

PADRÕES DE MORTALIDADE E DISCRIMINAÇÃO RACIAL: UM ESTUDO DE AGRUPAMENTOS NA REGIÃO METROPOLITANA FLUMINENSE

Fernanda Silva de Oliveira¹

¹Universidade Federal Rural do Rio de Janeiro, Rio de Janeiro, Brasil (hd.fernandaoliveira@gmail.com)

Resumo: Este estudo investiga disparidades raciais na mortalidade por homicídios, câncer de mama e causas maternas no Rio de Janeiro (2015-2022). Via algoritmo K-Modes e metodologia da área de Humanidades Digitais, identificaram-se perfis de vulnerabilidade entre pretos e pardos. Os resultados indicam que o racismo estrutural gera óbitos precoces e evitáveis. Conclui-se pela necessidade de políticas interseccionais e da execução da PNSIPN para mitigar iniquidades históricas.

Palavras-chave: Iniquidade racial; Saúde da população negra; K-Modes; Rio de Janeiro; Humanidades Digitais

INTRODUÇÃO

A organização social do Brasil contemporâneo é indissociável de seu passado colonial e escravocrata, que sedimentou o racismo estrutural como um eixo ordenador das relações de poder e do acesso a recursos. Esse fenômeno reflete-se diretamente na saúde pública, onde as desigualdades sociais e raciais determinam perfis epidemiológicos distintos para a população branca e a população negra (pretos e pardos) (Travassos; Noronha, 2007). No sistema de saúde, a cor da pele frequentemente atua como uma barreira simbólica e prática, resultando em iniquidades que vão desde a precariedade no pré-natal até disparidades no tratamento de doenças crônicas (Leal et al., 2017).



Na Região Metropolitana do Rio de Janeiro, essas disparidades são agravadas por um cenário de segregação territorial e violência institucional. A mortalidade da população negra nesse contexto não é um dado meramente estatístico, mas o reflexo de falhas sistemáticas na implementação de políticas de equidade. Enquanto a juventude negra é vitimada pela letalidade violenta de forma desproporcional (Ramos et al., 2023), as mulheres negras enfrentam obstáculos específicos no acesso a diagnósticos oncológicos e cuidados maternos qualificados. Estudos indicam que o pertencimento racial influencia diretamente o prognóstico de patologias como o câncer de mama, onde o diagnóstico tardio e a biologia da doença se entrelaçam com fatores socioeconômicos (Stringer-Reasor et al., 2021).

Para compreender a complexidade desses padrões de óbito, as Humanidades Digitais oferecem ferramentas que permitem transpor a análise puramente descritiva. O uso de técnicas de Machine Learning e ciência de dados permite identificar agrupamentos de indivíduos com características similares, evidenciando como a raça se combina com a baixa escolaridade e a vulnerabilidade territorial (Barbieri Filho, 2021). A aplicação do algoritmo K-Modes, especificamente desenhado para lidar com variáveis categóricas comuns em bancos de dados de saúde, possibilita a extração de conhecimentos valiosos sobre a segmentação das causas de morte (Huang, 1998).

Diante deste panorama, o presente trabalho investiga as dinâmicas da mortalidade de pretos e pardos na Região Metropolitana Fluminense entre os anos de 2015 e 2022, utilizando dados do Sistema de Informações sobre Mortalidade (SIM) e indicadores de segurança pública (ISP, 2022). O objetivo central deste estudo é analisar, através da análise de *clusters*, os perfis de óbitos por homicídios, câncer de mama e causas maternas. Pretende-se, assim, visibilizar as lacunas de proteção estatal e fornecer subsídios para o fortalecimento da Política Nacional de Saúde Integral da População Negra (PNSIPN), buscando caminhos para a mitigação das mortes evitáveis que assolam esse grupo populacional.

MATERIAL E MÉTODOS

A presente investigação fundamenta-se em uma abordagem quantitativa e exploratória, ancorada nos preceitos das Humanidades Digitais. Esta área

interdisciplinar permite a integração de métodos computacionais avançados ao pensamento crítico das ciências humanas e sociais, possibilitando que grandes volumes de dados sejam interpretados sob a ótica das iniquidades raciais e sociais (Barbieri Filho, 2021). A pesquisa caracteriza-se como um estudo retrospectivo e descritivo, focado na análise de padrões de mortalidade da população negra na Região Metropolitana do Rio de Janeiro.

Para a execução deste estudo, os dados foram extraídos do Sistema de Informações sobre Mortalidade (SIM), gerido pelo Departamento de Informática do Sistema Único de Saúde (DATASUS). O recorte temporal compreende o período entre 2015 e 2022, intervalo que permite observar tendências recentes e os impactos de crises sociais e sanitárias na saúde da população negra. Adicionalmente, foram consultadas bases do Instituto de Segurança Pública (ISP, 2022) para contextualizar os indicadores de letalidade violenta no território fluminense.

A amostra foi constituída por óbitos de indivíduos autodeclarados pretos e pardos, seguindo a classificação racial do IBGE. Foram selecionadas três causas específicas de morte, identificadas no conjunto de dados utilizado: Homicídios; neoplasia maligna da mama (câncer de mama) e óbitos relacionados à maternidade (ocorridos durante a gestação, parto ou puerpério)

O tratamento dos dados e a análise estatística foram realizados utilizando a linguagem de programação R, em ambiente RStudio. O processo de limpeza (*data cleaning*) envolveu a padronização de campos, o tratamento de valores ausentes (*missing data*) e a seleção das variáveis de interesse, a saber: Perfil Sociodemográfico: Idade, sexo, raça/cor e escolaridade; Contexto Familiar/Laboral: Estado civil e ocupação;

A variável "escolaridade" foi tratada como um indicador central de vulnerabilidade, uma vez que a literatura aponta uma correlação direta entre o nível de instrução e o acesso a diagnósticos precoces e medidas preventivas em saúde (Travassos; Noronha, 2007).

A principal ferramenta analítica empregada foi a análise de agrupamentos (*clustering*). Diferente de estudos tradicionais que utilizam o K-Means, modelo de



aprendizado de máquina não supervisionado voltado para dados numéricos, esta pesquisa optou pelo algoritmo K-Modes. Esta escolha justifica-se pelo fato de as bases de dados de saúde e segurança pública serem compostas predominantemente por variáveis categóricas (como estado civil, raça e escolaridade).

O algoritmo K-Modes atua substituindo as médias de clusters por modas, utilizando medidas de dissimilaridade baseadas na correspondência simples de atributos (Huang, 1998). O objetivo do algoritmo é minimizar a função de custo, agrupando os óbitos que compartilham características semelhantes em "perfis de vulnerabilidade". Para determinar o número ideal de *clusters* (k), utilizou-se o método do cotovelo (Elbow Method), que analisa a redução da variância intra-grupo conforme o aumento do número de agrupamentos.

No que tange ao processamento computacional, a pesquisa utilizou a linguagem R, um ambiente do RStudio. Para a modelagem dos perfis de mortalidade, empregou-se o pacote *klaR*, que disponibiliza a implementação do algoritmo K-Modes necessária para o tratamento de variáveis categóricas. Por fim, a síntese visual dos agrupamentos e a geração de evidências gráficas sobre as iniquidades de saúde foram realizadas com a biblioteca *ggplot2*, permitindo uma representação clara das vulnerabilidades identificadas nos *clusters*.

Por se tratar de uma pesquisa que utiliza bases de dados secundárias, de acesso público e sem identificação nominal dos sujeitos, o estudo dispensa a submissão ao Comitê de Ética em Pesquisa, conforme as normativas nacionais. A análise dos resultados foi conduzida sob a perspectiva da interseccionalidade, cruzando os perfis gerados pelo K-Modes com os marcos teóricos da Política Nacional de Saúde Integral da População Negra (PNSIPN), buscando identificar as falhas no atendimento e as barreiras impostas pelo racismo institucional (Leal et al., 2017).

RESULTADOS E DISCUSSÃO

As inferências obtidas através da técnica de agrupamento expuseram as camadas estruturais e multifacetadas das disparidades raciais que permeiam a Região Metropolitana do Rio de Janeiro. A análise subsequente detalha os achados para cada uma das três dimensões de mortalidade investigadas: homicídios, neoplasia de

mama e causas maternas. Dessa forma, estabeleceu-se um diálogo entre os dados quantitativos, os determinantes sociais da saúde e o referencial bibliográfico pertinente.

Em relação aos óbitos por homicídios, os *clusters* formados evidenciaram uma forte clivagem demográfica, racial e de gênero, sublinhando a natureza seletiva da violência letal no território fluminense. Identificou-se um agrupamento majoritário, denominado Perfil 1, composto predominantemente por indivíduos do sexo masculino e de cor parda (75,1%), no qual a mortalidade por homicídio representa 53,5% dos casos. Sob a ótica geográfica, observa-se uma concentração acentuada deste perfil em municípios da Baixada Fluminense, com destaque para Japeri e Queimados (ambos com 71%), seguidos por Belford Roxo e Nova Iguaçu (66% cada). Como demonstrado na Figura 1, que mapeia a espacialização desses dados, o Perfil 1 (destacado na tonalidade alaranjada) localiza-se primordialmente nas zonas mais densamente urbanizadas da metrópole. Esse padrão cartográfico corrobora a tese de que a violência letal atinge de forma brutal a juventude negra residente em áreas marcadas por iniquidades socioeconômicas profundas e pela precarização do acesso a serviços e direitos fundamentais (Santos et al., 2021).

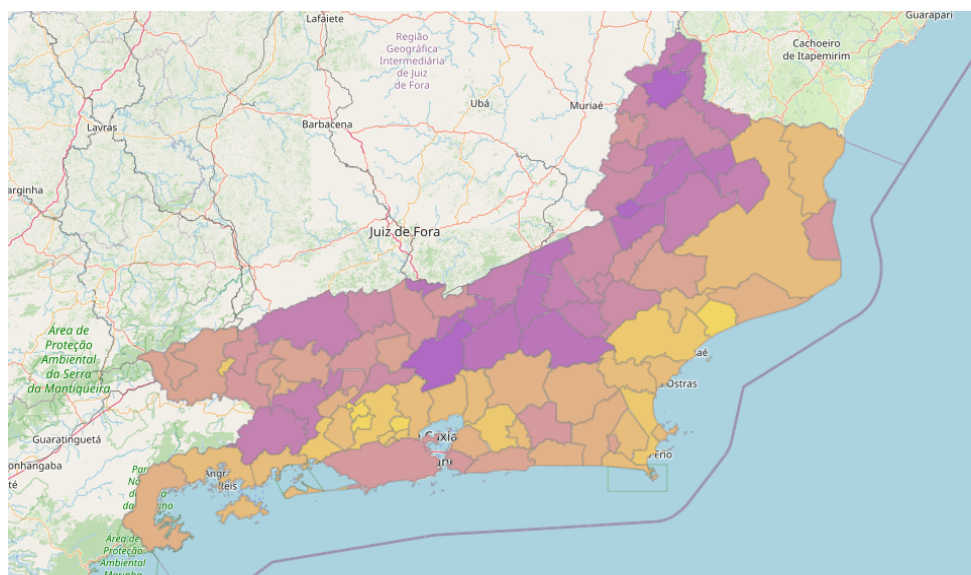


Figura 1. Mapa com a distribuição dos agrupamentos referentes aos óbitos de jovens homens por homicídio..

Em contrapartida à letalidade violenta observada no primeiro grupo, o Perfil 2 estabelece uma antítese sociodemográfica, sendo constituído majoritariamente por indivíduos do sexo masculino de cor branca (72%). Neste agrupamento, a causa predominante do óbito desloca-se das agressões para os eventos acidentais, que perfazem 57,8% das ocorrências. A Tabela 1, apresentada a seguir, sintetiza de forma comparativa as variáveis que definem cada um desses perfis, evidenciando as assimetrias fundamentais entre eles.

Tabela 1. Composição das variáveis nos agrupamentos de mortalidade por causas externas.

Variáveis	Perfil 1	Perfil 2
Gênero	Masculino (85,6%)	Masculino (66,8%)
Circunstância do Óbito	Homicídio (53,5%)	Acidente (57,8%)
Raça/cor	Pardos (75,1%)	Branco (72%)

Esta diferenciação é de extrema relevância para a compreensão da dinâmica social fluminense, uma vez que demonstra que a violência letal não se manifesta como um fenômeno aleatório ou puramente circunstancial. Pelo contrário, os dados reforçam a existência de um evento direcionado, que encontra eco em estudos contemporâneos sobre a "seletividade penal" e a vulnerabilidade da juventude negra, apontada como o principal alvo tanto da violência urbana quanto daquela perpetrada pelo braço armado do Estado (Ramos et al., 2023).

Nessa perspectiva, a preservação da vida dos jovens negros emerge como um dos pilares mais urgentes da Política Nacional de Saúde Integral da População Negra (PNSIPN). O histórico de extermínio e a marginalização desses corpos refletem uma herança estrutural onde o Estado, por vezes, falha em sua função protetiva. Estatísticas do Instituto de Segurança Pública (ISP) corroboram essa gravidade: em 2022, aproximadamente 30% das mortes registradas entre a população negra no estado do Rio de Janeiro tiveram como origem intervenções de agentes estatais (ISP, 2022). Tais evidências consolidam a tese de que marcadores como a cor da pele e a localização geográfica (territórios periféricos) funcionam como



determinantes que reduzem drasticamente as chances de sobrevivência deste grupo.

Ao transpor a análise para o campo das neoplasias malignas de mama, observa-se que as desigualdades raciais assumem contornos distintos, porém igualmente excludentes. Embora a literatura aponte uma incidência epidemiológica frequentemente elevada em mulheres brancas, os padrões de mortalidade revelados por esta pesquisa expõem a precariedade no acesso e as barreiras na qualidade do atendimento direcionado às mulheres negras.

A análise de clustering identificou quatro perfis heterogêneos para esta causa de óbito, conforme detalhado na Tabela 2. Os agrupamentos classificados como Perfil 1 e Perfil 2 são compostos de forma predominante por mulheres pardas e caracterizam-se por apresentarem os menores índices de instrução formal, concentrando-se nos níveis de Ensino Fundamental I e II.

Um dado identificado nesses perfis é a alta incidência de registros ignorados ou em branco para a variável "escolaridade". Este fenômeno não deve ser interpretado apenas como uma falha burocrática, mas sim como um sintoma da precariedade informacional em contextos de vulnerabilidade social. Nestes cenários, a desarticulação entre as redes de atenção e a falta de rigor no preenchimento de documentos de saúde pública comprometem a produção de evidências, dificultando a formulação de intervenções eficazes (Travassos; Noronha, 2007). A ausência desses dados mascara a realidade de mulheres que, devido ao racismo institucional, enfrentam itinerários terapêuticos fragmentados e diagnósticos tardios, o que impacta diretamente na letalidade da doença entre a população parda da Baixada Fluminense.

Já o Perfil 3 é constituído majoritariamente por mulheres brancas (94,9%), com níveis de instrução que abrangem o Ensino Médio e Superior, e óbitos registrados em idades avançadas, predominantemente acima de 71 anos. Esta configuração sugere uma correlação entre o status socioeconômico e a facilitação do acesso a serviços de saúde, o que favorece o diagnóstico precoce e a continuidade terapêutica, resultando em maior sobrevida. Em oposição, o Perfil 4 concentra mulheres pretas (53,9%), com escolaridade predominante de Ensino Médio e

mortalidade incidente em uma faixa etária mais jovem, entre 51 e 60 anos. A literatura indica que mulheres negras apresentam desfechos de mortalidade por câncer de mama mais desfavoráveis, frequentemente associados ao diagnóstico em estágios avançados e a impedimentos no fluxo de atendimento especializado (Stringer-Reasor et al., 2021).

Tabela 2. Distribuição de variáveis nos agrupamentos de óbitos por câncer de mama.

Variáveis	Perfil 1	Perfil 2	Perfil 3	Perfil 4
Raça/cor	Pardas (63,3%)	Pardas (60,6%)	Brancas (94,9%)	Preta (53,9%)
Escolaridade	Ignorado (33,9%)	Fundamental I (42,5%)	Ensino Médio (34,7%)	Ensino Médio (41,3%)
Faixa Etária (anos)	41 a 50 (51,4%)	61 a 70 (40%)	71 + (42,2%)	51 a 60 (62,8%)

Estudos apontam que a disparidade nos desfechos clínicos reflete trajetórias distintas de cuidado: enquanto mulheres brancas e com maior instrução formal tendem a ser acompanhadas de forma contínua, as mulheres negras enfrentam barreiras estruturais que reduzem as possibilidades de recuperação (Wheeler et al., 2013). O racismo institucional opera como um entrave na agilidade de exames de rastreamento e na oferta de terapias de alta complexidade, evidenciando a carência de protocolos de saúde que considerem as especificidades deste grupo populacional (Werneck, 2016). Nesse sentido, a desigualdade racial manifesta-se na forma como o sistema de saúde gerencia a jornada da paciente, exercendo influência direta nos índices de letalidade da neoplasia de mama.

A análise da mortalidade materna revelou agrupamentos delimitados por critérios de raça/cor e nível de instrução, evidenciando padrões de vulnerabilidade específicos para as mulheres negras durante o ciclo gravídico-puerperal. Conforme detalhado na Tabela 3, os perfis identificados permitem observar a estratificação dos riscos obstétricos na Região Metropolitana do Rio de Janeiro.

O Perfil 2 é constituído majoritariamente por mulheres pretas (60,8%), com nível de escolaridade concentrado no Ensino Fundamental II e faixa etária entre 30 e 40

anos. Este agrupamento evidencia as disparidades no acesso a informações e a serviços de assistência pré-natal de qualidade. A literatura associa esses indicadores à persistência de mortes evitáveis, muitas vezes resultantes de falhas na rede de atenção e de processos de discriminação no atendimento institucional (Leal et al., 2017).

Em uma configuração distinta, o Perfil 3 reúne predominantemente mulheres pardas (89,9%), em uma faixa etária mais jovem (20 a 30 anos) e com Ensino Médio completo. Já os Perfis 1 e 4 são formados majoritariamente por mulheres brancas, com representações de 65,4% e 82,6%, respectivamente, também apresentando o Ensino Médio como grau de instrução predominante, porém com variações nas faixas etárias.

Tabela 3. Distribuição de variáveis nos agrupamentos de óbitos maternos.

Variáveis	Perfil 1	Perfil 2	Perfil 3	Perfil 4
Raça/cor	Brancas (65,4%)	Pretas (60,8%)	Pardas (89,9%)	Brancas (82,6%)
Escolari-da de	Ensino Médio (54%)	Fundamental II (64,1%)	Ensino Médio (46,5%)	Ensino Médio (49,8%)
Faixa Etária (anos)	30 a 40 (88%)	30 a 40 (61,9%)	20 a 30 (58,7%)	20 a 30 (81,3%)

Os dados encontrados corroboram a tendência nacional de que as mulheres negras figuram como as principais vítimas da mortalidade materna. No contexto do Rio de Janeiro, a razão de mortalidade materna entre mulheres pretas chega a ser 4,3 vezes superior à observada entre mulheres brancas (Alves; Guimarães, 2021). Tais indicadores refletem o acesso desigual aos serviços obstétricos e a incidência de negligência institucional durante o pré-natal, o parto e o puerpério (Brasil, 2009).

A continuidade de óbitos por causas preveníveis nos grupos de maior vulnerabilidade racial indica que o racismo institucional atua como um determinante nos desfechos de saúde. Os perfis mapeados reforçam a necessidade de viabilizar a execução da Política Nacional de Saúde Integral da População Negra (PNSIPN), com ênfase em estratégias que considerem as assimetrias sociais e territoriais que incidem sobre a vida e a morte da população negra no estado



CONCLUSÃO

Este estudo demonstra que a aplicação do algoritmo K-Modes, sob a ótica das Humanidades Digitais, é uma ferramenta estratégica para converter grandes volumes de dados de saúde em diagnósticos sociais precisos. A identificação de agrupamentos específicos permitiu evidenciar que a mortalidade na Região Metropolitana do Rio de Janeiro não é distribuída de forma aleatória; ela possui uma assinatura racial e territorial que define quem são as principais vítimas da violência letal e das negligências no sistema de saúde.

Os resultados comprovam que o racismo estrutural atua como um determinante social que encurta a vida da população negra em diferentes ciclos. Na segurança pública, a seletividade da violência letal concentra-se em jovens negros periféricos, enquanto na saúde, o racismo institucional manifesta-se no diagnóstico tardio do câncer de mama e em óbitos maternos evitáveis. Esses padrões de vulnerabilidade, revelados pela análise de clusters, validam a necessidade de abandonar abordagens universalistas e adotar estratégias de cuidado que considerem as especificidades de raça, gênero e território.

Apesar da eficácia do modelo, a pesquisa ressalta que a qualidade das políticas públicas baseadas em evidências depende diretamente da qualidade dos registros administrativos. A recorrência de campos ignorados na variável "escolaridade" é um entrave que mascara iniquidades. Portanto, é fundamental que o Estado invista na capacitação dos agentes responsáveis pelo preenchimento de documentos como a Declaração de Óbito, garantindo que o Sistema de Informações sobre Mortalidade (SIM) reflita com fidelidade as lacunas de proteção social.

Dessa forma, os dados produzidos pela pesquisa reforçam a urgência da implementação plena da Política Nacional de Saúde Integral da População Negra (PNSIPN). O uso de inteligência de dados deve servir como subsídio para que a gestão pública direcione recursos e protocolos sensíveis à equidade racial, transformando a lógica de funcionamento das instituições para que o sistema de saúde seja, de fato, equânime e não discriminatório. Mais do que identificar padrões,



o objetivo final desta pesquisa é oferecer suporte científico para intervenções que busquem interromper o ciclo histórico de mortes evitáveis da população negra fluminense.

REFERÊNCIAS

- Alves, L. G. R.; Guimarães, R. M. (2021). Race inequalities in maternal mortality in the city of Rio de Janeiro, Brazil: 2010-2019. *Revista da Associação Médica Brasileira*, 67(1), 120–124.
- Brasil. Ministério da Saúde. Portaria nº 992, de 13 de maio de 2009. Institui a Política Nacional de Saúde Integral da População Negra. *Diário Oficial da União*, Brasília, DF, 14 maio 2009.
- Brasil. Ministério da Saúde. Sistema de Informações sobre Mortalidade (SIM). Disponível em: <https://dados.gov.br/dados/conjuntos-dados/sim-1979-2019>. Acesso em: 22 jul. 2025.
- Dos Santos, J. R. R.; Dias, C. M.; Filho, A. C. Machine learning and national health data to improve evidence: Finding segmentation in individuals without private insurance. *Health Policy and Technology*, v. 10, n. 1, p. 79–86, mar. 2021.
- Huang, Z. (1998). Extensions to the k means algorithm for clustering large datasets with categorical features. *Data Mining and Knowledge Discovery*, 2(3), 283-304.
- Instituto de Segurança Pública (ISP). (2022). ISP Dados Visualização. Grupos Vulneráveis. Disponível em: <https://www.ispvisualizacao.rj.gov.br/>. Acesso em: 15 mar. 2025.
- Leal, M. D. C. et al. (2017). A cor da dor: iniquidades raciais na atenção pré-natal e ao parto no Brasil. *Cadernos de Saúde Pública*, 33(supl 1).
- Ramos, S. et al. (2023). *Pele Alvo: a bala não erra o negro*. Rio de Janeiro: CESeC.
- Stringer-Reasor, E. M. et al. (2021). Disparities in Breast Cancer Associated With African American Identity. *American Society of Clinical Oncology Educational Book*, 41, e29–e46.
- Travassos, C.; Noronha, J. C. (2007). Desigualdades sociais e saúde no Brasil: uma análise da produção científica. *Ciência & Saúde Coletiva*, 12(6), 1477-1488.



- Wheeler, S. B., Reeder-Hayes, K. E., & Carey, L. A. (2013). Disparities in Breast Cancer Treatment and Outcomes: Biological, Social, and Health System Determinants and Opportunities for Research. *The Oncologist*, 18(9), 986–993.
- Werneck, J. (2016). Racismo institucional e saúde da população negra. In: *Saúde da população negra no Brasil: uma abordagem interseccional*. Rio de Janeiro: Fiocruz.