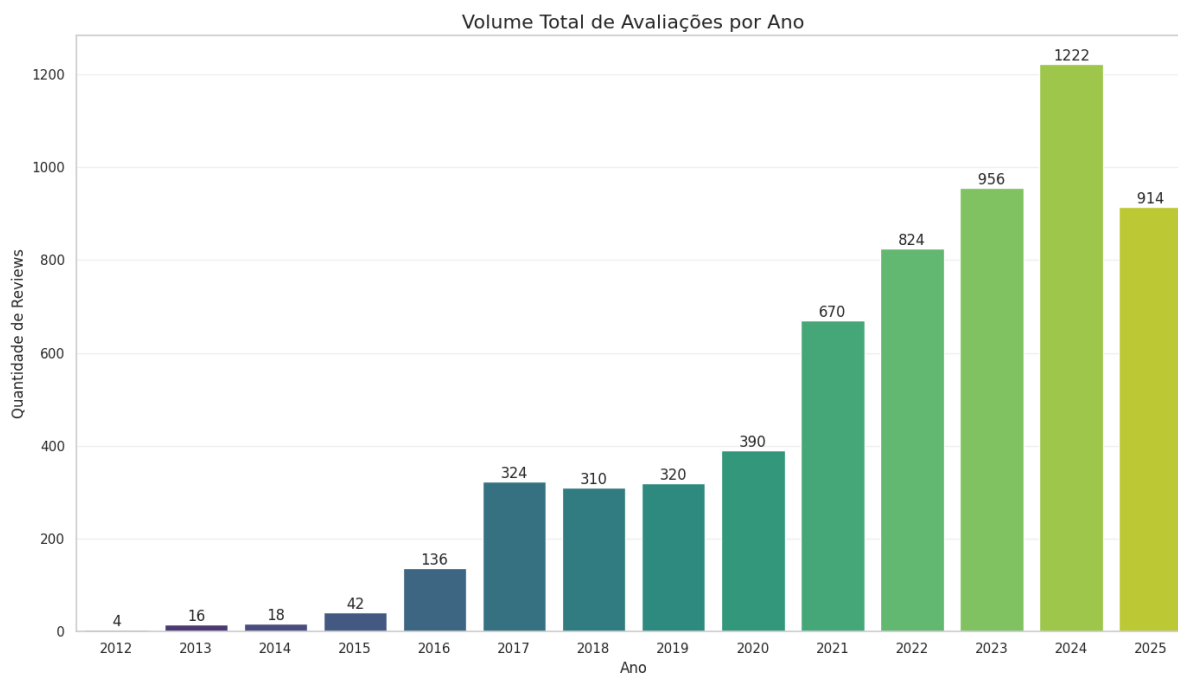


Figura 2: Volume Total de Avaliações por Ano (2012–2025).

Geograficamente, notou-se uma disparidade na qualidade percebida. A UPA Nordeste lidera o ranking com média 3.12, enquanto a UPA Venda Nova apresenta a situação mais crítica, com média 2.10.

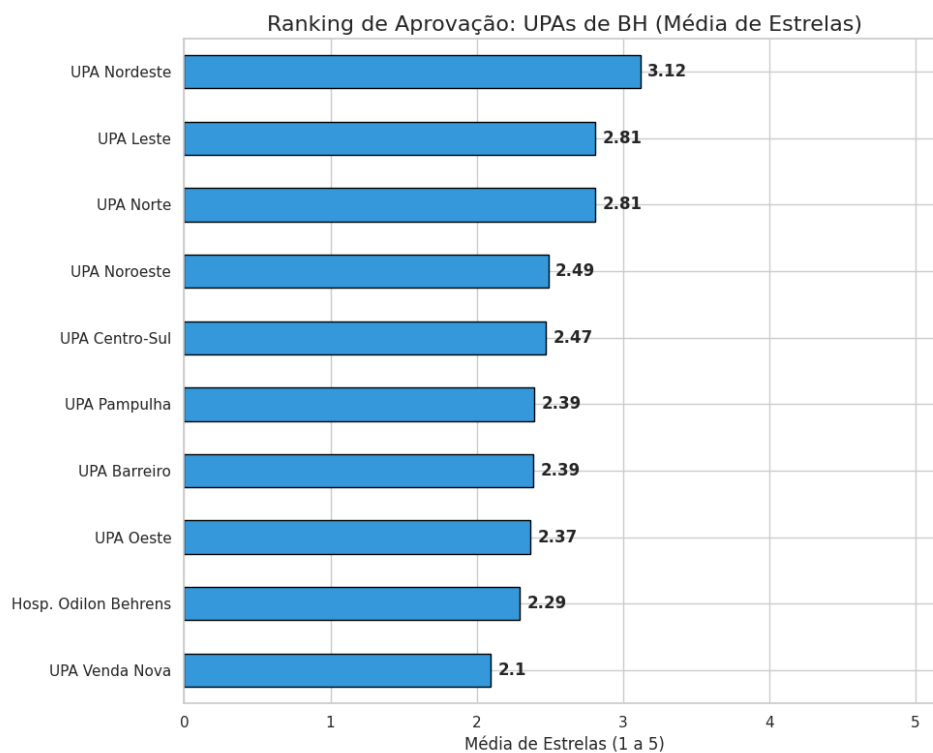


Figura 3: Ranking de média de aprovação por unidade.

4.2. Análise Temporal

A análise de séries temporais demonstra uma degradação progressiva do serviço a partir de 2023, coincidindo com o aumento exponencial do volume de reclamações em 2024.

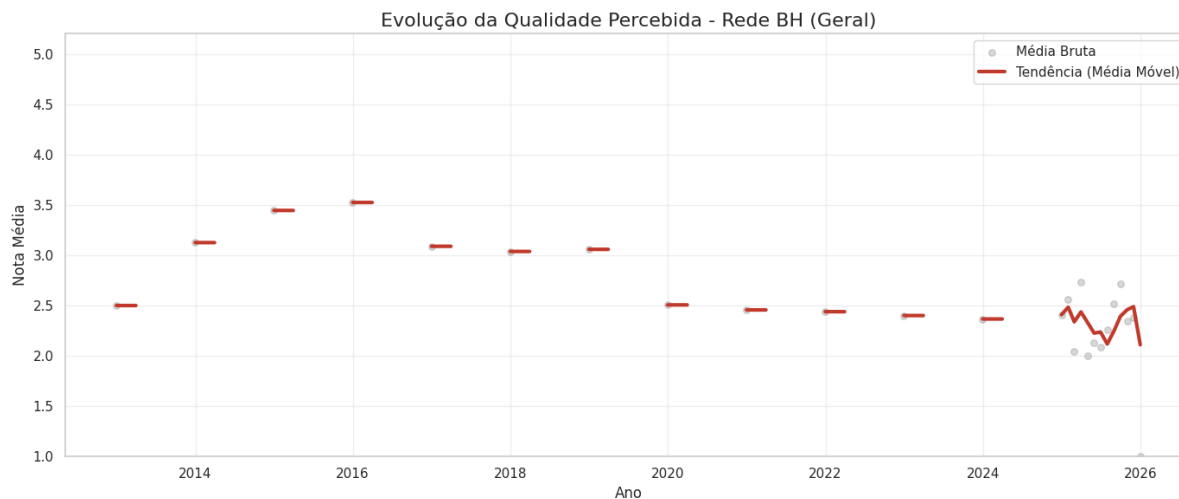
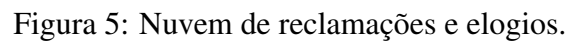


Figura 4: Evolução temporal da nota média (linha vermelha representa a média móvel).

Para compreender as causas da insatisfação, aplicou-se técnicas de PLN, gerando a nuvem de palavras.



A Figura 6 apresenta os **Tri-gramas** mais frequentes nas reclamações, evidenciando que a queixa principal é a demora no atendimento médico e a sensação de descaso.

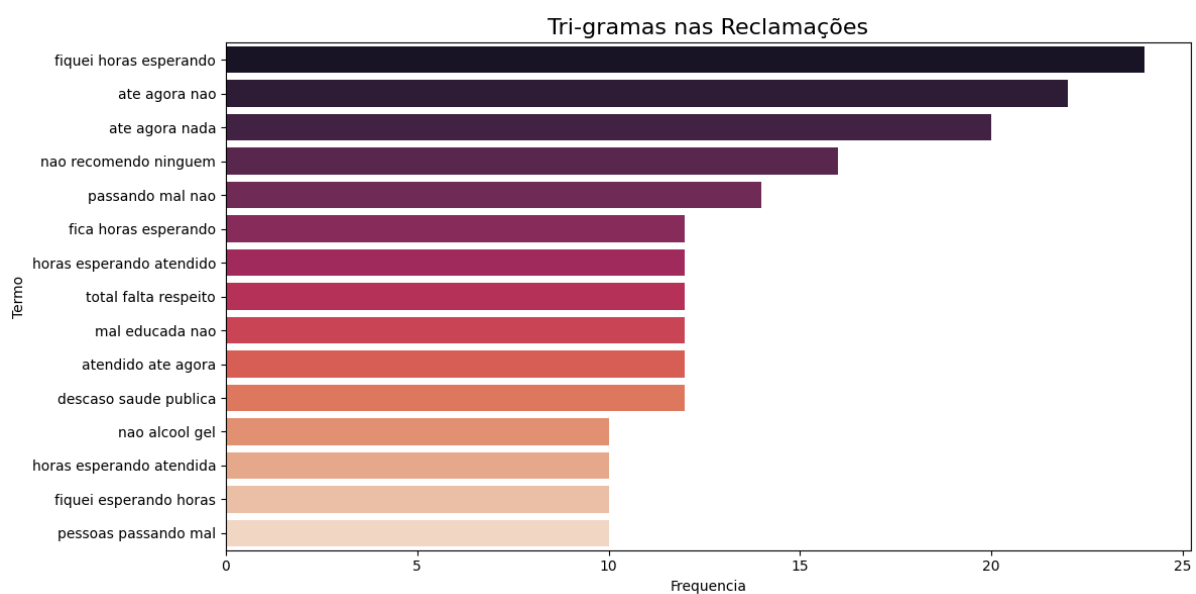


Figura 6: Tri-gramas mais frequentes nas avaliações negativas, indicando problemas de fluxo.

Complementarmente, a Figura 7 reforça a correlação entre os termos "horas" e "espera", além da recorrência de termos como "falta respeito".

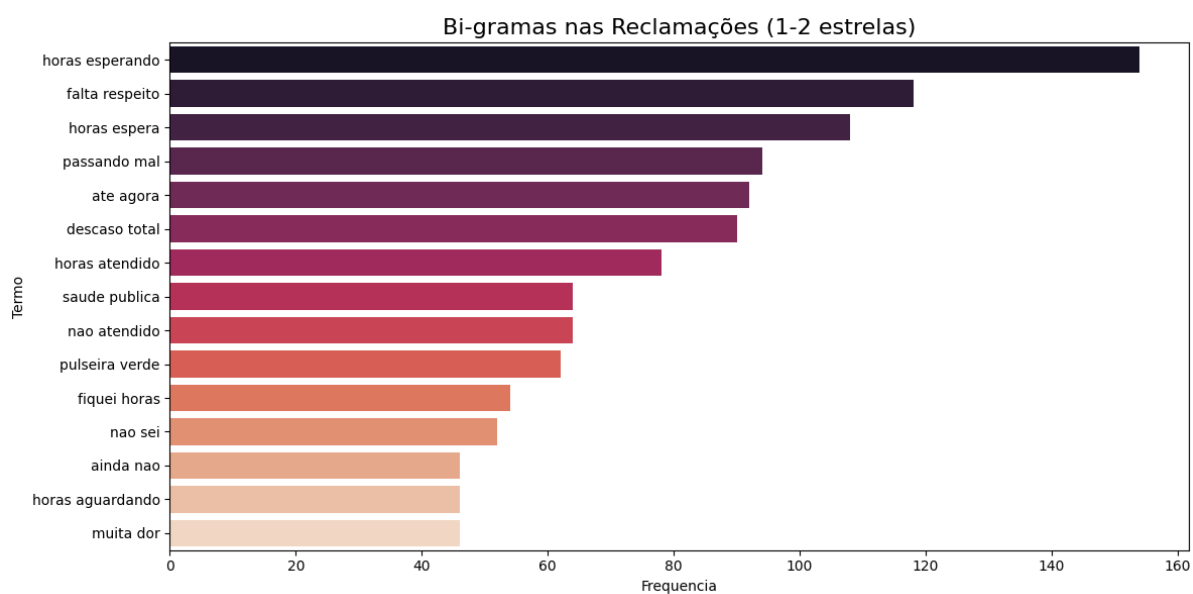


Figura 7: Bi-gramas mais frequentes nas reclamações (1 e 2 estrelas).

4.4. Estudo de Caso: A Correlação Financeira

Realizou-se o cálculo do Coeficiente de Correlação de Pearson entre a série temporal de “Nota Média Mensal” e “Despesa Mensal Executada”.

Ao isolar os dados do **Hospital Odilon Behrens (HOB)**, encontrou-se uma correlação **negativa de -0.22**.

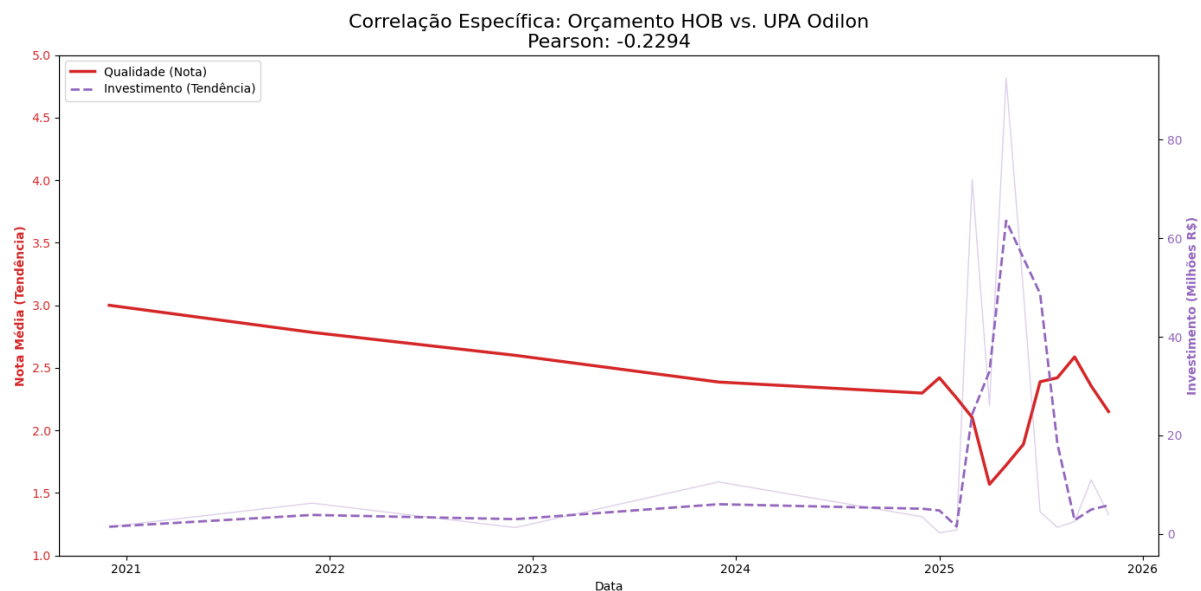


Figura 8: Correlação entre Investimento (Roxo) e Qualidade (Vermelho) no HOB.

Este resultado sugere um fenômeno de **“Gasto Reativo”**: os picos de despesa coincidem com os momentos de pior avaliação, indicando gastos para contenção de crises (superlotação) e não investimentos preventivos.

5. Conclusão

Este trabalho demonstrou a viabilidade do uso de dados de redes sociais para a gestão da saúde pública. A metodologia de *scraping* diagnosticou crises com precisão. Os resultados refutam a hipótese de que o aumento de volume financeiro garante satisfação imediata, apontando para problemas de gestão de fluxo. O código fonte disponível publicamente permite a replicação do estudo para outros municípios.

Os resultados obtidos foram validados externamente pelos eventos observados em 2025, que corroboram as tendências apontadas pelos modelos de dados. A queda acentuada na nota média a partir de 2024 e a explosão de termos como “Demora” e “Atendimento” na análise textual coincidem com o surto de Síndrome Respiratória Aguda Grave (SRAG) e a decretação de emergência em saúde pública pela prefeitura em maio de 2025.

Adicionalmente, a correlação negativa identificada no Hospital Odilon Behrens é explicada pelo déficit financeiro mensal de R\$ 40 milhões revelado pela imprensa em setembro de 2025. O aumento das despesas observado nos dados oficiais não refletiu melhorias, mas sim o pagamento de repasses atrasados e despesas emergenciais para evitar o colapso do atendimento. Por fim, a retirada da Guarda Municipal das UPAs em agosto de 2025 gerou um aumento na percepção de insegurança, refletindo-se em avaliações mais críticas e na insatisfação generalizada captada por este estudo.

Referências

- Dredze, M. (2012). How social media will change public health. *IEEE Intelligent Systems*, 27(4):81–84.
- Silva, T., Viana, A., and Benevenuto, F. (2019). Urban mobility in the age of social media. In *Proceedings of the AAAI Conference on Web and Social Media*, volume 13, pages 1–10.
- Prefeitura de Belo Horizonte (2025). Portal de Dados Abertos da PBH. Disponível em: <https://dados.pbh.gov.br/>. Acesso em: 01 dez. 2025.