



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS SOCIAIS APLICADAS
COORDENAÇÃO DE CIÊNCIAS ECONÔMICAS

**ANÁLISE DEMOGRÁFICA E SOCIOECONÔMICA DA DISSEMINAÇÃO DO
CORONAVIRUS NO BRASIL UTILIZANDO A PNAD COVID19**

Recife
Setembro/2021

UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS SOCIAIS APLICADAS
COORDENAÇÃO DE CIÊNCIAS ECONÔMICAS

**ANÁLISE DEMOGRÁFICA E SOCIOECONÔMICA DA DISSEMINAÇÃO DO
CORONAVIRUS NO BRASIL**

Trabalho de Conclusão de Curso apresentado à
Universidade Federal de Pernambuco como
requisito para a obtenção do título de bacharel em
Ciências Econômicas.

Discente: Luiz José Carneiro Leão Neto

Orientador: Prof. José Ricardo Bezerra Nogueira

Recife
Setembro/2021

RESUMO

Utilizando dados da PNAD COVID-19, este trabalho visa entender como determinadas características demográficas e socioeconômicas da população se correlacionam com a contaminação pelo Sars-Cov-2, responsável pela pandemia que se estende de 2020 até os dias atuais. É possível verificar que certas características possuem correlação positiva e, dentre essas, umas são mais fortemente relacionadas do que outras, enquanto algumas são negativamente associadas, também em diferentes escalas.

Palavras-chave: COVID-19, população, características demográficas e socioeconômicas, correlação.

ABSTRACT

Using data from PNAD COVID-19, this work aims to understand how certain demographic and socioeconomic characteristics of the population correlate with contamination by Sars-Cov-2, responsible for the pandemic that extends from 2020 to the present day. It is possible to verify that certain characteristics have a positive correlation and, among these, some are more strongly related than others, while some are negatively associated, also at different scales.

Keywords: COVID-19, population, demographic and socioeconomic characteristics, correlation.

Sumário

RESUMO	3
ABSTRACT	4
1. INTRODUÇÃO	6
2. OBJETIVOS E ESTUDOS ANTERIORES.....	8
3. ESTRUTURA DA MONOGRAFIA	9
4. REVISÃO DA LITERATURA	10
4.1 CORONAVÍRUS.....	10
4.2 CASOS POSITIVOS	10
4.3 PANDEMIA.....	11
4.4 AUXÍLIO EMERGENCIAL.....	12
5. MÉTODO.....	13
5.1 BASE DE DADOS.....	13
5.2 VARIÁVEIS UTILIZADAS	14
5.2.1 VARIÁVEL DEPENDENTE	14
5.2.2 VARIÁVEIS INDEPENDENTES.....	15
5.2.3 VARIÁVEIS DESCRITIVAS.....	16
5.3 MODELO ESTATÍSTICO	17
5.4 SOFTWARES.....	18
6. IMPLEMENTAÇÃO	19
6.1 ANÁLISE DESCRITIVA	19
6.2 ANÁLISE DE REGRESSÃO LOGÍSTICA.....	25
7. DISCUSSÃO DOS RESULTADOS	28
8. CONSIDERAÇÕES FINAIS	31
9. REFERÊNCIAS	32

1. INTRODUÇÃO

O primeiro caso de COVID-19 foi reportado na cidade de Wuhan, na China, em dezembro de 2019 (World Health Organization, 2020). Desde então, a doença se transformou em uma pandemia no ano de 2020, chegando a 2021 com mais de 220 milhões de casos confirmados ao redor do mundo e culminando em mais de 4,6 milhões de óbitos.¹

No Brasil, o primeiro caso do novo coronavírus foi confirmado no dia 26 de fevereiro de 2020². Mais de um ano se passou e o país se tornou o terceiro colocado em número de casos, ultrapassando 20 milhões, e o segundo em número de óbitos, se aproximando dos 600 mil, ainda segundo dados da Organização Mundial de Saúde (OMS).

Os números gerais são lastimáveis e o vírus “não apresentou seletividade contagiosa”, mas a contaminação é sentida de forma diferente pelas pessoas e círculos sociais, de acordo com raça, classe e gênero, como contou Fernanda Matheus Estrela et al (2020, p.3). Dessa forma, uma grande preocupação por parte dos pesquisadores é tentar entender razões e características que tornam algumas pessoas mais suscetíveis à contaminação pela doença, sejam essas de natureza biológica, socioeconômica ou demográfica. Assim:

A crítica à concepção universalista sobre os sujeitos sociais, o espaço e o movimento considera a necessidade de estabelecer relações com outros marcadores sociais, como raça, gênero, classe social, sexualidade, territórios e dinâmica social e econômica. Ou seja, analisar e intervir sobre os fenômenos decorrentes da circulação e transmissão do Sars-CoV-2 não se resume a identificar o vírus, compreender sua disseminação e controlá-lo. A colocação em cena da COVID-19 em diferentes contextos, espaços e linguagens, especialmente em situações de extrema desigualdade socio sanitária, expõe a multiplicidade e especificidade do fenômeno pandêmico desde sua dimensão macrossocial até a capilaridade micropolítica nas formas e estratégias de produção do cotidiano. (MATTA, 2021).

Identificadas pelas autoridades responsáveis por guiar a sociedade por essa pandemia, já foram observadas características que se mostraram associadas a maiores riscos de contaminação em outras crises sanitárias. Se não a maior, certamente uma das maiores preocupações da OMS é para com os idosos, que segundo a organização

¹ Dados obtidos no website da Organização Mundial de Saúde (OMS). Disponível em: <https://COVID19.who.int>. Acesso em 09 de setembro de 2021.

² Disponível em: <https://saude.abril.com.br/medicina/coronavirus-primeiro-caso-brasil/>. Acesso em 09 de setembro de 2021.

enfrentam mais elevadas ameaças e desafios³. Pessoas acima de 60 anos de idade possuem uma maior probabilidade de apresentarem casos mais graves da doença, bem como uma taxa de fatalidade muito maior pela COVID-19 (WHO, 2020).

Outra preocupação por parte da OMS é com a chamada “população vulnerável”, grupo que inclui pessoas em situação de rua, vivendo em moradias superlotadas, favelas e comunidades, trabalhadores migrantes, refugiados, pessoas com deficiência, pessoas vivendo em lugares fechados, em localizações remotas e vivendo na pobreza ou extrema pobreza. Dentre as causas citadas pela entidade para a preocupação com essa parcela da população estão a menor possibilidade de acesso a serviço de saúde de qualidade, maior probabilidade de exposição ao vírus e menor alcance a informações adequadas (WHO, 2020).

Assim, dada a preocupação apresentada por pesquisadores, este estudo analisou a existência ou não de uma associação entre determinadas características da população com a contaminação pela COVID-19 no Brasil, utilizando dados da PNAD COVID-19.

³ Disponível em: <https://www.euro.who.int/en/health-topics/health-emergencies/coronavirus-COVID-19/publications-and-technical-guidance/vulnerable-populations/health-care-considerations-for-older-people-during-COVID-19-pandemic>. Acesso em 09 de setembro de 2021.

2. OBJETIVOS E ESTUDOS ANTERIORES

A Pesquisa Nacional por Amostra de Domicílios COVID-19 (PNAD COVID-19) foi um instrumento criado pelo Instituto Brasileiro de Geografia e Estatística (IBGE) em 2020, com o objetivo de monitorar o avanço da pandemia e seu impacto no país, principalmente no mercado de trabalho⁴.

Dessa forma, a maioria dos trabalhos previamente realizados que utilizaram a pesquisa, apresentaram temas alinhados ao seu principal objetivo, com análises mais voltadas ao mercado de trabalho e às desigualdades nele existentes.

Por sua vez, este trabalho busca adicionar ao debate, objetivando entender como características demográficas e socioeconômicas se correlacionam com a contaminação pelo coronavírus no Brasil, utilizando dados da PNAD COVID-19, mais especificamente caracterizar a população acometida pela doença e analisar associações existentes entre as características demográficas e socioeconômicas e a contaminação pelo vírus.

⁴ Disponível em: https://www.ibge.gov.br/estatisticas/investigacoes-experimentais/estatisticas-experimentais/27946-divulgacao-semanal-pnadCOVID1?t=o-que-e&utm_source=COVID19&utm_medium=hotsite&utm_campaign=COVID_19. Acesso em 09 de setembro de 2021.

3. ESTRUTURA DA MONOGRAFIA

Esta monografia foi estruturada em três grandes blocos, que foram subdivididos em tópicos. O primeiro, apresentando o trabalho, contém a introdução, os objetivos, a comparação com estudos anteriores e o referencial teórico ou revisão de literatura. No segundo se encontra a descrição do trabalho, explicando o método utilizado, como foram criadas as bases de dados para os cálculos, tanto da análise descritiva quanto da análise de regressão, bem como os *softwares* usados. A terceira e última parte traz os resultados e discute os valores encontrados, concluindo o estudo.

Assim, a estrutura da monografia ficou da seguinte forma:

1. INTRODUÇÃO;
2. OBJETIVOS E ESTUDOS ANTERIORES;
3. ESTRUTURA DA MONOGRAFIA;
4. REVISÃO DA LITERATURA;
5. MÉTODO;
6. IMPLEMENTAÇÃO;
7. DISCUSSÃO DOS RESULTADOS;
8. CONSIDERAÇÕES FINAIS;
9. REFERÊNCIAS.

4. REVISÃO DA LITERATURA

4.1 CORONAVÍRUS

Segundo a OMS⁵:

A COVID-19 é uma doença infecciosa causada por um tipo recém-descoberto de coronavírus. A maior parte das pessoas infectadas pelo vírus sentirão dificuldades respiratórias de forma leve a moderada e não precisarão de tratamentos especiais para se recuperar. Idosos e pessoas com comorbidades como doenças cardiovasculares, diabetes, doenças respiratórias crônicas e câncer têm maior probabilidade de desenvolverem casos mais graves. (Tradução nossa)

A COVID-19 é causada pelo Sars-Cov-2, que pertence a uma família de outros vírus já conhecidos. Alguns integrantes dessa família já afetaram seres humanos e causaram surtos em algumas regiões, como o Mers-Cov no Oriente Médio e o Sars-Cov na China, mas nenhum deles na proporção desse que ficou conhecido como “novo” coronavírus.⁶

A nomenclatura “SARS” vem do termo em inglês que significa Síndrome Respiratória Aguda Grave (no original: Severe Acute Respiratory Syndrome). A OMS informa que, geralmente, o primeiro sintoma apresentado por indivíduos contaminados pelo vírus é febre, que pode vir acompanhada de dores de cabeça, no corpo e indisposição. Em um segundo momento, se desenvolvida a doença, o indivíduo pode passar a apresentar dificuldades respiratórias e, desses, entre 10% e 20% necessitam de intubação ou respiração mecânica. Ainda segundo a organização, a taxa de mortalidade é estimada em 3% das pessoas que apresentaram quadros suspeitos e prováveis de SARS.⁷

4.2 CASOS POSITIVOS

Em guia publicado no dia 11 de setembro de 2020, para casos suspeitos de COVID-19, a OMS recomenda a realização de Testes de Amplificação de Ácido Nucleicos como os de Transcrição Reversa Seguida de Reação em Cadeia de Polimerase, significados das siglas em inglês NAAT e RT-PCR, respectivamente (World Health Organization, pg. 4, 2020). Ainda segundo a organização, o paciente pode ser considerado

⁵ Disponível em: https://www.who.int/health-topics/coronavirus#tab=tab_1. Acesso em: 09 de setembro de 2021.

⁶ Disponível em: <https://saude.abril.com.br/medicina/as-diferencas-e-semelhancas-entre-o-sars-cov-2-e-outros-coronavirus/>. Acesso em: 09 de setembro de 2021.

⁷ Disponível em: https://www.who.int/health-topics/severe-acute-respiratory-syndrome#tab=tab_1. Acesso em: 09 de setembro de 2021. Tradução nossa.

positivo para a doença a partir da detecção do RNA viral (material genético do vírus) ou da resposta imune do corpo humano à infecção, como anticorpos ou outros marcadores, em exames realizados (OMS, pg. 1, 2020).

4.3 PANDEMIA

Uma pandemia é uma epidemia ocorrendo em todo o mundo ou em uma área muito ampla, cruzando fronteiras internacionais, e geralmente afetando um grande número de pessoas (LAST, 2001, tradução nossa). No dia 12 de março de 2020, a OMS anunciou o surto da COVID-19 como uma pandemia, após a confirmação de mais de 20 mil casos e quase mil mortes na região europeia.⁸ Segundo dados da Universidade Johns Hopkins, mais de 192 países já foram afetados pelo coronavírus, com mais de 220 milhões de casos confirmados e 4,6 milhões de óbitos.⁹

Algumas medidas foram tomadas por líderes mundiais, na tentativa de controlar a disseminação do vírus. Desde o fechamento de fronteiras com outros países até o confinamento quase completo da população, as ações foram se intensificando com o passar dos meses para que o impacto causado pela doença fosse minimizado¹⁰.

Alguns lugares se saíram melhores do que outros nesse sentido. Um estudo publicado na Bloomberg¹¹ vem avaliando os países e os ranqueando através de pontuação dada para os seguintes critérios: casos mensais por 100 mil habitantes, taxa de mortalidade mensal, mortes totais por 1 milhão de habitantes, taxa de testes positivos, acesso a vacinas da COVID e doses ministradas por 100 habitantes. O trabalho vem sendo atualizado desde novembro de 2020 e atualmente mostra a Noruega como o país que melhor lida com a crise de saúde causada pela doença, enquanto a Malásia é o pior dos 53 países avaliados. No momento, o Brasil ocupa a 35^a posição, mas já esteve ranqueado em 51º no mês de maio de 2021.

⁸ Disponível em: <https://www.euro.who.int/en/health-topics/health-emergencies/coronavirus-COVID-19/news/news/2020/3/who-announces-COVID-19-outbreak-a-pandemic> . Acesso em 09 de setembro de 2021.

⁹ Disponível em: <https://coronavirus.jhu.edu/map.html> . Acesso em: 09 de setembro de 2021.

¹⁰ Disponível em: <https://saude.estadao.com.br/noticias/geral,o-que-os-paises-estao-fazendo-para-controlar-a-pandemia-do-coronavirus,70003235232>. Acesso em: 09 de setembro de 2021.

¹¹ Disponível em: <https://www.bloomberg.com/graphics/covid-resilience-ranking/>.

4.4 AUXÍLIO EMERGENCIAL

Além do problema com a saúde da população, outra consequência avassaladora da pandemia foi a crise econômica. No Brasil, o Governo Federal estabeleceu medidas com o objetivo de tentar de minimizar o impacto econômico causado pela pandemia. Através da lei nº 13.982, de 2 de abril de 2020, em seu Artigo 2º, foi definida a concessão de um auxílio emergencial no valor de R\$ 600,00 aos trabalhadores que atenderem aos requisitos listados. O auxílio foi, inicialmente, estabelecido para 3 meses a partir da data de publicação da lei nº 13.982, mas se estendeu até dezembro de 2020. Ao todo foram mais de R\$ 210 bilhões investidos e mais de 67 milhões de pessoas beneficiadas durante o período¹².

O auxílio também retornou em abril de 2021¹³, com um valor mais baixo do que o original (entre R\$ 150 a R\$ 375) e custando cerca de R\$ 9 bilhões por mês para os cofres públicos, sendo estendido até o mês de novembro¹⁴.

¹² Disponível em: <https://agenciabrasil.ebc.com.br/economia/noticia/2020-09/caixa-detalha-calendario-de-pagamentos-do-auxilio-emergencial-extensao>.

Acesso em: 09 de setembro de 2021.

¹³ Disponível em: <https://g1.globo.com/economia/noticia/2021/06/08/auxilio-emergencial-nova-etapa-deve-custar-r-11-bilhoes-fora-do-teto-de-gastos-diz-guedes.ghtml> . Acesso em: 09 de setembro de 2021.

¹⁴ Disponível em: <https://auxilio.caixa.gov.br/#/calendario> . Acesso em: 09 de setembro de 2021.

5. MÉTODO

5.1 BASE DE DADOS

Os dados utilizados para o estudo foram retirados da PNAD COVID-19, realizada pelo IBGE em 2020. Para realizar a pesquisa, o IBGE utilizou como base a amostra do primeiro trimestre de 2019 da PNAD Contínua. Segundo o Instituto:

A amostra original da PNAD Contínua foi obtida por um plano amostral conglomerado em dois estágios com estratificação das unidades primárias de amostragem (UPAs). No primeiro estágio foram selecionadas UPAs com probabilidade proporcional ao número de domicílios dentro de cada estrato definido. No segundo estágio foram selecionados 14 domicílios particulares permanentes (que podem estar ou não ocupados) dentro de cada UPA da amostra do primeiro estágio. O sorteio dos domicílios em cada UPA foi feito por amostragem aleatória simples, considerando os endereços listados no Cadastro Nacional de Endereços para Fins Estatísticos (CNEFE) atualizado para cada UPA. (IBGE, 2020)

Dos 211 mil domicílios presentes na amostra base, 193.622 (92%) foram contatados para a PNAD COVID-19, com 3.364 municípios distribuídos por todos os estados do país sendo contemplados.

A PNAD COVID-19 passou por ajustes ao longo dos meses em que foi aplicada, a partir maio de 2020, dessa forma a base de dados utilizada foi a mais recente divulgada, datada de novembro de 2020.

A base de microdados é dividida em 7 blocos e cada um deles está atrelado a um tipo de informação desejada. Na primeira parte, a PNAD nos fornece dados para a identificação e controle de pessoas e domicílios. O segundo bloco apresenta características gerais dos moradores e possui informações como idade, sexo e escolaridade. A terceira parte da pesquisa faz questionamentos sobre a doença, com as perguntas desse bloco sendo voltadas aos sintomas sentidos, testes realizados e providências tomadas pelos indivíduos. O quarto grupo de perguntas é voltado a informações sobre o mercado de trabalho e, como já dito, é um dos principais objetivos desta PNAD. Após esse bloco, os três seguintes buscam informações complementares sobre as condições da população e questionam sobre outras fontes de remuneração, empréstimos e condições de moradia das pessoas.

Para fins desse trabalho, a PNAD COVID-19 se mostrou a melhor disponível em termos de domínio público, pois, assim como a PNAD Contínua, apresenta indicadores socioeconômicos aliados a questões de trabalho da população. Ainda, a primeira se apresentou mais compatível para esse estudo em relação à segunda por trazer, também, informações sobre sintomas da Covid-19 apresentados ou não pelos integrantes da

amostra, que tipo de providência e cuidados foram tomados por quem obteve sintomas e que procedimentos foram feitos com aqueles que necessitaram de atendimento médico, além de dados dos testes realizados e seus resultados.

5.2 VARIÁVEIS UTILIZADAS

Para realização do trabalho, foram selecionadas variáveis de característica socioeconômica, demográfica ou de saúde, que foram distribuídas em dois grupos: variável dependente e variáveis independentes. Ainda, foi realizada uma análise descritiva dos dados, visando uma maior compreensão da amostra.

5.2.1 VARIÁVEL DEPENDENTE

A variável dependente do estudo foi denominada de “COVID” e responde se um indivíduo testou positivo para a doença ou não. Casos positivos foram sinalizados em pessoas que assinalaram “Positivo” para as variáveis “B009B”, “B009D”, “B009F”, que questionam o resultado para indivíduos que responderam “Sim” para as seguintes perguntas: “Fez o exame coletado com cotonete na boca e/ou nariz (SWAB)?”, “Fez o exame de coleta de sangue através de furo no dedo?”, “Fez o exame de coleta de sangue através da veia no braço?”.

É importante salientar que a PNAD COVID-19 apresentou questionamentos a respeito de sintomas associados à COVID-19, mas optou-se pela não utilização dessas variáveis no estudo, visto que o resultado dos testes são parâmetros mais precisos para a definição de um caso positivo da doença.

5.2.2 VARIÁVEIS INDEPENDENTES

A análise de regressão utilizou de dados binários, por isso as variáveis independentes foram agrupadas da seguinte forma:

Variável	Descrição	Categorias	Código
UF	Região de moradia do entrevistado	Nordeste	1
		Outra região	0
A003	Sexo	Mulher	1
		Homem	0
A004	Raça	Pardos	1
		Não Pardos	0
A005	Escolaridade	Ensino Básico	1
		Ensino Superior	0
Grupo de Risco	Indivíduo com comorbidade diagnosticada	Sim	1
		Não	0
Idosos	Indivíduos com 60 anos ou mais de idade	Sim	1
		Não	0
Pobres	Renda domiciliar per capita abaixo da linha de pobreza de US\$ 5,50	Sim	1
		Não	0

Para as variáveis de Sexo e Raça, foram utilizados os dados brutos presentes na PNAD COVID-19, tendo como referência o sexo feminino e os pardos, respectivamente.

Por outro lado, foi necessária alguma manipulação da base de dados para a criação das demais variáveis.

A região de moradia utilizada como base para o estudo foi o Nordeste, representando uma variável que categoriza os indivíduos que moram em algum dos nove estados ou não.

Para a categoria de Ensino Básico, na variável da Escolaridade, foram contempladas todas as categorias da PNAD COVID-19 com nível de estudo máximo de ensino superior incompleto, enquanto a categoria Ensino Superior contemplou pessoas com ensino superior completo ou acima.

No Grupo de Risco, houve um grupamento das variáveis “A002”, “B101”, “B102”, “B103”, “B104”, “B105”, “B106” da PNAD COVID-19, sendo considerado pertencente ao grupo de risco qualquer indivíduo que tenha respondido “Sim” para ao menos uma das seguintes perguntas: “Algum médico já lhe deu o diagnóstico de diabetes?”, “Algum médico já lhe deu o diagnóstico de hipertensão?”, “Algum médico já lhe deu o diagnóstico de asma/bronquite/enfisema/doenças respiratória crônica ou doença de pulmão?”, “Algum médico já lhe deu o diagnóstico de doenças do coração (infarto, angina,

insuficiência cardíaca, arritmia)?”, “Algum médico já lhe deu o diagnóstico de depressão?”, “Algum médico já lhe deu o diagnóstico de câncer?”.

Já a variável Pobres passou pelas seguintes manipulações: primeiramente, a renda domiciliar foi calculada através da soma da renda total individual de todas as pessoas vivendo no mesmo domicílio. Para a renda total individual, por sua vez, foram somadas as variáveis “C01012”, “D0013”, “D0023”, “D0033”, “D0043”, “D0053”, “D0063” e “D0073”, que representam os rendimentos referentes ao salário ou de outras fontes como aposentadorias, pensões e programas assistenciais do governo, por parte dos moradores de 14 anos ou mais. Assim, a renda domiciliar per capita é a renda domiciliar dividida pelo número de indivíduos de cada domicílio. O fator de conversão da moeda utilizado foi o PPP – consumo privado para o ano de 2020, disponibilizado pelo Banco Mundial¹⁵, no valor de 2,45, e o mês foi considerado como um período de 30 dias. Dessa forma, a classificação Pobre foi feita para os residentes em domicílios com renda domiciliar per capita abaixo da linha de US\$ 5,50 por dia, estabelecida pelo Banco Mundial (R\$ 404,25 ao mês)¹⁶.

Por fim, foram considerados idosos aqueles que possuem 60 anos ou mais.

5.2.3 VARIÁVEIS DESCRITIVAS

Para a elaboração da análise descritiva da amostra, optou-se pelo grupamento diferente de algumas variáveis, para uma análise mais ampla.

No caso das Unidades da Federação, utilizou-se a região à qual pertence o estado, fazendo desta uma análise regional entre as partes Norte, Nordeste, Centro-Oeste, Sudeste e Sul do país.

Outra variável que foi objeto de ajuste para uma melhor compreensão foi a da idade dos indivíduos, sendo agrupada em 3 faixas etárias: até 19 anos, de 20 a 59 anos e 60 anos ou mais.

A variável referente à renda, nessa parte da análise, foi dividida em Classe Social, utilizando de critério já apresentado pelo IBGE¹⁷ em um de seus estudos. Foi considerado o total de renda recebida pelos integrantes do domicílio, incluindo benefícios e valores previdenciários pagos pelo Governo. Assim, os grupos de renda utilizados para a divisão

¹⁵ Disponível em: <https://data.worldbank.org/indicator/PA.NUS.PRVT.PP?end=2020&start=2011>. Acesso em: 03 de janeiro de 2022.

¹⁶ Renda mensal de referência = $5,50 \times 30 \times 2,45 = 404,25$

¹⁷ Disponível em: <https://www.ibge.gov.br/estatisticas/multidominio/condicoes-de-vida-desigualdade-e-pobreza/17374-indicadores-sociais-minimos.html?=&t=resultados>. Acesso em: 09 de setembro de 2021.

foram, em classes de salário-mínimo: até 2 (Classe E), mais de 2 a 5 (Classe D), mais de 5 a 10 (Classe C), mais de 10 a 20 (Classe B) e mais de 20 (Classe A). O salário-mínimo de referência utilizado para o estudo foi de R\$ 1.045,00, estabelecido na Medida Provisória nº 919, de 30 de janeiro de 2020.

Para a Raça e a Escolaridade dos indivíduos, foram mantidas as classificações da PNAD COVID-19, enquanto o Grupo de Risco continuou como na divisão anterior.

5.3 MODELO ESTATÍSTICO

A análise estatística elaborada pelo trabalho se deu na forma de uma regressão logística, que, ao invés de modelar uma resposta direta e quantitativa para a variável dependente, modela a probabilidade desta variável pertencer a determinada categoria (JAMES, 2013).

Para esse método de regressão utiliza-se a chamada função logística, que possui a notação abaixo:

$$p(X) = \frac{e^{\beta_0 + \beta_1 X}}{1 + e^{\beta_0 + \beta_1 X}} \quad (1) \quad ,$$

com “ $p(X)$ ” representando a probabilidade ligada a “ X ”, “ e ” sendo a constante de Euler (que vale aproximadamente 2,718) e “ $\beta_0 + \beta_1$ ” como os coeficientes ou parâmetros. A equação acima pode ser rearranjada para encontrar

$$\frac{p(X)}{1 - p(X)} = e^{\beta_0 + \beta_1 X} \quad (2) \quad ,$$

onde o termo “ $p(X) / 1 - p(X)$ ” é conhecido como “chance”.

Outra forma de observar o resultado da regressão logística é através do termo denominado de Logit, conhecido como log das chances,

$$\log \left(\frac{p(X)}{1 - p(X)} \right) = \beta_0 + \beta_1 X \quad (3) \quad ,$$

que pode ser encontrado ao tirar o logaritmo dos dois lados de (2) (JAMES, 2013).

Hosmer (2013) afirma que é importante salientar que o objetivo do uso de um modelo de regressão logística é o mesmo de qualquer outro: encontrar o melhor modelo para descrever a relação entre uma variável dependente (resposta) e um conjunto de variáveis independentes (explicativas). Ainda segundo o autor, o que diferencia o modelo de regressão logística para um de regressão linear é que, na primeira, a variável dependente é binária (dicotômica). Assim:

Essa diferença entre regressão logística e linear é refletida tanto na forma do modelo quanto nos seus pressupostos. Uma vez que essa diferença é contabilizada, os métodos empregados em uma análise por regressão logística seguem, mais ou menos, os mesmos princípios gerais usados na regressão linear.

5.4 SOFTWARES

Todos os cálculos e ajustes necessários durante o estudo foram realizados através do *software* Rstudio, pela linguagem R (v. 4.1.2), por meio de suas bibliotecas dplyr (v. 1.0.7), srvyr (v. 1.1.0), readr (v. 2.1.1), survey (v. 4.1-1) e jtools (v. 2.1.4).

6. IMPLEMENTAÇÃO

A implementação do trabalho foi realizada em dois passos: primeiramente uma análise descritiva dos dados, seguida por uma análise de regressão logística.

A etapa inicial, contendo os primeiros cálculos e manipulações realizados com os dados, proporcionou um maior entendimento sobre a população, resumindo a informação e apresentando medidas sumárias.

Realizada na segunda etapa, a regressão logística, segundo Kleinbaum (2002), é um procedimento popularmente utilizado na análise de dados epidemiológicos, ainda mais para problemas multivariáveis, como o desse trabalho. O autor cita que um questionamento típico feito pelos pesquisadores para esse tipo de problema é: “qual é a relação de uma ou mais variáveis do estudo com o resultado da doença?”, que pode ser respondido por esse tipo de modelagem matemática.

6.1 ANÁLISE DESCRITIVA

A base de microdados utilizada para o trabalho, a PNAD COVID-19, possui 381.438 observações. Para uma melhor compreensão das informações levantadas pela pesquisa, torna-se importante um estudo descritivo da população, a respeito de características demográficas e socioeconômicas dos dados.

A Tabela 1 abaixo mostra como estão distribuídos todos os indivíduos presentes na amostra, com seus pesos, a partir das frequências absolutas e relativas para cada característica.

Tabela 1 – Frequência demográfica e socioeconômica da população brasileira - PNAD COVID-19

Região do Brasil	Frequência - Absoluta	Frequência - Relativa
Norte	18.425.105	8,71%
Nordeste	57.337.173	27,09%
Centro-Oeste	16.451.989	7,77%
Sudeste	89.214.762	42,15%
Sul	30.223.340	14,28%
Gênero	Frequência - Absoluta	Frequência - Relativa
Mulher	108.188.952	51,12%
Homem	103.463.417	48,88%
Raça	Frequência - Absoluta	Frequência - Relativa
Parda	97.363.765	46,00%
Branca	93.550.048	44,20%
Preta	18.503.853	8,74%
Amarela	1.544.538	0,73%
Indígena	624.391	0,30%
Ignorado	65.775	0,03%
Escolaridade	Frequência - Absoluta	Frequência - Relativa
Sem instrução	23.009.122	10,87%
Fundamental incompleto	64.442.686	30,45%
Fundamental completo	13.472.590	6,37%
Médio incompleto	20.149.046	9,52%
Médio completo	50.267.381	23,75%
Superior incompleto	12.515.405	5,91%
Superior completo	21.878.257	10,34%
Pós-graduação, mestrado ou doutorado	5.917.881	2,80%
Faixa Etária	Frequência - Absoluta	Frequência - Relativa
Até 19 anos	59.597.149	28,16%
20 a 59 anos	121.507.211	57,41%
60 anos ou mais	30.548.009	14,43%
Grupo de Risco	Frequência - Absoluta	Frequência - Relativa
Não	163.932.839	77,45%
Sim	47.719.530	22,55%
Classe Social	Frequência - Absoluta	Frequência - Relativa
A	3.273.242	1,55%
B	10.260.983	4,85%
C	31.455.288	14,86%
D	92.643.914	43,77%
E	74.018.943	34,97%

A partir dos dados acima, podemos traçar o perfil geral da população observada na base do IBGE. As maiores frequências foram de indivíduos do Sudeste (42,15%), do sexo Feminino (51,12%), raça Parda (46,00%), com Ensino Fundamental Incompleto (30,45%), da Faixa Etária entre 20 a 59 anos (57,41%), não pertencente ao Grupo de Risco (77,45%) e da Classe Social D (43,77%).

Podemos, ainda, analisar mais detalhadamente a Idade e a Renda Domiciliar dos indivíduos presentes na amostra.

A Tabela 2 traz os dados analisando mais a fundo a primeira variável:

Tabela 2 – Distribuição da população brasileira por Idade

Idade (anos)	
Média	34,7
Desvio Padrão	21,26
Mínimo	0
25%	17
50%	33
75%	50
Máximo	111

A média de idade dos indivíduos presentes na PNAD COVID-19 foi de aproximadamente 35 anos, o indivíduo mais novo constando na base ainda não havia completado um ano de idade, enquanto o mais velho possuía 111 anos e o Desvio Padrão foi de pouco mais de 21 anos. Detalhando melhor a distribuição das idades, 25% das pessoas possuíam 17 anos ou menos, a mediana foi de 33 anos, enquanto outros 25% estão com 50 anos ou mais (quartil de 75%).

Já a Renda Domiciliar, que é dada pela soma dos ganhos totais de todos os indivíduos vivendo no mesmo domicílio, como explicado anteriormente, teve os indicadores da sua distribuição apresentados na Tabela 3:

Tabela 3 – Distribuição da população brasileira por nível de Renda Domiciliar

Renda Domiciliar (R\$/mês)	
Média	4.165,67
Desvio Padrão	4.917,23
Mínimo	0,00
25%	1.680,00
50%	2.845,00
75%	4.798,00
Máximo	190.000,00

A variável apresentou uma média de pouco mais de quatro mil reais (R\$ 4.165,67) e uma grande discrepância entre seus valores mínimo e máximo: R\$ 0,00 foi o menor valor apresentando, enquanto o maior foi de R\$ 190.000,00. Essa diferença resultou em um Desvio Padrão de R\$ 4.917,23. O valor recebido pelos 25% de Renda Domiciliar mais baixa da amostra foi de R\$ 1.680,00 ou menos, enquanto os 25% de maior renda vivem em domicílio que recebe R\$ 4.798,00 ou mais. A mediana foi de R\$ 2.845,00.

Em documento divulgado em 2020, o Banco Mundial mostra uma análise sobre a quantidade de pessoas vivendo em 3 condições que a instituição denomina de linhas de pobreza: rendas per capita menores que US\$1,90, US\$ 3,20 e US\$ 5,50 por dia (World Bank, 2020). Utilizando-se dos dados presentes na PNAD COVID-19, a realização de análise semelhante resultou nos dados apresentados na Tabela 4 abaixo.

Tabela 4 – Percentual da população brasileira vivendo abaixo das linhas de pobreza

Classificação	Frequência Relativa
Acima da Linha de Pobreza	81,56%
Menos de US\$ 5,50 / dia (R\$ 404,25 / mês)	18,44%
Menos de US\$ 3,20 / dia (R\$ 235,20 / mês)	7,97%
Menos de US\$ 1,90 / dia (R\$ 139,65 / mês)	3,28%

Para chegar ao resultado, a variável utilizada foi a Renda Domiciliar Per Capita dos indivíduos, que é a Renda Domiciliar calculada anteriormente, dividida pela quantidade de pessoas morando no domicílio, critério também já adotado pelo IBGE¹⁸. Além disso, o fator de conversão da moeda foi o PPP – consumo privado para o ano de 2020, como já dito anteriormente.

A análise mostra mais de 18% da população abaixo da linha de US\$ 5,50 por dia, quase 8% vivendo com menos de US\$ 3,20 e 3,28% na classe denominada de extrema pobreza.

Outra forma de análise que se mostra importante para o estudo é a divisão das observações em outros dois grupos: casos negativos e casos positivos de COVID-19. As tabelas abaixo mostram como ficou distribuída a amostra da PNAD COVID-19 a partir da divisão entre os diferentes resultados para os testes da doença.

¹⁸ Disponível em: <https://agenciadenoticias.ibge.gov.br/agencia-sala-de-imprensa/2013-agencia-de-noticias/releases/26956-ibge-divulga-o-rendimento-domiciliar-per-capita-2019>.

Acesso em: 09 de setembro de 2021.

Analisando primeiramente os casos por região do país, alguns pontos se mostram importantes de destacar.

Tabela 5 – Frequências de Casos Negativos / Não Testados e Positivos de Covid-19 por Região do Brasil

Variável e Categorias	Frequência - Casos negativos e não testados		Frequência - Casos Positivos	
	Absoluta	Relativa	Absoluta	Relativa
Região do Brasil				
Norte	17.649.199	8,60%	775.906	11,94%
Nordeste	55.536.899	27,07%	1.800.274	27,71%
Centro-Oeste	15.629.803	7,62%	822.186	12,65%
Sudeste	86.871.704	42,34%	2.343.058	36,06%
Sul	29.467.211	14,36%	756.129	11,64%

Como observado na Tabela 5, a participação da região Sudeste tem uma diminuição considerável na frequência relativa de casos negativos para casos positivos, bem como da região Sul. Isso está atrelado ao aumento das participações das regiões Norte e Centro-Oeste no grupo de pessoas que testaram positivo para a doença. A região Nordeste manteve a participação relativamente estável.

Na Tabela 1, na variável de Escolaridade, foi observado que a categoria de maior frequência foi a de Fundamental Incompleto. Analisando sob uma nova ótica:

Tabela 6 – Frequências de Casos Negativos / Não Testados e Positivos de Covid-19 por Escolaridade

Variável e Categorias	Frequência - Casos negativos e não testados		Frequência - Casos Positivos	
	Absoluta	Relativa	Absoluta	Relativa
Sem instrução	22.694.781	11,06%	314.341	4,84%
Fundamental incompleto	63.351.277	30,88%	1.091.409	16,80%
Fundamental completo	13.124.123	6,40%	348.467	5,36%
Médio incompleto	19.646.039	9,58%	503.007	7,74%
Médio completo	48.306.632	23,55%	1.960.749	30,18%
Superior incompleto	11.923.565	5,81%	591.841	9,11%
Superior completo	20.606.316	10,04%	1.271.941	19,58%
Pós-graduação, mestrado ou doutorado	5.502.082	2,68%	415.798	6,40%

Na Tabela 6, essa característica se manteve para os indivíduos que não tiveram teste diagnosticando como um caso positivo de COVID-19, chamados neste trabalho de Casos Negativos e Não Testados, com uma frequência de 30,88%. Porém, ao olhar-se para os casos positivos, a categoria da Escolaridade com maior aparição foi a de Ensino Médio Completo (30,18%), o que mostra uma mudança no perfil geral da população através dessa divisão.

Já na variável relacionada ao Gênero dos indivíduos, observamos:

Tabela 7 – Frequências de Casos Negativos / Não Testados e Positivos de Covid-19 por Gênero

Variável e Categorias	Frequência - Casos negativos e não testados		Frequência - Casos Positivos	
	Absoluta	Relativa	Absoluta	Relativa
Mulher	104.689.234	51,03%	3.499.718	53,86%
Homem	100.465.583	48,97%	2.997.834	46,14%

A frequência das mulheres é maior para os dois casos, mas houve um aumento, mesmo que modesto, na sua participação nos casos positivos.

A Raça foi a variável que obteve frequências mais semelhantes ao dividir a amostra em grupos de casos negativos e positivos, mas ainda assim apresentou um ponto que deve ser destacado.

Tabela 8 – Frequências de Casos Negativos / Não Testados e Positivos de Covid-19 por Raça

Variável e Categorias	Frequência - Casos negativos e não testados		Frequência - Casos Positivos	
	Absoluta	Relativa	Absoluta	Relativa
Parda	94.383.967	46,01%	2.979.798	45,86%
Branca	90.646.822	44,18%	2.903.226	44,68%
Preta	17.948.856	8,75%	554.997	8,54%
Amarela	1.512.965	0,74%	31.573	0,49%
Indígena	598.792	0,29%	25.598	0,39%
Ignorado	63.414	0,03%	2.361	0,04%

Apesar de pequena, a frequência relativa dos indígenas aumentou em quase 35% no grupo dos casos positivos em relação aos casos negativos.

Outra variável que demonstrou comportamento interessante ao dividir-se a amostra em Casos Negativos e Casos Positivos foi a da Faixa Etária dos indivíduos.

Tabela 9 – Frequências de Casos Negativos / Não Testados e Positivos de Covid-19 por Faixa Etária

Variável e Categorias	Frequência - Casos negativos e não testados		Frequência - Casos Positivos	
	Absoluta	Relativa	Absoluta	Relativa
Até 19 anos	58.845.006	28,68%	752.143	11,58%
20 a 59 anos	116.513.523	56,79%	4.993.688	76,85%
60 anos ou mais	29.796.287	14,52%	751.722	11,57%

Mesmo a categoria predominante sendo igual nos dois grupos (20 a 59 anos), há um grande aumento da frequência relativa dessa categoria, de 56,79% para 76,85%, acompanhado de uma queda substancial no indicador da faixa etária até 19 anos (de 28,68% para 11,58%).

Os integrantes do Grupo de Risco também aumentaram sua participação no bloco de pessoas que testaram positivo para o coronavírus, em relação aos que não obtiveram resultado positivo para o teste.

Tabela 10 – Frequências de Casos Negativos / Não Testados e Positivos de Covid-19 por Grupo de Risco

Variável e Categorias	Frequência - Casos negativos e não testados		Frequência - Casos Positivos	
	Absoluta	Relativa	Absoluta	Relativa
Não	159.295.332	77,65%	4.637.507	71,37%
Sim	45.859.484	22,35%	1.860.046	28,63%

Apesar dos não pertencentes ao grupo ainda continuarem sendo maioria nos dois blocos, a diferença que antes era de 22,35% - 77,65%, alterou-se para 28,63% - 71,37%.

Por fim, a análise na variável da Classe Social:

Tabela 11 – Frequências de Casos Negativos / Não Testados e Positivos de Covid-19 por Classe Social

Variável e Categorias	Frequência - Casos negativos e não testados		Frequência - Casos Positivos	
	Absoluta	Relativa	Absoluta	Relativa
A	3.042.495	1,48%	230.746	3,55%
B	9.670.447	4,71%	590.536	9,09%
C	30.080.050	14,66%	1.375.238	21,17%
D	89.875.952	43,81%	2.767.962	42,60%
E	72.485.872	35,33%	1.533.071	23,59%

O destaque fica por parte da diferença da frequência da Classe E no primeiro e segundo bloco, havendo uma queda de aproximadamente 12% em sua participação, pautada por um aumento na frequência das Classes A, B e C nos casos positivos. Nesse quadro, a Classe D continuou sendo a mais presente nos casos negativo e positivos.

Assim, o perfil das pessoas que apresentaram casos positivos da COVID-19, traçando um paralelo com o geral da população apresentado, foi de indivíduos do Sudeste (36,06%), do sexo feminino (53,86%), raça Parda (45,86%), com Ensino Médio Completo (30,18%), da Faixa Etária entre 20 a 59 anos (76,85%), não pertencente ao Grupo de Risco (71,37%) e da Classe Social D (42,60%).

6.2 ANÁLISE DE REGRESSÃO LOGÍSTICA

Após uma melhor compreensão de como está distribuída a população de acordo com as variáveis determinadas acima, esse segundo momento de análise objetivou mensurar a relação entre algumas características demográficas e socioeconômicas e as chances de contaminação pelo novo coronavírus. Essa relação será apresentada em termos

da Razão das Chances – do inglês Odds Ratio (OR) – e do Logit, presentes no resultado da regressão logística.

As chances (Odds) de um evento são dadas pela razão entre a probabilidade desse evento ocorrer e a probabilidade de não ocorrer. No caso de um evento com 75% de probabilidade de ocorrer, por exemplo, as chances desse evento são iguais a 3¹⁹. Assim, a Razão das Chances (OR) compara as chances de um evento acontecer quando determinado critério é atendido e as chances de acontecer quando não é atendido. Para esse trabalho, as ORs representam as relações entre as chances de um indivíduo contrair a COVID-19 quando possui as características selecionadas e quando não possui.

Porém, ao comparar, por exemplo, dois eventos com chances iguais a 1 e 4, respectivamente, os resultados da OR podem ser dados por 0,25 ou 4, a depender de como o cálculo seja arranjado, o que gera um problema de simetria. Assim, uma forma de acrescentar essa simetria ao estudo é utilizando o log da Razão das Chances, que é chamado de Logit. Para o exemplo anterior, utilizando o Logit, os resultados encontrados seriam $\text{Log}(0,25) = -0,602$ ou $\text{Log}(4) = 0,602$, que apresentam a simetria desejada.

A tabela abaixo apresenta os resultados da análise de regressão:

Tabela 12 – Resultados da Regressão Logística

Variável	Logit	[0.025	0.975]	P	Razão das Chances	[0.025	0.975]
Nordeste	0,1560	0,0797	0,2324	0,00	1,1689	1,0829	1.2616
Mulheres	0.0643	0.0214	0.1072	0,00	1.0664	1.0217	1.1132
Pardos	0.1100	0.0487	0.1714	0,00	1.1163	1.0499	1.1869
Ensino Básico	-0.8101	-0.8752	-0.7450	0,00	0.4448	0.4168	0.4747
Grupo de Risco	0.4250	0.3620	0.4881	0,00	1.5297	1.4363	1.6291
Idosos	-0.5092	-0.5890	-0.4294	0,00	0.6010	0.5549	0.6509
Pobres	-0.5716	-0.6856	-0.4576	0,00	0.5646	0.5038	0.6328

As variáveis consideradas no estudo que apresentaram relação positiva com a contaminação pela COVID-19 foram as do gênero Feminino (OR: 1,0664; IC 95%: 1,0217 – 1,1132), dos Pardos (OR: 1,1163; IC 95%: 1,0499 – 1,1869), dos residentes do Nordeste (OR: 1,1689; IC 95%: 1,0829 – 1,2616) e dos pertencentes ao Grupo de Risco (OR: 1,5297; IC 95%: 1,4363 – 1,6291). Já as negativamente relacionadas foram as pertencentes às categorias dos Pobres (OR: 0,5646; IC 95%: 0,5038 – 0,6328), dos sem

¹⁹ Odds = 75% / 25% = (3/4) / (1/4) = 3.

Ensino Superior (OR: 0,4448; IC 95%: 0,4168 – 0,4747) e dos Idosos (OR: 0,6010; IC 95%: 0,5549 – 0,6509).

O modelo de regressão logística apresentado obteve $\text{Pseudo-R}^2 = 0,02$. O Pseudo-R^2 para a regressão logística é uma medida análoga ao R^2 utilizado em problemas de regressão linear, servindo como índice de adequação e capacidade preditiva do modelo (SMITH, 2013). O indicador varia de 0 a 1, com valores mais próximos de 1 indicando modelos mais explicativos quanto à variável dependente.

7. DISCUSSÃO DOS RESULTADOS

As variáveis selecionadas para o estudo carregam informações socioeconômicas amplamente debatidas ao longo do tempo e que, por isso, permitem que sejam feitas algumas projeções quanto às suas respectivas relações com a contaminação pela COVID-19. Os resultados da regressão logística permitem a realização de uma análise a respeito de quais dessas variáveis seguiram aquilo que foi imaginado, quais não seguiram e as razões para tal.

Através dos dados da PNAD COVID-19, foi possível observar, neste trabalho, uma menor propensão à contaminação pelo Sars-Cov-2 por parte da população pobre (OR: 0,5646; IC 95%: 0,5038 – 0,6328). O resultado vai de encontro ao esperado de que essa parcela da população está mais suscetível a ser contaminada, devido às condições socioeconômicas precárias vividas pela população de renda mais baixa, com problemas de saneamento e muitas vezes superlotação de moradias.

Em livro divulgado no mês de abril de 2021, a Fundação Oswaldo Cruz (Fiocruz) analisou, dentre outras coisas, as taxas de contaminação e letalidade para regiões com baixa e alta concentração de favelas. No trabalho, o órgão conta:

Nossa expectativa inicial (hipótese) era de que, com a evolução da pandemia, a ocorrência da doença fosse maior em áreas com grande concentração de favelas. Entretanto, os resultados indicam que tanto casos como óbitos apresentaram ocorrência maior em áreas sem favelas e com baixa concentração de favelas (MATTA, 2021).

A Fundação caracteriza as favelas como regiões de alta densidade populacional e oferta insuficiente de serviços públicos, entre outros, dando razão à sua hipótese de que seus habitantes sofreriam mais com a pandemia. Porém, a Fiocruz credita o resultado encontrado à baixa testagem por parte da população mais pobre, devido à menor possibilidade dessa de pagar pelos testes diagnósticos (MATTA, 2021).

O problema sugerido pelos números ligados à renda dos indivíduos é reforçado pelo resultado encontrado em relação à variável ligada à escolaridade desses. Nesse estudo, os indivíduos que não possuem Ensino Superior apresentaram menores chances de contaminação pelo Sars-Cov-2, em comparação com aqueles que possuem Ensino Superior ou mais (OR: 0,4448; IC 95%: 0,4168 – 0,4747). A conhecida correlação dessas duas variáveis (Renda e Escolaridade)²⁰ é a base para o reforço da tese que a falta de

²⁰ Disponível em: https://www.ipea.gov.br/portal/index.php?option=com_content&view=article&id=20067.

informações esteja impactando os valores e afastando a conclusão da hipótese que a população menos abastada esteja sofrendo mais com a pandemia em termos de contágio.

Com relação ao gênero dos indivíduos, o estudo apontou a existência de uma correlação entre os casos positivos e as mulheres (OR: 1,0664; IC 95%: 1,0217 – 1,1132), indicando que o gênero feminino está associado a uma maior probabilidade de contaminação pela COVID-19. As estatísticas descritivas elaboradas com os dados da PNAD COVID-19 mostraram um aumento da participação de indivíduos do sexo feminino para os casos positivos, se comparados aos demais.

Uma situação que pode explicar essa maior probabilidade de contaminação pelas mulheres está na caracterização da população dos profissionais pertencentes à chamada “linha de frente” do combate à pandemia, aqueles com maior exposição ao vírus por trabalharem na área de saúde ou em serviços essenciais. Um resumo técnico elaborado pelo Fundo das Nações Unidas para Atividades Populacionais (UNFPA) mostrou que, ao redor do mundo, 70% da força de trabalho dessa chamada “linha de frente” é composta por mulheres. A nível de Brasil, a Associação Brasileira de Ciência Política (ABCP) mostrou:

Dos trabalhadores de nível superior nessa área elas são 90,39% entre as enfermeiras; 95,31% entre as nutricionistas, e, no entanto, entre os médicos, elas representam apenas 35,94%. Já, entre os profissionais de nível médio elas são 77,88% das técnicas em fisioterapia e afins, 78,03% das atendentes de enfermagem, parteiras e afins, e 86,93% das técnicas e auxiliares de enfermagem. A situação do Serviço Social é praticamente idêntica. (ABCP, 2020).

Outra característica positivamente associada a uma maior probabilidade de contaminação pela COVID-19 foi a dos indivíduos Pardos, com 11,63% a mais de chances de adquirir a doença (OR: 1,1163; IC 95%: 1,0499 – 1,1869). Esse é um resultado que pode ser interpretado como controverso, dados os outros já apresentados para a população mais pobre e com baixa escolaridade, onde a população de cor parda está majoritariamente inserida, que mostraram uma correlação negativa com a contaminação pela doença. Porém, uma análise mais segmentada pode trazer uma explicação para esse caso. Apesar de enxergar um problema de baixa testagem para a população mais pobre, ao comparar indivíduos vivendo em situações econômicas semelhantes, a Fiocruz apontou para uma taxa de contaminação dos pretos e pardos duas vezes maior em relação à população branca nos bairros “sem favela”, onde a população possui maior condição de realizar testes. (MATTA, 2020)

Neste trabalho, residentes do Nordeste também apresentaram maior probabilidade de contaminação pela COVID-19 (OR: 1,1689; IC 95%: 1,0829 – 1,2616). Essa maior propensão a contrair a doença por parte dos indivíduos que residem nessa região traz uma nova ótica e corrobora com o que foi pensado anteriormente para as pessoas vivendo em situações mais precárias, visto que essa é uma das regiões mais pobres do Brasil.

Segundo Kerr (2020):

A epidemia, apesar de não ter poupado as regiões mais ricas, vem ocorrendo de forma mais acentuada nos estados das regiões mais pobres, como o Norte e Nordeste. O Nordeste, uma das regiões mais pobres do país, representa 27% da população brasileira e apresenta cerca de um terço de todos os casos (34%) e dos óbitos (32%). (KERR, 2020).

A pesquisa do IBGE também permitiu apontar que pessoas do Grupo de Risco estão mais suscetíveis à contaminação pelo Sars-Cov-2 (OR: 1,5297; IC 95%: 1,4363 – 1,6291). Essa parcela da população tem sido razão de uma das maiores preocupações por parte dos especialistas e esse trabalho mostrou que a preocupação é extremamente válida, visto que essa foi a categoria mais fortemente relacionada com uma maior probabilidade de contrair a COVID-19, com 53% a mais de chances. Essa relação positiva entre pessoas com morbidades se mostrou verdadeira também em pandemias anteriores de outros coronavírus, tanto do tipo Mers-Cov quanto do Sars-Cov, em outros lugares do mundo, como mostraram Badawi et al (2016) e Yu et al (2006), respectivamente.

É importante ressaltar novamente que a classificação de Grupo de Risco para este estudo não utilizou a idade acima de 60 anos como critério, visto que os Idosos formaram uma categoria analisada separadamente e que apresentou correlação negativa com as chances de contaminação pela COVID-19 (OR: 0,6010; IC 95%: 0,5549 – 0,6509), resultado que também vai de encontro ao esperado, devido à maior fragilidade do grupo. Buscando uma explicação para a divergência entre a expectativa e a realidade, uma razão apontada é que, segundo pesquisas, os jovens estão menos isolados que os idosos e, por isso, estão mais vulneráveis à contaminação (VASCONCELOS, 2020).

8. CONSIDERAÇÕES FINAIS

O presente trabalho objetivou acrescentar ao debate a respeito da pandemia no Brasil. O estudo permitiu a percepção da existência de relações entre variáveis socioeconômicas e a contaminação pelo coronavírus, bem como a compreensão da magnitude dessas.

De maneira geral, os estudos econômicos relacionados à COVID-19 acabam levantando alguns questionamentos à medida que buscam solucionar outros, bem como apontam para direções não esperadas anteriormente e que sugerem problemas muitas vezes não relacionados diretamente com a doença, mas que foram colocados em evidência por conta da pandemia.

Este estudo, então, trouxe alguns resultados condizentes com o esperado, outros que fugiram do imaginado e apresentou explicações para a convergência ou não entre a expectativa e a realidade. Além disso, as estatísticas descritivas detalharam bem as informações presentes na amostra, permitindo que fosse traçado um perfil geral da população do estudo com os dados presentes na PNAD COVID-19.

A partir da discussão apresentada, o trabalho pode auxiliar nas decisões a serem tomadas por parte dos responsáveis, visando proteger as parcelas mais vulneráveis da população, bem como mostrar números que indicam relativo sucesso nas medidas já aplicadas, como o isolamento ou distanciamento social. Em tempos vacinação, esses indicadores se tornam valiosos.

9. REFERÊNCIAS

Associação Brasileira de Ciência Política. **Boletim Nº 11 - Pandemia COVID-19 e as mulheres**. 2020. Disponível em: <

<https://cienciapolitica.org.br/web/noticias/2020/04/boletim-11-pandemia-covid-19-e-mulheres> >. Acesso em: 09 de setembro de 2021.

Badawi A, Ryoo SG. **Prevalence of comorbidities in the Middle East respiratory syndrome coronavirus (MERS-CoV): a systematic review and meta-analysis**. Int J Infect Dis. 2016;49:129-133. doi:10.1016/j.ijid.2016.06.015

BAQUI, Pedro et al. **Ethnic and regional variation in hospital mortality from COVID-19 in Brazil**. 2020. Disponível em: <https://www.medrxiv.org/content/10.1101/2020.05.19.20107094v1>. Acesso em 09 de setembro de 2021.

BREWER, M.; TASSEV, Iva V. Did the UK policy response to COVID-19 protect household incomes?. **EUROMOD Working Paper Series**, Essex, EM12/20, jun. 2020. Disponível em: < <https://www.euromod.ac.uk/sites/default/files/working-papers/em12-20.pdf> >. Acesso em: 09 de setembro de 2021.

BRONKA, Patryk et al. The COVID-19 Crisis Response Helps the Poor: the Distributional and Budgetary Consequences of the UK lock-down. **EUROMOD Working Paper Series**, Essex, EM11/20, jun. 2020. Disponível em: < <https://www.euromod.ac.uk/sites/default/files/working-papers/em11-20.pdf> >. Acesso em: 09 de setembro de 2021.

BROWN, Caitlin S.;RAVALLION, Martin. Inequality and the Coronavirus: Socioeconomic Covariates of Behavioral Responses and Viral Outcomes Across US Counties. **NBER Working Paper Series**, Cambridge, Working Paper 27549, jul.2020. Disponível em: < <https://www.nber.org/papers/w27549> >. Acesso em: 9 de setembro de 2021.

CAVALCANTE, João Roberto et al. **COVID-19 no Brasil: evolução da epidemia até a semana epidemiológica 20 de 2020.** 2020. Epidemiol. Serv. Saúde vol.29 no.4 Brasília set. 2020 Epub 05-Ago-2020. Disponível em: < http://scielo.iec.gov.br/scielo.php?script=sci_arttext&pid=S1679-49742020000400016 >. Acesso em: 09 de setembro de 2021.

ESTRELA, Fernanda Matheus et al. **Pandemia da COVID 19: refletindo as vulnerabilidades a luz do gênero, raça e classe.** Ciênc. saúde coletiva 25 (9) 28 Ago 2020 Set 2020 < <https://doi.org/10.1590/1413-81232020259.14052020> >. Acesso em: 09 de setembro de 2021.

HOSMER, David W. et al. **Applied Logistic Regression.** Nova Jersey: 2013.

Instituto Brasileiro de Geografia e Estatística. **PNAD COVID19 : plano amostral e ponderação.** Rio de Janeiro, 2020. Disponível em: < <https://biblioteca.ibge.gov.br/index.php/biblioteca-catalogo?view=detalhes&id=2101726> > Acesso em: 11 de setembro de 2021.

JAMES, Gareth et al. **An Introduction to Statistical Learning.** Nova York: 2013.

KAWAOKA KOMATSU, B.;MENEZES-FILHO,N. **Simulações de Impactos da COVID-19 e da Renda Básica Emergencial sobre o Desemprego, Renda, Pobreza e Desigualdade.** INSPER Policy Papers, São Paulo, n. 43, v. 14, abr. 2020. Disponível em: < <https://www.insper.edu.br/wp-content/uploads/2020/04/Policy-Paper-v14.pdf> >. Acesso em: 09 de setembro de 2021.

KERR, Lígia et al. **COVID-19 no Nordeste brasileiro: sucessos e limitações nas respostas dos governos dos estados.** 2020, vol.25, suppl.2, pp.4099-4120. Epub Sep 30, 2020. ISSN 1678-4561. < <https://doi.org/10.1590/1413-812320202510.2.28642020> >. Acesso em: 09 de setembro de 2021.

KLEINBAUM, D. G.; KLEIN, M. **Logistic Regression - A Self-Learning Text.** Atlanta: 2002.

LAST, J. M. **A Dictionary of Epidemiology**. 4 ed. Nova York, 2001.

MATTA, G.C., REGO, S., SOUTO, E.P., and SEGATA, J., eds. Os impactos sociais da COVID-19 no Brasil: populações vulnerabilizadas e respostas à pandemia [online]. Rio de Janeiro: Observatório COVID 19; Editora FIOCRUZ, 2021, 221 p. Informação para ação na COVID-19 series. ISBN: 978-65-5708-032-0. Disponível em: < <http://books.scielo.org/id/r3hc2> >. Acesso em: 09 de setembro de 2021.

MCNEELY, Connie L. et al. **Social Determinants and COVID-19 Disparities: Differential Pandemic Effects and Dynamics**. 2020. World Medical & Health Policy. Disponível em: < <https://onlinelibrary.wiley.com/doi/full/10.1002/wmh3.370> >. Acesso em: 09 de setembro de 2021.

NASCIMENTO, Israel Júnior Borges do et al. **Clinical characteristics and outcomes among Brazilian patients with severe acute respiratory syndrome coronavirus 2 infection: an observational retrospective study**. *Sao Paulo Med. J.* [online]. 2020, vol.138, n.6, pp.490-497. Epub Nov 27, 2020. ISSN 1806-9460. < <http://dx.doi.org/10.1590/1516-3180.2020.00365.r1.08092020> >. Acesso em: 09 de setembro de 2021.

NUNES, Bruno Pereira et al. **Multimorbidade e população em risco para COVID-19 grave no Estudo Longitudinal da Saúde dos Idosos Brasileiros**. *Cad. Saúde Pública* 36 (12) 20 Nov 20202020. < <https://www.scielo.org/article/csp/2020.v36n12/e00129620/pt/> >. Acesso em: 09 de setembro de 2021.

OLIVEIRA, R. G. et al. **Ensaio: Desigualdades raciais e a morte como horizonte - considerações sobre a COVID-19 e o racismo estrutural**. Rio de Janeiro, 2020. Disponível em: < <https://portal.fiocruz.br/documento/ensaio-desigualdades-raciais-e-morte-como-horizonte-consideracoes-sobre-COVID-19-e-o> >. Acesso em: 09 de setembro de 2021.

SANTOS, Edy Lawson Silva; FRANÇA, Jairo Fernando Taufick. **A Cor da Pandemia – Um Estudo Sobre a Mortalidade por COVID-19 Entre Brancos e Negros no Brasil.** 2020. Disponível em:

< https://admpg.com.br/2020/anais/arquivos/08302020_180821_5f4c1f59d419c.pdf >.

Acesso em: 09 de setembro de 2021.

SMITH, T. J.; MCKENNA, C.M. **A Comparison of Logistic Regression Pseudo R² Indices.** Multiple Linear Regression Viewpoints, 2013, Vol. 39(2).

United Nations Population Fund. **COVID-19: Um Olhar para Gênero. Proteção da Saúde e dos Direitos Sexuais e Reprodutivos e Promoção da Igualdade de Gênero.**

United Nations Population Fund HQ. 2020. Disponível em:

< <https://brazil.unfpa.org/pt-br/publications/um-olhar-para-g%C3%AAnero> >.

Acesso em: 09 de setembro de 2021.

VASCONCELOS, Anselmo César et al. **Fatores associados ao comportamento da população durante o isolamento social na pandemia de COVID-19.** Ciênc. saúde coletiva 25 (suppl 1) 05 Jun 2020Jun 2020. Disponível em: <

<https://doi.org/10.1590/1413-81232020256.1.10792020> >. Acesso em: 09 de setembro de 2021.

World Bank. 2020. **Poverty and Shared Prosperity 2020: Reversals of Fortune.** Washington, DC: World Bank. doi: 10.1596/978-1-4648-1602-4. License: Creative Commons Attribution CC BY 3.0 IGO.

World Bank. 2020. Poverty and Shared Prosperity 2020:

Reversals of Fortune. Washington, DC: World Bank. doi: 10.1596/978-1-4648-1602-4. License: Creative

Commons Attribution CC BY 3.0 IGO

World Health Organization. **Actions for consideration in the care and protection of vulnerable population groups for COVID-19.** Maio, 2020. Disponível em: <

<https://www.who.int/westernpacific/internal-publications-detail/WPR-DSE-2020-021-eng> >. Acesso em: 09 de setembro de 2021.

World Health Organization. **Coronavirus disease 2019 (COVID-19) Situation Report**
- **94**. Abril, 2020. Disponível em: < <https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200423-sitrep-94-COVID-19.pdf> >. Acesso em: 09 de setembro de 2021.

World Health Organization. **Diagnostic testing for SARS-CoV-2**. Setembro, 2020.

World Health Organization. **Guidance on COVID-19 for the care of older people and people living in long-term care facilities, other nonacute care facilities and home care**. Julho, 2020. Disponível em: < <https://apps.who.int/iris/handle/10665/331913> >. Acesso em: 09 de setembro de 2021.

Yu CM, Wong RS, Wu EB, Kong SL, Wong J, Yip GW, Soo YO, Chiu ML, Chan YS, Hui D, Lee N, Wu A, Leung CB, Sung JJ. **Cardiovascular complications of severe acute respiratory syndrome**. Postgrad Med J. 2006 Feb;82(964):140-4. doi: 10.1136/pgmj.2005.037515. PMID: 16461478; PMCID: PMC2596695.