

Análise de Dados DATASUS

Arthur Luz, Leonardo Ripes, Morgana Weber e Vitor Garcez

¹Escola Politécnica – Pontifícia Universidade Católica do RS (PUCRS)
Av. Ipiranga, 6681 – 90.619-900 – Porto Alegre – RS – Brasil

{arthur.musskopf, leonardo.ripes, morgana.weber, vitor.hugo}@edu.pucrs.br

Resumo. *Este trabalho tem como intuito explorar dados públicos do DATASUS e IBGE para analisar os índices de suicídio no Brasil durante os anos principais da pandemia de COVID-19 (2020 e 2021), comparando-os com o ano de 2019. Dessa forma, também será levado em conta as informações de raça-cor, a fim de verificar se houve alguma parcela mais atingida na sociedade. Este artigo compõe a disciplina de Coleta, Preparação e Análise de Dados, da Escola Politécnica da Pontifícia Universidade Católica do Rio Grande do Sul.*

1. Introdução

O DATASUS (Departamento de Informática do Sistema Único de Saúde) surgiu em 1991, e passou a exercer a função de controle e processamento das contas referentes à saúde, que antes era da Empresa de Tecnologia e Informações da Previdência Social (DATA-PREV). A partir disso, foi definida a responsabilidade do DATASUS de prover os órgãos do SUS de sistemas de informação e suporte de informática, necessários ao processo de planejamento, operação e controle [DATASUS 2023].

Dentre as competências descritas para o DATASUS está a competência de desenvolver, pesquisar e incorporar tecnologias de informática que possibilitem a implementação de sistemas e a disseminação de informações necessárias às ações de saúde [DATASUS 2023].

O sistema TABNET disponibiliza os dados de forma pública, de modo que possam ser utilizados para subsidiar análises objetivas da situação sanitária do país. Além disso, podem auxiliar nas tomadas de decisão e na definição de políticas públicas, servindo como base de evidências para a elaboração de programas de ações de saúde [DATASUS 2023].

Tendo em vista a disponibilidade pública dos dados de saúde da população brasileira do Sistema Único de Saúde (SUS), optou-se por analisar dados gerados durante a pandemia para verificar se houve aumento nas taxas de suicídios ao longo da pandemia de COVID-19 (anos 2020 e 2021). Como *baseline*, foi utilizado o ano de 2019, pré-pandemia. Além disso, foram utilizados dados demográficos do IBGE, a fim de utilizar as informações de raça-cor e analisar como cada parcela dessa população foi afetada.

2. Base de Dados

Para a seleção dos dados foi utilizada a biblioteca *PySUS*, uma biblioteca de auxílio para a análise dos dados do SUS [Coelho]. Assim, optou-se por utilizar a biblioteca para *download* dos dados, visto que a plataforma do DATASUS possui algumas limitações como por exemplo, o excesso de filtros, impedindo uma visão macro.

Tabela 1. Significado das colunas do *Dataframe*

Coluna	Significado
RACACOR	Raça/cor referenciada como 1(branca), 2(preta), 3(amarela) 4(parda) e 5(indígena)
SEXO	Gênero referenciado como 0 (ignorado), 1 (masculino) e 2 (feminino)
DTOBITO	Data do óbito
DTNASC	Data do nascimento
UF	Unidade Federativa
IDADE_ANOS	Idade em anos, calculada com a data de nascimento e óbito
CIRCOBITO	Circunstância do óbito referenciada por: 9(ignorado), 1(acidente) 2(suicídio), 3(homicídio) e 4(outros)

Desse modo, foram coletados apenas dados necessários para realizar a avaliação, sendo eles raça-cor, sexo, data do óbito, data de nascimento, Unidade Federativa (UF) e a circunstância do óbito.

2.1. Pré-processamento

A coluna “CIRCOBITO”, referente à circunstância do óbito, possuía identificadores para cada tipo de óbito. Para o caso do presente trabalho, foram selecionados apenas os dados contendo o valor 2 para “CIRCOBITO”, valor referente para os casos de óbito por suicídio. Assim obteve-se o *Dataframe* contendo as informações detalhadas na Tabela 1, que seguem a documentação disponibilizada pela Fiocruz.

Também foram coletados os dados da população atual de cada ano para cada UF. Neste caso o número total de pessoas foi multiplicado por 1.000, pois os valores populacionais eram a cada 1.000 pessoas. Nos dados da população atual haviam apenas os percentuais de pessoas por raça/cor, sendo necessário extrair este valor de percentual para numérico.

Foram eliminados 500 casos de raça não identificada, que representavam 1,23% do total dos dados obtidos até então. Por fim, foi adicionada uma coluna chamada “Taxa de suicídio”, que é o cálculo do total de suicídios em determinada UF e determinado ano dividido pelo total da população em determina UF em determinado ano. Para gerar o infográfico na ferramenta *Power BI*, os dados de censo e população foram unificados em um tabela. O código completo foi disponibilizado na plataforma *GitHub*¹.

3. Análise dos Dados

Após a extração dos dados e a visualização dos comportamentos através de gráficos no *Power BI*, foi possível verificar alguns *insights* sobre o aumento da proporção de suicídios ao longo dos 3 anos analisados, 2019, 2020 e 2021 e analisar se houve algum grupo mais afetado, como por exemplo, homens ou mulheres, pessoas pretas, brancas ou pardas.

Durante a pandemia referentes aqui aos anos de 2020 e 2021, a população brasileira aumentou em 3,6%, ou seja, em proporção, a população brasileira se manteve constante em relação a taxa populacional entre os anos (Figura 1). Em proporção populacional de cada Estado, os estados com maiores índices de suicídio acumulado entre os anos estão: Rio Grande do Sul, Piauí e Santa Catarina (Figura 2).

¹ <https://github.com/Leonardo-ripes/TF-Coleta/tree/main>

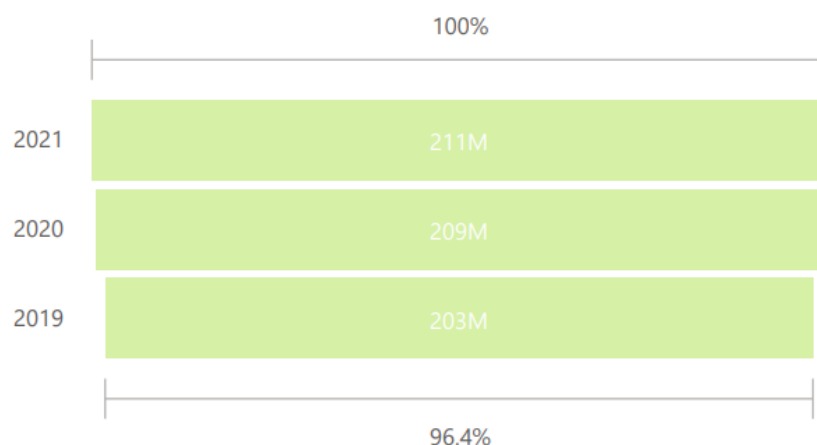


Figura 1. Crescimento populacional por ano

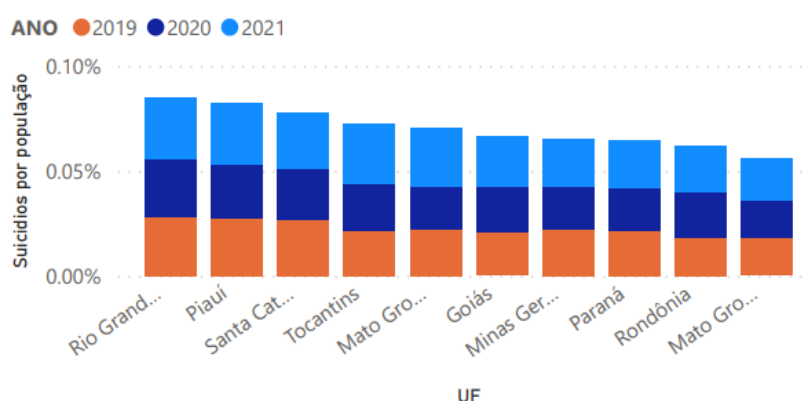


Figura 2. Top 10 estados com maior taxa de suicídios

Pôde-se observar também que a pandemia influenciou na quantidade de suicídios ocorridos durante os anos (Figura 3). Em comparação, 2021 obteve um salto de 16% em relação ao primeiro ano de pandemia (2020). Ao longo de 2020 e 2021, notou-se que o Amazonas foi a região com maior crescimento no índice de suicídios, seguido por Rio de Janeiro, Alagoas e Bahia (Figura 4).

Analisando a proporção de suicídios por etnia e ano, pode-se observar que em 2021 o índice de pessoas pardas cometendo suicídio aumentou em 4% comparado ao ano de 2019. Da mesma forma a porcentagem de pessoas brancas se suicidando diminuiu.

Quando analisado por gênero, o índice de suicídios se manteve proporcional. Assim, pode-se descartar a hipótese de que um dos gêneros foi mais afetado que outro em relação a suicídios durante a pandemia.

Por fim, pode-se perceber que o grupo com maior número de suicídios são jovens de 21 a 40 anos de idade (Figura 7). Entretanto, não houve nenhuma alteração significativa na taxa de suicídios por idade pré-pandemia e pandemia.

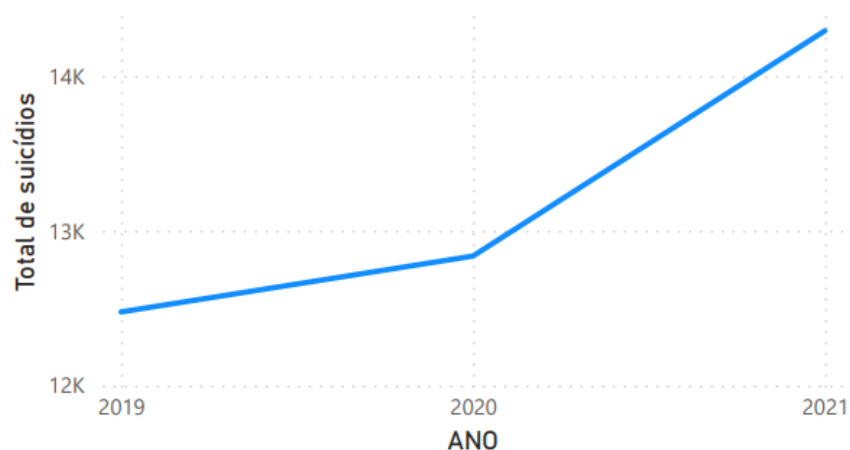


Figura 3. Suicídios por ano

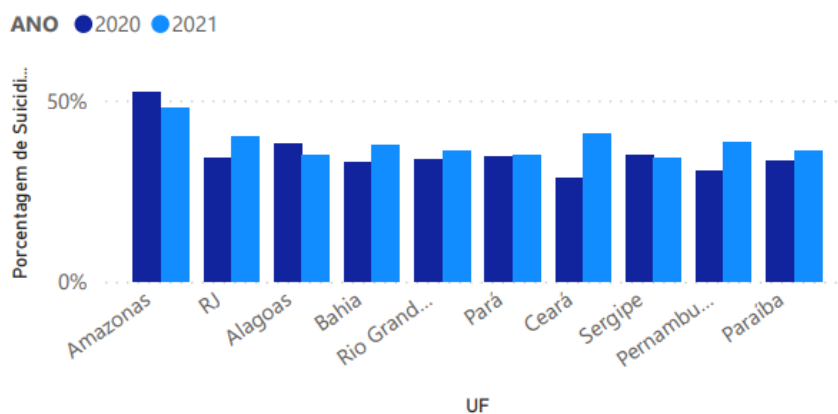


Figura 4. Top 10 estados com maior crescimento no índice de suicídios

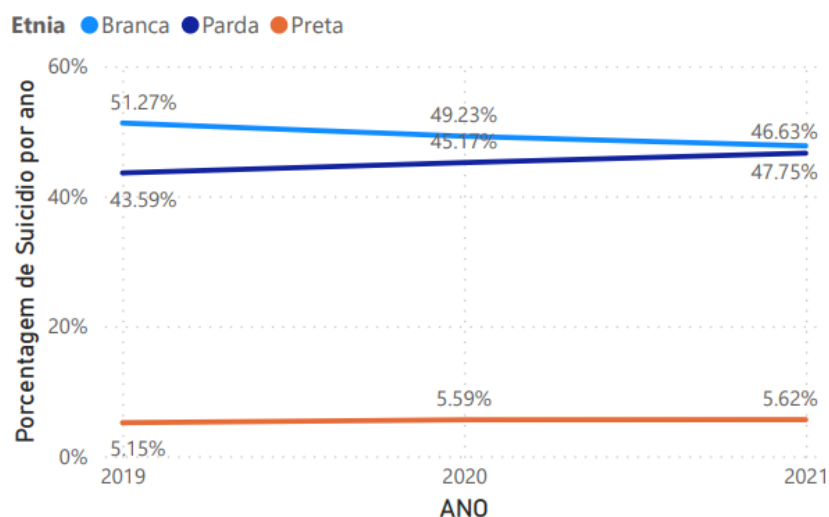


Figura 5. Índice de suicídios agrupado por etnia e ano

4. Conclusão

Os Estados mais pobres obtiveram os maiores aumentos em relação ao índice de suicídios e a população. Além disso, se enquadram também neste aumento, a população preta e

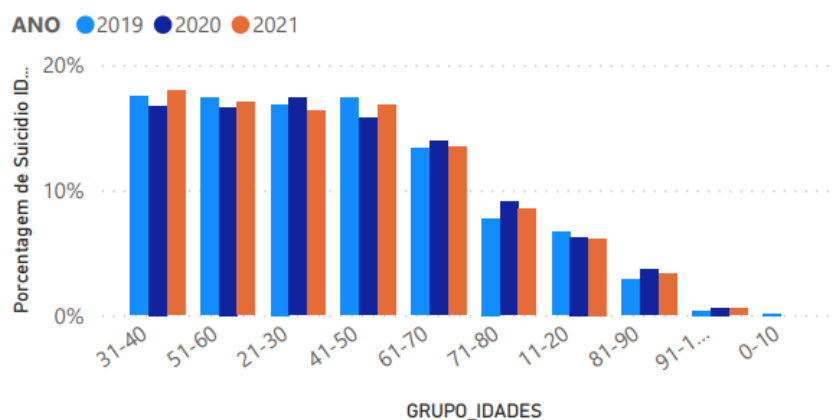


Figura 6. Índice de suicídios agrupado por gênero e ano

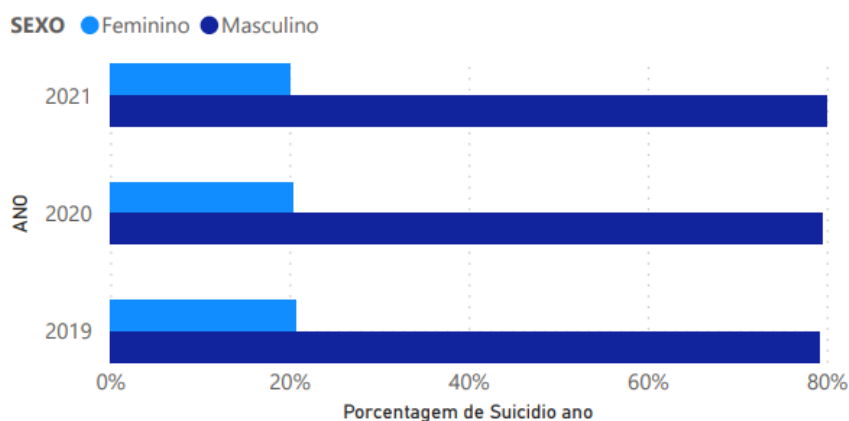


Figura 7. Índice de suicídios agrupado por idade e ano

parda. Portanto, para futuras pesquisas pode-se aprofundar utilizando dados demográficos como por exemplo a renda per capita para determinar se a população mais pobre foi mais afetada em relação aos suicídios ou não.

Referências

Coelho, F. C. Pysus's documentation. <https://pysus.readthedocs.io/en/latest/>. 25 de novembro de 2023.

DATASUS (2023). Histórico.