

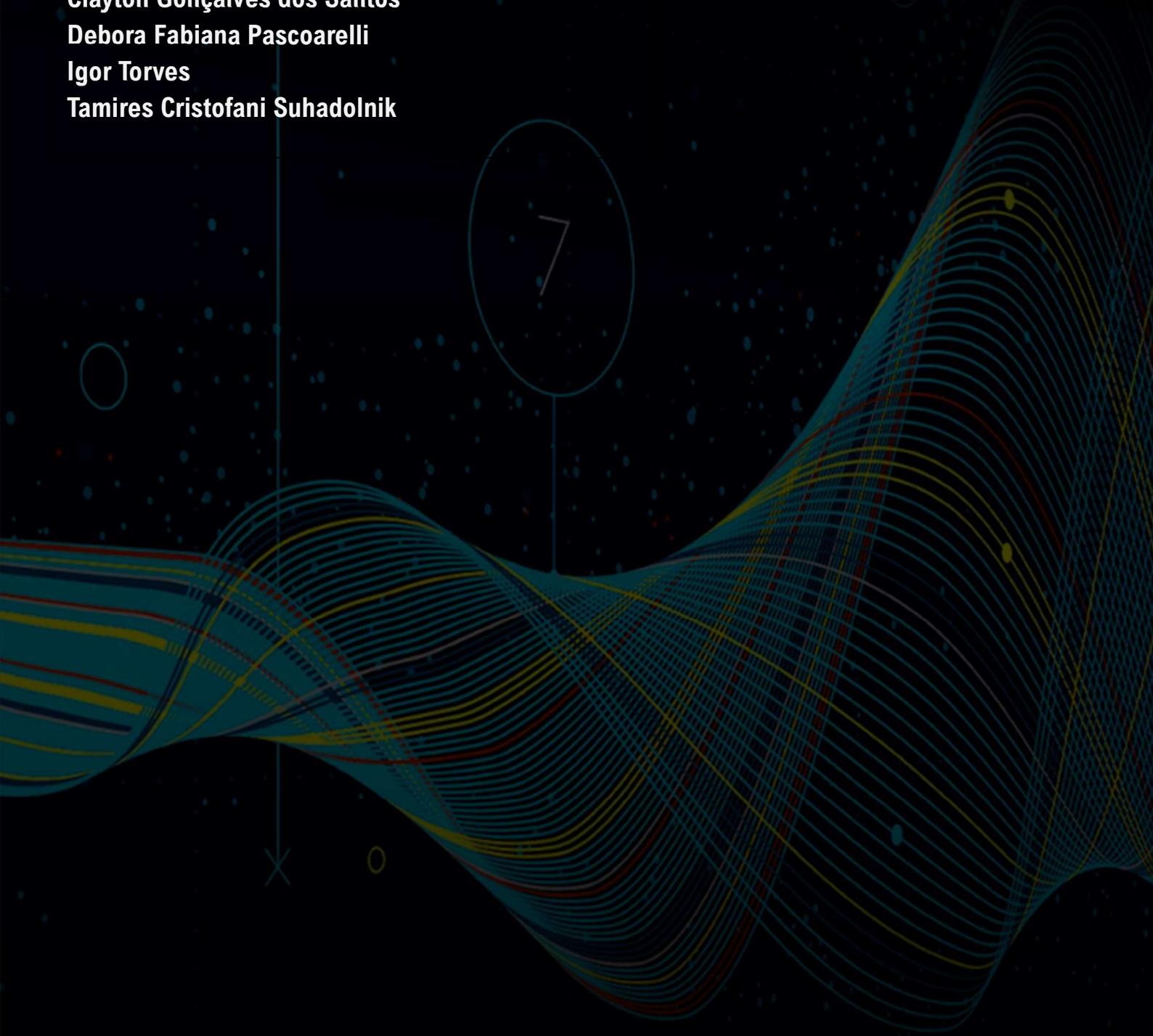
André Antônio Campos

Clayton Gonçalves dos Santos

Debora Fabiana Pascoarelli

Igor Torves

Tamires Cristofani Suhadolnik



DATA ANALYTICS

TECH CHALLENGE

FASE 03

Introdução

Neste relatório, nossa equipe foi contratada por um hospital de grande porte para realizar uma análise detalhada sobre o comportamento da população durante a pandemia de COVID-19, a fim de identificar indicadores críticos para auxiliar no planejamento de ações em caso de um novo surto. Utilizando a base de dados do **PNAD-COVID 19**, fornecida pelo IBGE, desenvolvemos uma abordagem estruturada para fornecer insights importantes, baseados em dados confiáveis e abrangentes.

Este trabalho explora a importância de compreender o comportamento da população durante a pandemia de COVID-19. Iniciaremos com uma **Análise Exploratória dos Dados**, focando em indicadores essenciais para essa compreensão, como a **Identificação e Controle** e as principais características gerais dos moradores em nível nacional.

Além disso, nosso projeto visa fornecer informações valiosas sobre as causas e resultados das infecções, as estratégias de recuperação implementadas, e as possíveis melhorias que poderiam ter sido feitas. O objetivo é gerar insights que ajudem a mitigar futuras ocorrências semelhantes e aprimorar a resposta a situações de emergência.

Período Analisado:

- Ano: 2020
- Meses: Maio, Junho e Julho

A análise será focada em três áreas principais:

- **Características clínicas dos sintomas** apresentados pela população;
- **Comportamento da população** ao longo do período da pandemia;
- **Características econômicas** da sociedade durante o surto.

Neste documento, apresentaremos a organização da base de dados, as perguntas selecionadas a partir da pesquisa do PNAD-COVID e uma análise das principais ações que o hospital deverá adotar para uma resposta eficaz a possíveis novos surtos da doença.

GitHub Path:

https://github.com/Tamireees/Tech_Challenge_Fase3/blob/main/TechChallenge_Fase3.ipynb

Fonte: base de dados do PNAD-COVID-19 do IBGE (<https://covid19.ibge.gov.br/pnad-covid/>)

Sumário

1.	Importação da base de dados do BigQuery.....	4
2.	Análise Exploratória: Características gerais dos moradores.....	4
	Distribuição etária dos Moradores	
	Proporção de Moradores por Sexo	
	Escolaridade da População	
	Distribuição Geográfica por UF	
	Distribuição por Situação Domiciliar	
	Distribuição de Cor e Raça	
	Análise Correlação: Cor e Raça x Nível de Escolaridade	
3.	Análise Exploratória: Sintomas da COVID-19.....	11
	Proporção de Sintomas por Mês	
	Distribuição de Sintomas por Faixa Etária	
	Sintomas por Região Geográfica	
	Análise de Coocorrência de Sintomas	
	Análise de Sintomas por Condição Socioeconômica	
	Porcentagem de sintomas por situação do domicílio	
	Pacientes com Covid-19 que permaneceram internados por mais de 1(um) dia	
	Análise por Faixa Etária	
	Análise por Gênero dos Pacientes:	
	Análise por Região	
	Pacientes que foram sedados ou entubados ou colocado em respirador artificial	
	Análise de Pacientes Entubados por Faixa Etária	
	Análise de Pacientes Entubados por Gênero	
	Análise de Pacientes Entubados por Região	
4.	Análise Exploratória: Características de trabalho das pessoas.....	28
	No trabalho (único ou principal) que tinha nessa semana, era	
	Tem carteira de trabalho assinada ou é funcionário público estatutário?	
	Cargo ou função, realiza no seu trabalho	
	Número da faixa do rendimento	
	Relação entre Faixa de Rendimento e Idade dos Moradores	
	Relação entre Faixa de Rendimento e Sexo	
	Relação entre Faixa de Rendimento e Escolaridade	
	Relação entre Faixa de Rendimento e Cor/Raça	
	Relação entre Faixa de Rendimento e Localização (Estado/UF)	
5.	Análise Exploratória: Rendimento de outras fontes dos moradores.....	37
	Análise dos Auxílios emergenciais relacionados ao coronavírus	
	Análise da somatória de valores recebidos por Nível de Escolaridade	
	Distribuição Percentual dos Valores Recebidos por Escolaridade	
	Análise da somatória de valores recebidos por Faixa Etária e Cor/Raça	
6.	Análise Exploratória: Características da habitação.....	41
	Análise Domiciliar	
7.	Conclusão Final.....	43

1. Importação da base de dados do BigQuery

Para iniciar a análise, utilizamos o Google BigQuery, uma plataforma robusta de armazenamento e consulta de dados em nuvem. O BigQuery facilita a manipulação de grandes volumes de dados de forma eficiente, permitindo consultas SQL para extrair as informações necessárias.

A base de dados do PNAD-COVID 19, fornecida pelo IBGE, foi importada para o BigQuery para garantir maior escalabilidade e acessibilidade durante o processo de análise. Isso nos permite trabalhar com dados de várias fontes e realizar consultas complexas com rapidez e precisão, essencial para a análise de indicadores clínicos, comportamentais e econômicos.

Além disso, o uso do BigQuery otimiza o processo de extração, transformação e carregamento (ETL), facilitando a limpeza, organização e preparação dos dados para análise exploratória e modelagem posterior.

```
!pip install google-cloud-bigquery pandas
!pip install db-dtypes
```

```
from google.cloud import bigquery
import db_dtypes
import pandas as pd

import os
os.environ['GOOGLE_APPLICATION_CREDENTIALS'] = r'Dados/credenciais.json'

import warnings
warnings.filterwarnings('ignore')
```

```
client = bigquery.Client()
```

```
query = """
    SELECT *
    FROM `plucky-tract-434600-b8.tech_challeng_fase_03.PNAD_COVID_19`
    LIMIT 100
"""
```

2. Análise Exploratória: Características gerais dos moradores

A análise dos dados de moradores foi realizada com foco em diversas características demográficas e socioeconômicas. Os dados coletados incluem informações sobre a localização, faixa etária, sexo, cor/raça e escolaridade dos indivíduos, abrangendo os meses de maio a julho de 2021.

- Localização:

Os dados incluem a sigla da unidade federativa, informações sobre se a residência está localizada em uma capital e detalhes sobre a região metropolitana. A classificação das áreas em urbanas e rurais foi feita com base na variável de situação domiciliar, que permite entender melhor a distribuição dos moradores entre ambientes urbanos e rurais.

- Faixa Etária:

A idade dos moradores foi capturada, permitindo a análise da distribuição etária. Essa informação é crucial para compreender as características demográficas da população e sua relação com outras variáveis, como saúde e acesso a serviços.

- Sexo:

Os dados incluem uma categorização de sexo, permitindo observar a proporção de homens e mulheres na amostra. Essa análise pode ajudar a identificar padrões de comportamento ou necessidades específicas entre os gêneros.

- Cor/Raça:

A categorização por cor/raça inclui cinco grupos principais: Branca, Preta, Amarela, Parda e Indígena, além de uma categoria de dados ignorados. Essa variável é importante para examinar as desigualdades sociais e econômicas que podem existir entre diferentes grupos raciais.

- Escolaridade:

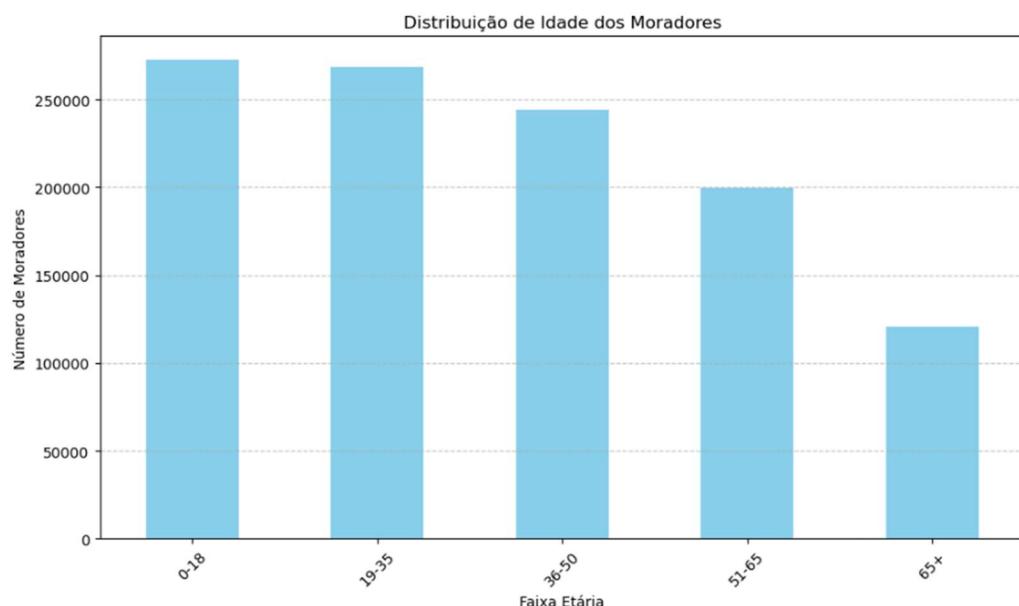
A análise da escolaridade foi feita com base em diversas categorias, desde a ausência de instrução até pós-graduação. Esta informação é essencial para entender o nível educacional da população e suas implicações no acesso a oportunidades e recursos.

- Conclusão

Esta **análise exploratória** oferece uma visão abrangente das características gerais da população, permitindo identificar tendências e padrões que podem influenciar diretamente a qualidade de vida e o acesso a serviços essenciais. Com base nos insights gerados, será possível formular hipóteses para investigações mais aprofundadas sobre o impacto dessas variáveis em diferentes segmentos da sociedade. Os resultados obtidos poderão servir como uma base sólida para a criação de **políticas públicas** mais eficientes e **intervenções sociais** direcionadas, contribuindo para uma resposta mais eficaz em futuras crises de saúde pública e para a melhoria do bem-estar social.

- **Distribuição etária dos Moradores**

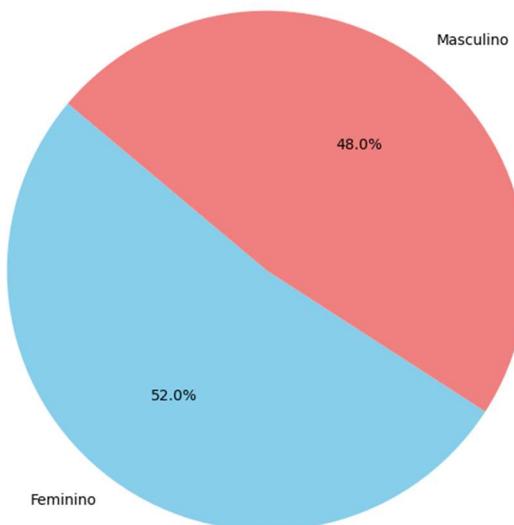
```
faixa_etaria
0-18      272577
19-35     268379
36-50     244507
51-65     199437
65+       120797
Name: count, dtype: int64
```



- Proporção de Moradores por Sexo

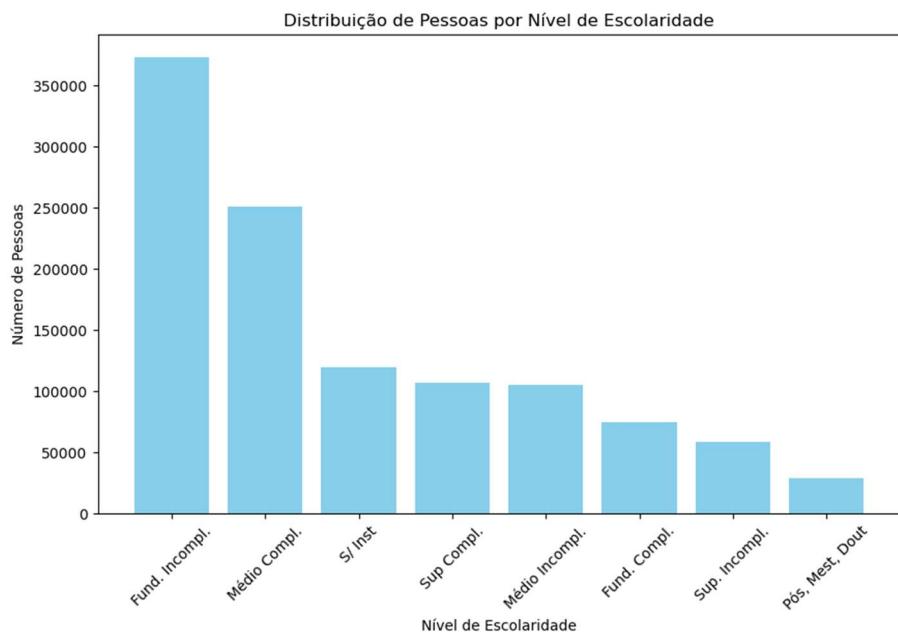
```
id_sexo_label
Feminino      51.985033
Masculino     48.014967
Name: proportion, dtype: float64
```

Proporção de Moradores por Sexo



- Escolaridade da População

```
id_escolaridade
Fund. Incompl.    372947
Médio Compl.     250790
S/ Inst          118887
Sup Compl.       106409
Médio Incompl.   104715
Fund. Compl.     74622
Sup. Incompl.    58131
Pós, Mest, Dout  28241
Name: count, dtype: int64
```

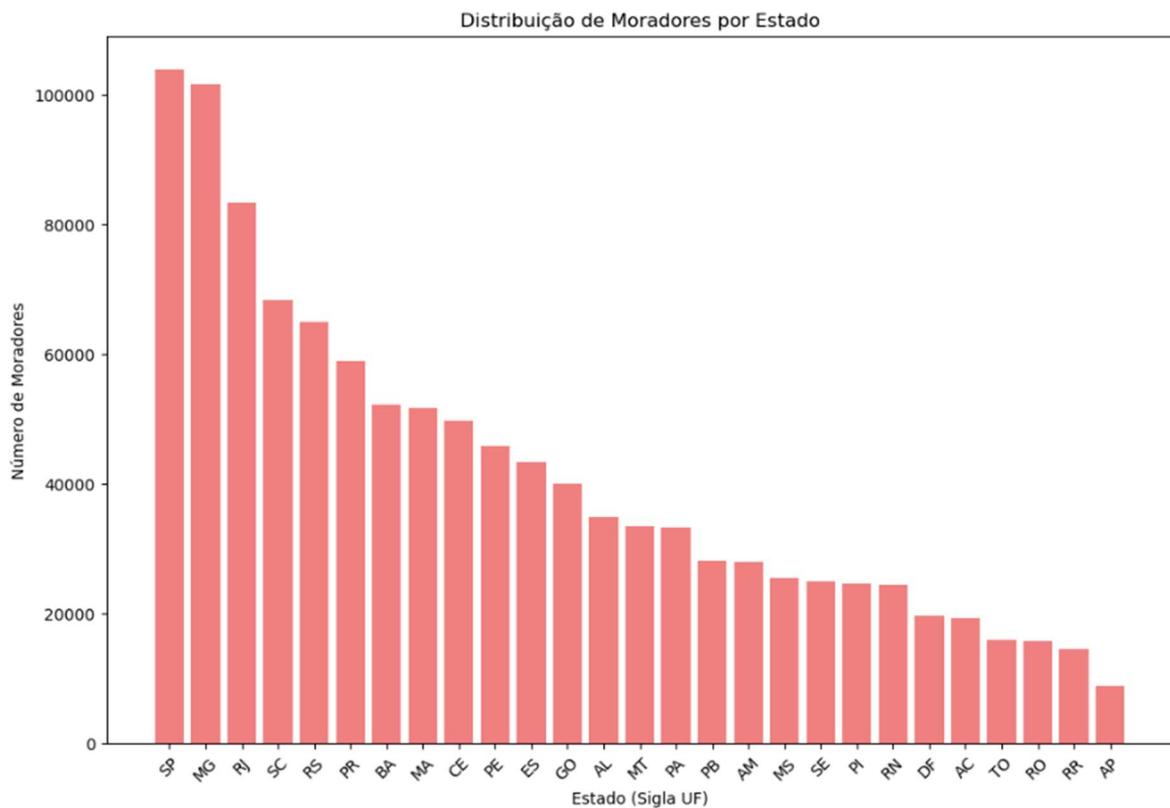


- **Distribuição Geográfica por UF**

```

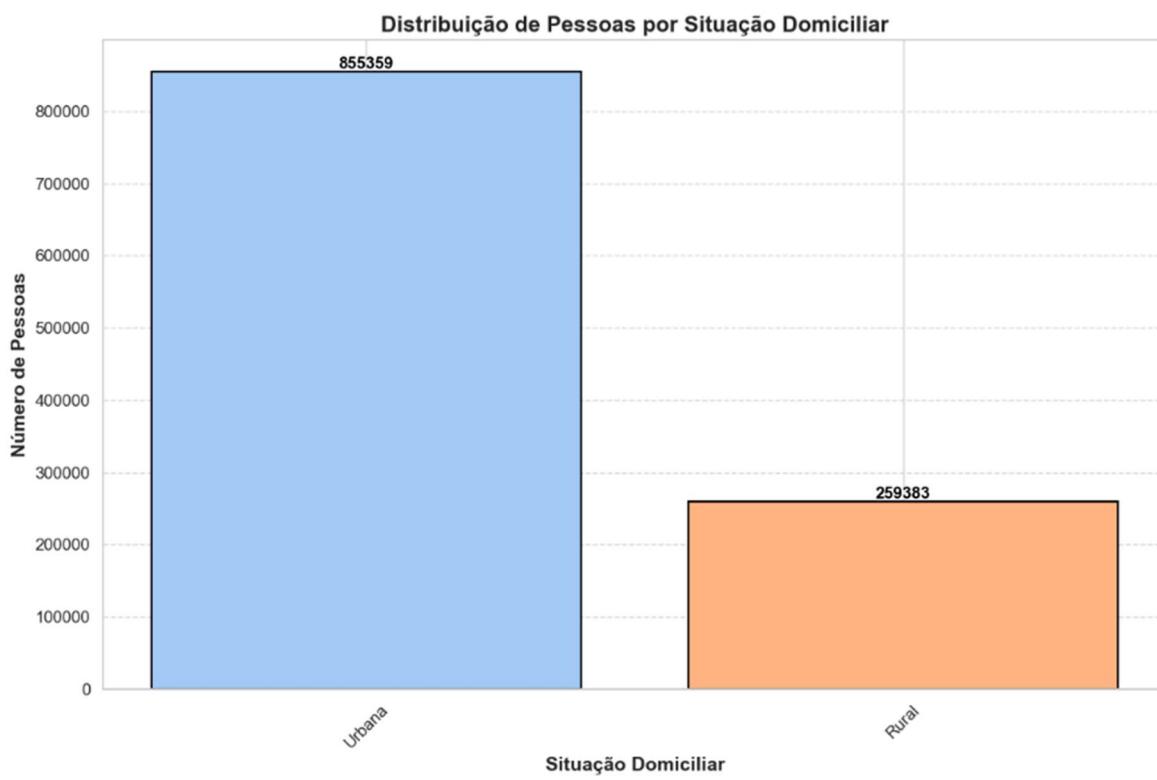
sigla_uf
SP      103897
MG      101557
RJ      83367
SC      68356
RS      64868
PR      58938
BA      52131
MA      51728
CE      49691
PE      45860
ES      43389
GO      40062
AL      34828
MT      33454
PA      33235
PB      28156
AM      27975
MS      25473
SE      24969
PI      24516
RN      24385
DF      19682
AC      19203
TO      15848
RO      15775
RR      14537
AP      8862
Name: count, dtype: int64

```

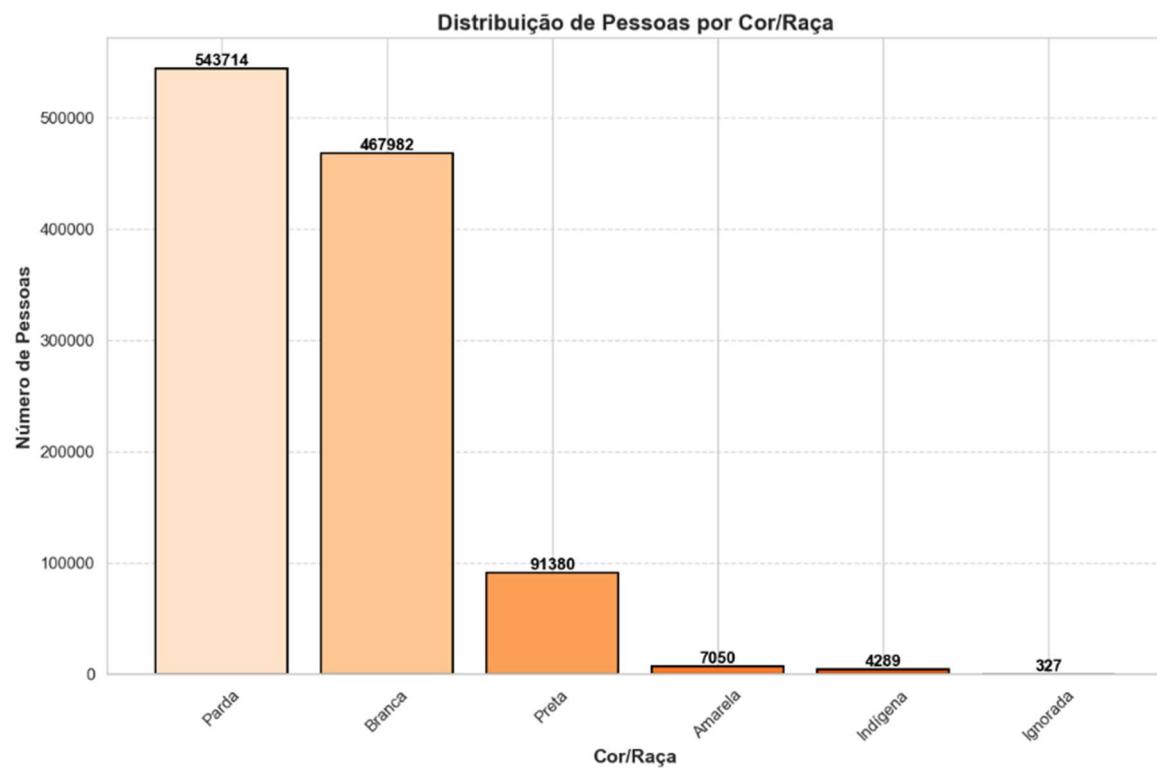


- **Distribuição por Situação Domiciliar**

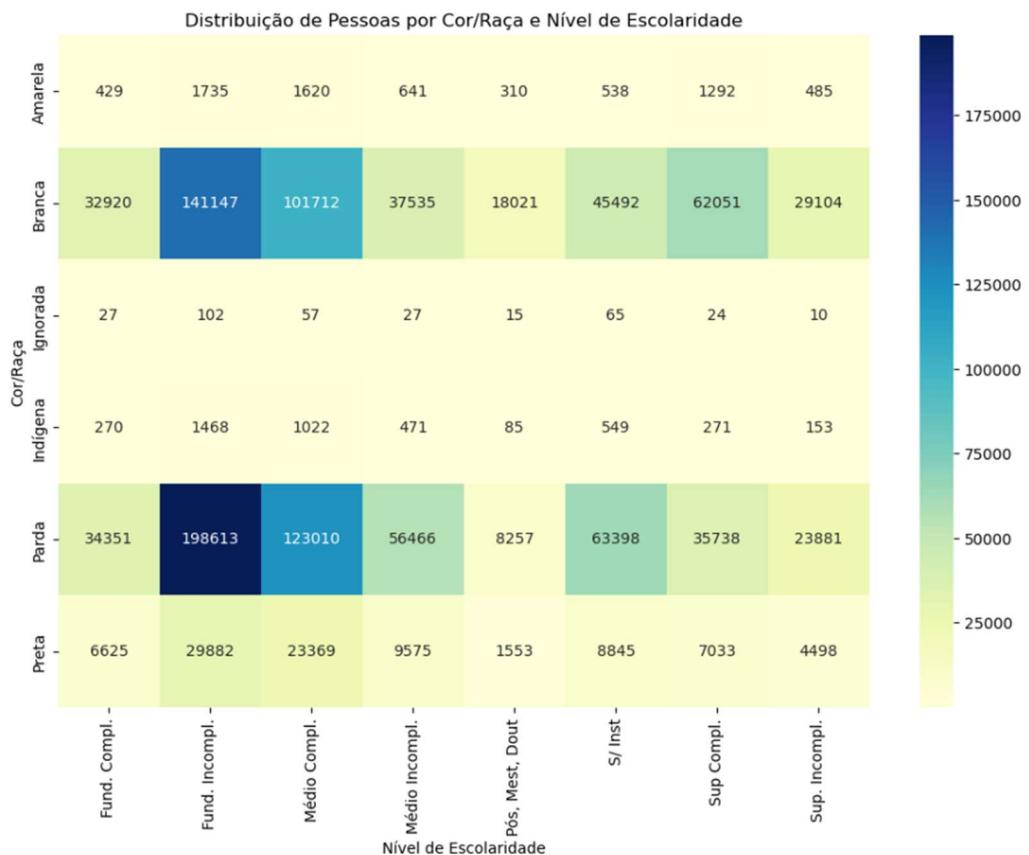
```
id_situacao_domicilio
Urbana      855359
Rural       259383
Name: count, dtype: int64
```



- Distribuição de Cor e Raça**



- **Análise Correlação: Cor e Raça x Nível de Escolaridade**



- **Conclusão - Cor e Raça x Nível de Escolaridade:**

A partir do gráfico de calor (heatmap) acima e sobre a distribuição de pessoas por cor/raça e nível de escolaridade, podemos observar algumas tendências importantes. As seguintes conclusões podem ser tiradas:

- Predominância de pessoas pardas e brancas:

As pessoas que se identificam como "Pardas" e "Brancas" têm a maior representação em praticamente todos os níveis de escolaridade. O número mais alto é observado entre pessoas pardas com o ensino fundamental incompleto, seguido de perto por pessoas brancas com o mesmo nível de escolaridade.

- Escolaridade baixa é comum em todos os grupos raciais:

Níveis de escolaridade mais baixos (ensino fundamental incompleto e completo) são os mais representados em todas as categorias de raça/cor. Há uma concentração significativa de pessoas em "Parda" e "Branca" nos níveis de escolaridade incompletos, indicando uma grande quantidade de pessoas que não completaram o ensino básico.

- Menor representação nas categorias "Amarela", "Indígena" e "Preta":

Pessoas das categorias "Amarela", "Indígena" e "Preta" aparecem em menor número, com distribuições significativamente menores em todos os níveis de escolaridade, se comparadas às categorias "Parda" e "Branca".

- Níveis de escolaridade superiores e pós-graduação:

As representações em níveis mais altos de escolaridade (como ensino superior completo ou incompleto, e pós-graduação) são menores para todos os grupos raciais, mas ainda são maiores para as categorias "Parda" e "Branca". Isso pode refletir barreiras de acesso à educação superior.

- Distribuição homogênea na categoria "Ignorada":

A categoria de cor/raça "Ignorada" é a menos representada em todos os níveis de escolaridade, sugerindo que uma pequena parcela da população não se identifica com as categorias estabelecidas ou não respondeu à questão sobre cor/raça.

3. Análise Exploratória: Sintomas da COVID-19

Inicialmente, vamos analisar a quantidade de pessoas que tiveram os sintomas listados abaixo:

- *Na semana passada teve febre?*

```
id_febre_sm
Não      1088212
Sim      20825
Ignorado 4458
Não Sabe 1247
Name: count, dtype: int64
```

Relevância do sintoma:

Prevalência da Febre: A grande maioria dos entrevistados (aproximadamente 97,6%) não apresentou febre, enquanto apenas cerca de 1,9% relatou ter febre. Isso sugere que, na amostra, a febre não foi um sintoma comum.

Dados Ignorados e "Não Sabe": Os dados ignorados (4.458) e aqueles que não sabem (1.247) representam um total de 5.705 pessoas. Isso indica que, mesmo entre os entrevistados, há uma quantidade considerável de incerteza ou falta de informação sobre sintomas.

- *Na semana passada teve tosse?*

```
id_tosse_sm
Não      1079433
Sim      29554
Ignorado 4481
Não Sabe 1274
Name: count, dtype: int64
```

Relevância do sintoma:

Prevalência da Tosse: Assim como na febre, a maioria dos entrevistados (aproximadamente 96,8%) não relatou tosse, enquanto cerca de 2,7% disseram ter tosse. Isso indica que a tosse também não foi um sintoma prevalente na amostra.

Dados Ignorados e "Não Sabe": Os dados ignorados (4.481) e os que não sabem (1.274) totalizam 5.755 pessoas. Isso mostra que ainda há uma quantidade significativa de incerteza ou falta de resposta, semelhante ao que observamos com a febre.

- *Na semana passada teve dor de garganta?*

```
id_dor_garganta_sm
Não      1086101
Sim      22769
Ignorado 4472
Não Sabe 1400
Name: count, dtype: int64
```

Relevância do sintoma:

Prevalência de Dor de Garganta: A maioria dos entrevistados (aproximadamente 97,4%) não relatou dor de garganta, enquanto cerca de 2,0% afirmaram ter esse sintoma. Isso indica que a dor de garganta não foi um sintoma comum na amostra.

Dados Ignorados e "Não Sabe": Os dados ignorados (4.472) e os que não sabem (1.400) totalizam 5.872 pessoas. Esse número expressivo sugere uma quantidade considerável de incerteza entre os entrevistados.

- Na semana passada teve dificuldade para respirar?

```
id_dif_respirar_sm
Não           1097018
Sim            11858
Ignorado      4544
Não Sabe       1322
Name: count, dtype: int64
```

Relevância do sintoma:

Prevalência da Dificuldade para Respirar: A grande maioria dos entrevistados (aproximadamente 98,4%) não relatou dificuldade para respirar, enquanto apenas cerca de 1,1% afirmou ter esse sintoma. Isso sugere que a dificuldade para respirar não foi comum na amostra analisada.

Dados Ignorados e "Não Sabe": Os dados ignorados (4.544) e os que não sabem (1.322) totalizam 5.866 pessoas. Assim como nas análises anteriores, isso indica uma quantidade significativa de incerteza ou falta de resposta.

- Na semana passada teve dor de cabeça?

```
id_dor_cabeca_sm
Não           1065819
Sim            42936
Ignorado      4489
Não Sabe       1498
Name: count, dtype: int64
```

Relevância do sintoma:

Prevalência da Dor de Cabeça: Aproximadamente 95,6% dos entrevistados não relataram dor de cabeça, enquanto cerca de 3,9% afirmaram ter esse sintoma. Isso indica que a dor de cabeça não foi o sintoma mais comum na amostra.

Dados Ignorados e "Não Sabe": Os dados ignorados (4.489) e os que não sabem (1.498) somam 5.987 pessoas. Esse número expressivo sugere que uma parte significativa da população entrevistada pode estar incerta sobre seus sintomas.

- Na semana passada teve dor no peito?

```
id_dor_peito_sm
Não           1099122
Sim            9623
Ignorado      4544
Não Sabe       1453
Name: count, dtype: int64
```

Relevância do sintoma:

Prevalência da Dor no Peito: A grande maioria dos entrevistados (aproximadamente 98,6%) não relatou dor no peito, enquanto apenas cerca de 0,9% afirmaram ter esse sintoma. Isso indica que a dor no peito não foi um sintoma comum na amostra.

Dados Ignorados e "Não Sabe": Os dados ignorados (4.544) e os que não sabem (1.453) totalizam 5.997 pessoas. Isso sugere uma quantidade considerável de incerteza ou falta de informação sobre os sintomas.

- Na semana passada teve náusea?

```
id_nausea_sm
Não      1099511
Sim      9284
Ignorado 4586
Não Sabe 1361
Name: count, dtype: int64
```

Relevância do sintoma:

Prevalência de Náusea: A maioria dos entrevistados (aproximadamente 98,6%) não relatou náusea, enquanto cerca de 0,8% afirmaram ter esse sintoma. Isso sugere que a náusea não foi um sintoma comum na amostra.

Dados Ignorados e "Não Sabe": Os dados ignorados (4.586) e os que não sabem (1.361) totalizam 5.947 pessoas. Essa quantidade significativa de respostas ignoradas ou incertas indica que pode haver dificuldades na identificação dos sintomas.

- Na semana passada teve nariz entupido ou escorrendo?

```
id_nariz_entupido_sm
Não      1076227
Sim      32592
Ignorado 4524
Não Sabe 1399
Name: count, dtype: int64
```

Relevância do sintoma:

Prevalência de Sintomas Nasais: Aproximadamente 96,5% dos entrevistados não relataram nariz entupido ou escorrendo, enquanto cerca de 2,9% afirmaram ter esse sintoma. Isso sugere que, embora não seja o sintoma mais comum, há uma porcentagem relevante que experimentou problemas nasais.

Dados Ignorados e "Não Sabe": Os dados ignorados (4.524) e os que não sabem (1.399) totalizam 5.923 pessoas. Isso mostra que uma parte considerável dos entrevistados pode estar incerta sobre seus sintomas.

- Na semana passada teve fadiga?

```
id_fadiga_sm
Não      1093677
Sim      15067
Ignorado 4552
Não Sabe 1446
Name: count, dtype: int64
```

Relevância do sintoma:

Prevalência de Fadiga: A maioria dos entrevistados (aproximadamente 98,1%) não relatou fadiga, enquanto cerca de 1,4% afirmaram ter esse sintoma. Isso sugere que a fadiga não foi um sintoma amplamente relatado na amostra.

Dados Ignorados e "Não Sabe": Os dados ignorados (4.552) e os que não sabem (1.446) totalizam 5.998 pessoas. Esse número significativo de respostas não informadas indica incerteza entre os entrevistados.

- Na semana passada teve dor nos olhos?

```
id_dor_olhos_sm
Não      1095551
Sim      11206
Ignorado 6533
Não Sabe 1452
Name: count, dtype: int64
```

Relevância do sintoma:

Prevalência de Dor nos Olhos: A maioria dos entrevistados (aproximadamente 98,3%) não relatou dor nos olhos, enquanto cerca de 1,0% afirmaram ter esse sintoma. Isso indica que a dor nos olhos não foi um sintoma comum na amostra.

Dados Ignorados e "Não Sabe": Os dados ignorados (6.533) e os que não sabem (1.452) somam 7.985 pessoas. Esse número indica uma quantidade considerável de incerteza ou falta de informação sobre esse sintoma.

- *Na semana passada teve perda de cheiro ou sabor?*

```
id_perda_chheiro_sabor_sm  
Não      1092954  
Sim      13946  
Ignorado 6372  
Não Sabe 1470  
Name: count, dtype: int64
```

Relevância do sintoma:

Prevalência de Perda de Cheiro ou Sabor: A maioria dos entrevistados (aproximadamente 98,0%) não relatou perda de cheiro ou sabor, enquanto cerca de 1,3% afirmaram ter esse sintoma. Isso sugere que, apesar de ser um sintoma associado à COVID-19, sua ocorrência nesta amostra foi relativamente baixa.

Dados Ignorados e "Não Sabe": Os dados ignorados (6.372) e os que não sabem (1.470) totalizam 7.842 pessoas. Isso indica uma quantidade significativa de incerteza ou falta de informação sobre esse sintoma.

- Na semana passada teve dor muscular?

```
id_dor_muscular_sm  
Não      1081564  
Sim      25318  
Ignorado 6396  
Não Sabe 1464  
Name: count, dtype: int64
```

Relevância do sintoma:

Prevalência de Dor Muscular: A maioria dos entrevistados (aproximadamente 97,0%) não relatou dor muscular, enquanto cerca de 2,3% afirmaram ter esse sintoma. Isso sugere que a dor muscular não foi um sintoma comum na amostra.

Dados Ignorados e "Não Sabe": Os dados ignorados (6.396) e os que não sabem (1.464) totalizam 7.860 pessoas. Essa quantidade significativa de respostas não informadas indica incerteza sobre os sintomas.

A partir da análise dessas perguntas, iremos identificar a prevalência de cada sintoma na população e os casos mais extremos. Em seguida, aprofundaremos a análise considerando fatores como a localidade dos entrevistados, e características socioeconômicas e demográficas, como renda, cor de pele, idade, entre outros, para compreender o impacto desses sintomas em diferentes grupos e regiões.

- **Conclusão dos Sintomas:**

Com base nas análises dos sintomas relatados na semana passada, observamos que a maioria dos entrevistados não apresentou sintomas clássicos da COVID-19, como febre, tosse, dificuldade para respirar e dor de garganta, com prevalências geralmente abaixo de 3%. Embora sintomas como dor muscular e perda de cheiro ou sabor tenham sido relatados, sua frequência também foi baixa.

Esses dados sugerem que muitos casos podem ser assintomáticos ou manifestar sintomas não típicos, o que destaca a importância de uma abordagem abrangente na vigilância e testagem

para a COVID-19. A considerável quantidade de respostas ignoradas ou de pessoas incertas sobre seus sintomas indica que a comunicação e a conscientização sobre a identificação de sinais da doença ainda são essenciais.

- **Proporção de Sintomas por Mês:**

```
# Porcentagem de sintomas por semana e mês
sintomas_semana = dados3.groupby(['ano', 'mes'])[colunas_sintomas].apply(lambda x: (x == 'Sim').sum() / x.count() * 100)
print(sintomas_semana)

  id_febre_sm  id_tosse_sm  id_dor_garganta_sm  id_dif_respirar_sm \
ano mes
2020 5          2.418510      3.463725          2.574820          1.454599
     6          1.716631      2.353450          1.840690          0.979621
     7          1.518094      2.207900          1.758875          0.791845

  id_dor_cabeca_sm  id_dor_peito_sm  id_nausea_sm \
ano mes
2020 5            5.094387      1.196086          1.102185
     6            3.563879      0.792876          0.756419
     7            3.007294      0.630457          0.663776

  id_nariz_entupido_sm  id_fadiga_sm  id_dor_olhos_sm \
ano mes
2020 5            3.842476      1.704236          1.387895
     6            2.638288      1.198363          0.948934
     7            2.371631      1.183082          0.713233

  id_perda_chheiro_sabor_sm  id_dor_muscular_sm
ano mes
2020 5            1.760634      3.046040
     6            1.154562      2.069662
     7            0.883472      1.766684
```

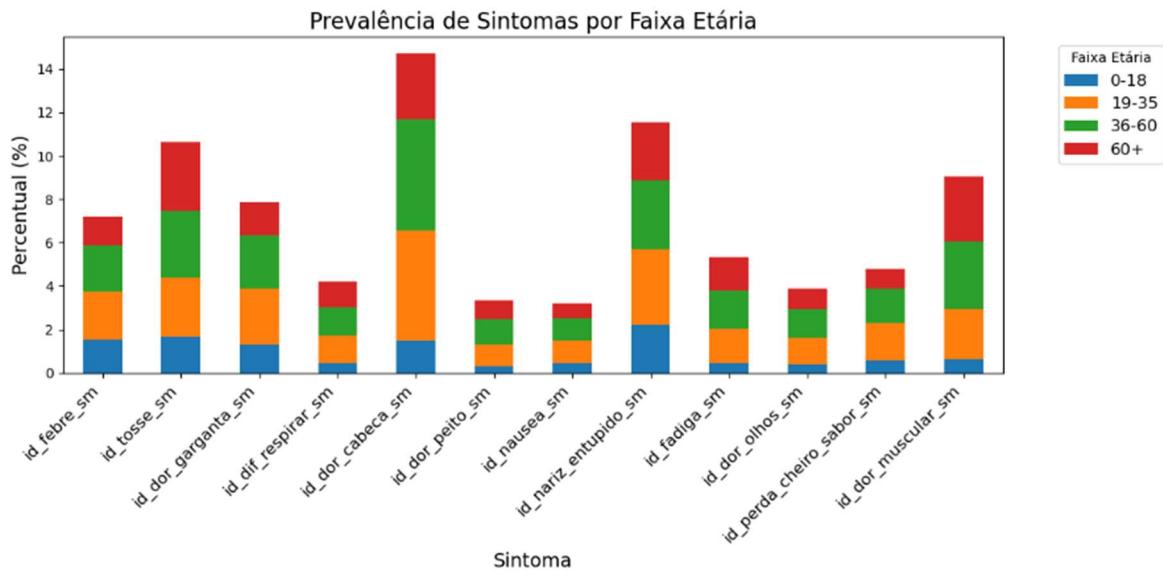
Conclusão:

O resultado apresenta a prevalência de sintomas em percentagem relatados por mês (maio, junho e julho de 2020) durante o ano de 2020. Com base nesses dados, podemos tirar as seguintes conclusões:

- Há uma tendência de redução na prevalência da maioria dos sintomas de maio para julho de 2020. Isso pode sugerir que o pico de casos ocorreu em maio, com um declínio nos meses subsequentes.
- Dor de cabeça foi o sintoma mais relatado em todos os meses, com uma prevalência de 5,09% em maio, reduzindo para 3,01% em julho. Outros sintomas comuns incluem tosse (3,46% em maio) e dor muscular (3,04% em maio), também mostrando uma redução ao longo do tempo.
- Sintomas como dor no peito, náusea, fadiga, e perda de cheiro/sabor foram relatados com menor frequência. A dor no peito foi consistentemente um dos sintomas menos comuns, variando de 1,19% em maio a 0,63% em julho.

A análise revela que a prevalência de sintomas relacionados à COVID-19 atingiu um pico em maio de 2020 e diminuiu gradualmente nos meses seguintes. Dor de cabeça, tosse, e dor muscular foram os sintomas mais comuns, enquanto a dificuldade respiratória e a perda de cheiro e sabor diminuíram significativamente com o tempo, sugerindo uma possível redução na transmissão ou efetividade das medidas de controle da pandemia ao longo desses meses.

- Distribuição de Sintomas por Faixa Etária:**



A análise dos sintomas por faixa etária revela os seguintes padrões e tendências:

- Incidência Geral de Sintomas:**

As faixas etárias entre 19-35 e 36-60 anos apresentam a maior porcentagem de sintomas relatados, com destaque para a faixa 36-60, que lidera em diversos sintomas. Isso pode indicar que adultos de meia-idade são mais propensos a experimentar sintomas mais intensos ou a relatá-los com mais frequência.

- Sintomas Respiratórios:**

- A tosse é mais frequente em todas as faixas etárias, especialmente em pessoas de 36-60 anos (3.08%) e em idosos (60+) (3.17%).

- A febre é mais comum na faixa de 19-35 anos (2.22%) e 36-60 anos (2.12%), enquanto a dificuldade para respirar é ligeiramente mais alta nas faixas de 36-60 anos e 60+.

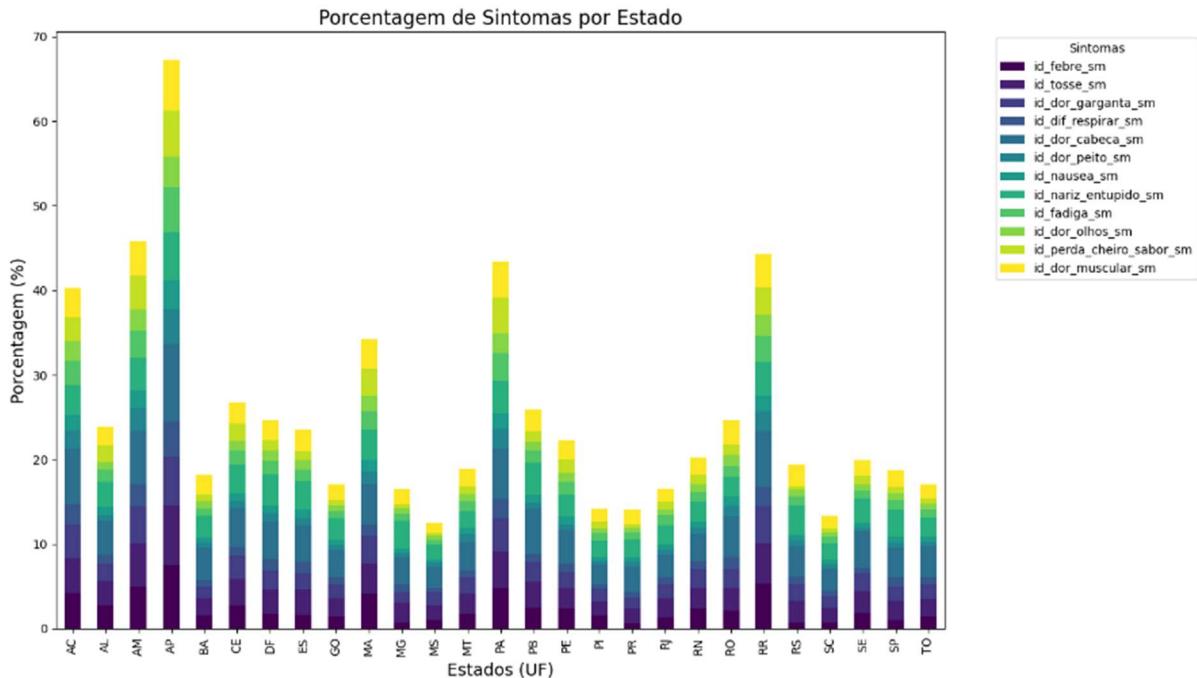
- Dor de Cabeça: É mais comum em adultos de 36-60 anos (5.11%) e em pessoas de 19-35 anos (5.06%).
- Fadiga: A fadiga também é maior nas faixas de 36-60 anos (1.78%) e 19-35 anos (1.56%).
- Dor Muscular: É mais relatada na faixa de 36-60 anos (3.13%), seguida pelos idosos (60+) (2.96%).
- Perda de Olfato/Paladar: É mais presente na faixa de 19-35 anos (1.72%) e diminui com a idade, sendo menos comum em crianças/adolescentes (0-18).

- Conclusão Geral:**

Os adultos jovens (19-35 anos) e os adultos de meia-idade (36-60 anos) são os grupos que apresentam maior incidência de sintomas da doença, incluindo sintomas respiratórios, dores de cabeça, dores musculares, e perda de olfato/paladar.

A faixa etária mais jovem (0-18 anos) e os idosos (60+) tendem a apresentar uma menor porcentagem de sintomas ou menos intensidade, indicando que os sintomas podem ser mais leves ou menos frequentes nesses grupos.

- **Sintomas por Região Geográfica:**



Analisando o gráfico de barras empilhadas, podemos fazer as seguintes observações:

- Variedade de sintomas por estado:

Estados como Amapá (AP) e Amazonas (AM) têm uma porcentagem significativamente maior de sintomas em relação a outros estados. Isso indica que esses estados registraram uma prevalência maior de sintomas entre seus habitantes.

Em contraste, estados como Espírito Santo (ES), Goiás (GO), e São Paulo (SP) apresentam porcentagens menores, sugerindo uma menor prevalência ou relato de sintomas nesses locais.

- Sintomas mais comuns:

O sintoma "id_dor_cabeca_sm" (dor de cabeça), que aparece em uma cor azul-clara, parece ser bastante prevalente em vários estados.

A perda de cheiro/sabor ("id_perda_chheiro_sabor_sm") e dores musculares ("id_dor_muscular_sm"), que aparecem nas cores verde e amarelo respectivamente, também são sintomas frequentes, especialmente em estados como Amapá (AP) e Roraima (RR).

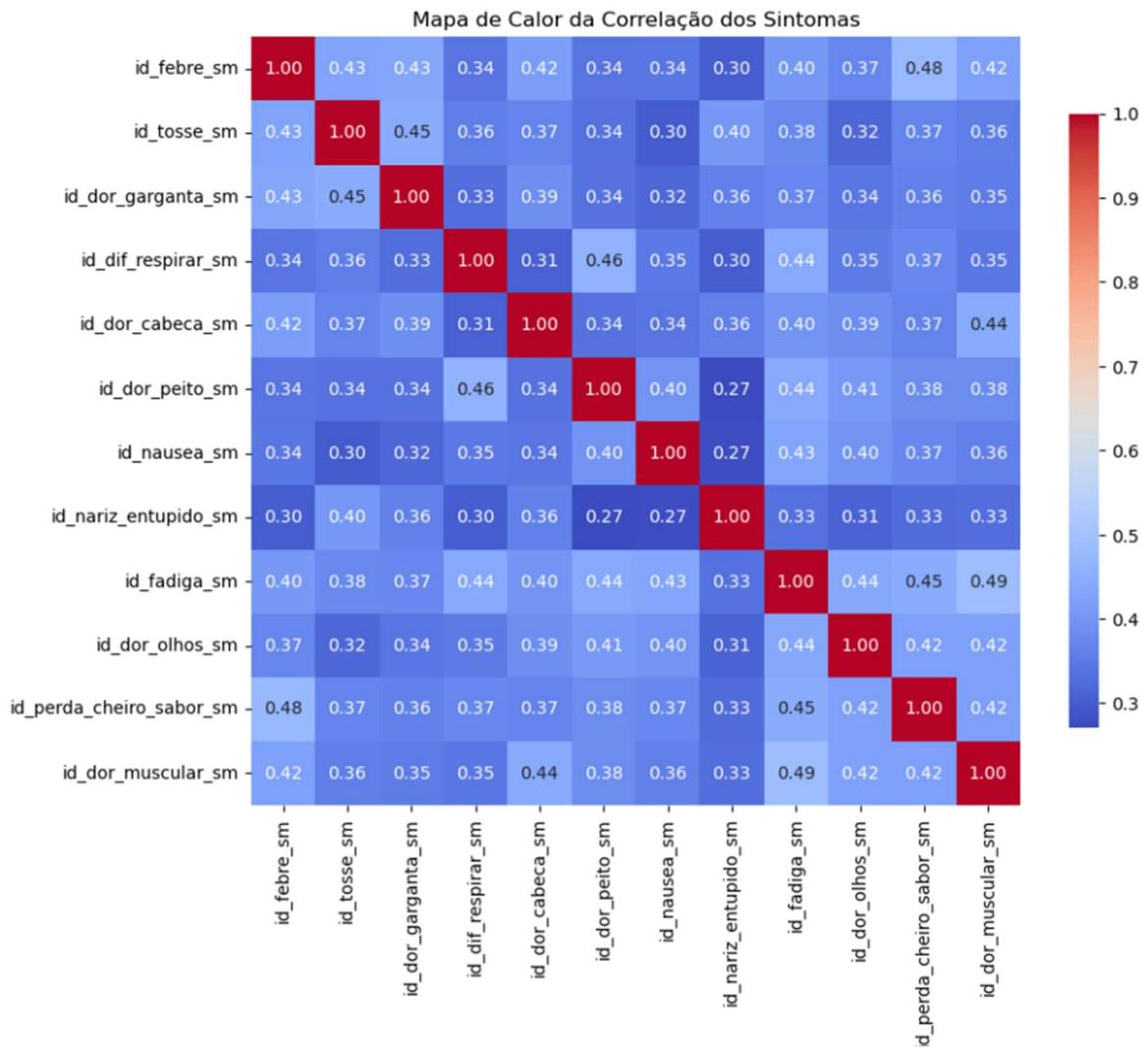
- Variação por região:

Os estados da região Norte, como AP, AM, e RR, têm uma prevalência maior de sintomas em comparação com os estados do Sudeste e Centro-Oeste. Isso pode sugerir uma maior incidência de doenças que causam esses sintomas ou diferenças na forma como os sintomas foram relatados em diferentes regiões do país.

Conclusão:

O gráfico evidencia variações significativas na prevalência de sintomas entre os estados do Brasil. Os estados da região Norte tendem a apresentar maiores taxas de sintomas, enquanto os estados do Sul e Sudeste apresentam taxas mais baixas.

- Análise de Coocorrência de Sintomas:



Conclusão sobre a Análise de Coocorrência de Sintomas:

A análise de coocorrência dos sintomas, conforme ilustrado no mapa de calor, revela padrões importantes que podem ajudar a entender melhor a interação entre diferentes sintomas em um determinado contexto, como em doenças respiratórias ou infecciosas.

- Sintomas com Alta Coocorrência:

Fadiga tem uma forte correlação com vários outros sintomas, como dor muscular (0.49), dor nos olhos (0.45) e dor de cabeça (0.44). Isso sugere que a fadiga é um sintoma central que tende a acompanhar outros desconfortos físicos em muitos casos.

A dor muscular e fadiga são altamente correlacionadas, indicando que, em muitos pacientes, o cansaço extremo está relacionado a dores musculares, o que é típico em doenças infecciosas ou estados febris.

- Sintomas com Correlação Moderada:

A febre tem uma correlação moderada com a perda de cheiro e sabor (0.48), sugerindo que, em alguns casos, a febre está associada a essa perda sensorial, um padrão comum em infecções virais, como a COVID-19.

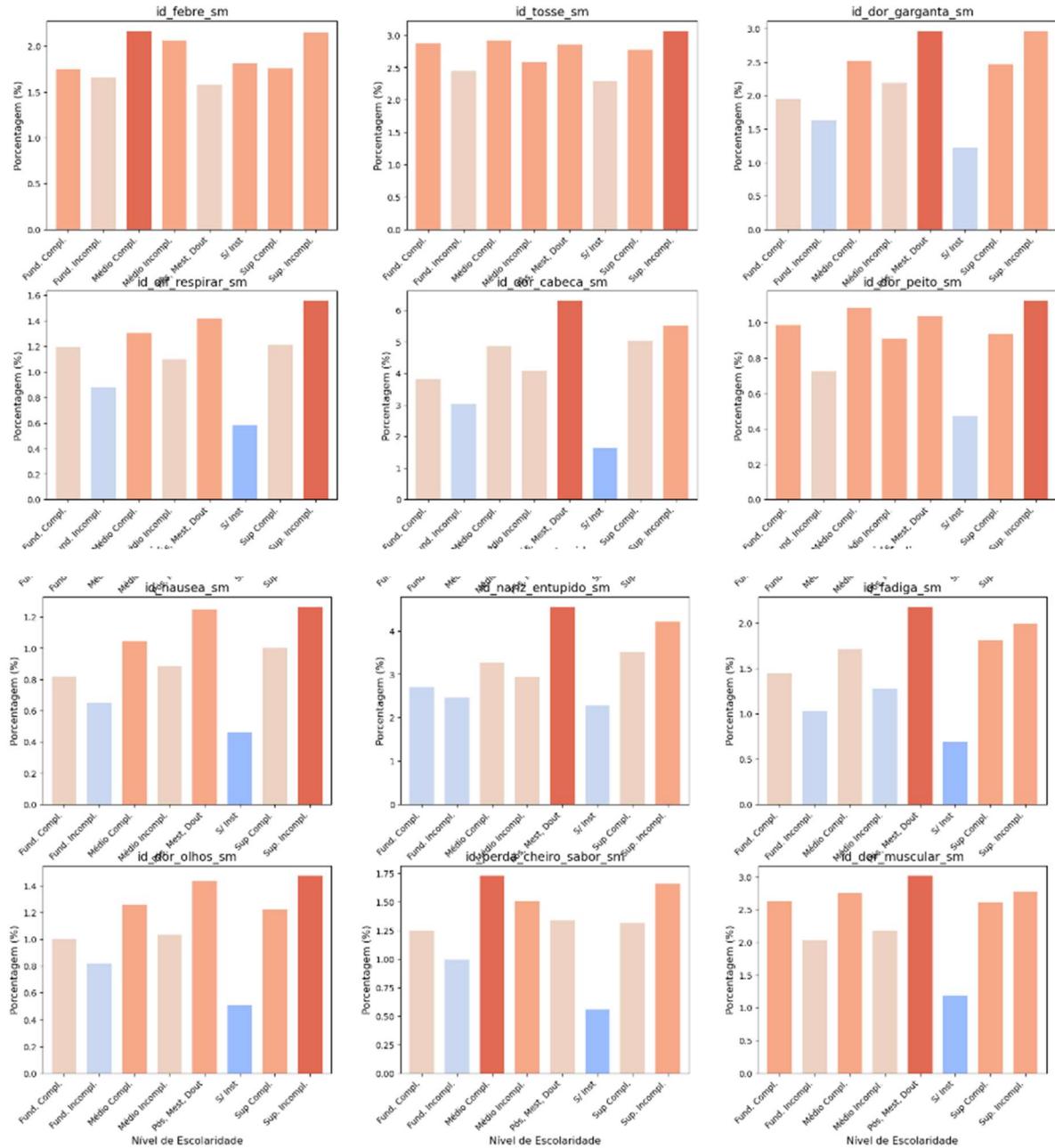
Dor de cabeça também mostra correlação moderada com diversos sintomas, o que pode indicar que ela frequentemente coexiste com outros sintomas em quadros clínicos mais amplos.

- Sintomas com Baixa Coocorrência:

A náusea tem uma correlação baixa com a maioria dos sintomas, como dor de garganta (0.30) e nariz entupido (0.27), sugerindo que esses sintomas aparecem de maneira mais isolada, não sendo comumente associados em um mesmo quadro clínico.

Nariz entupido também tem correlações baixas com outros sintomas, exceto com tosse (0.40), o que pode indicar que, embora seja um sintoma desconfortável, não tem tanta relevância em quadros que incluem sintomas mais severos.

- Análise de Sintomas por Condição Socioeconômica:



Conclusão referente aos sintomas por nível de escolaridade:

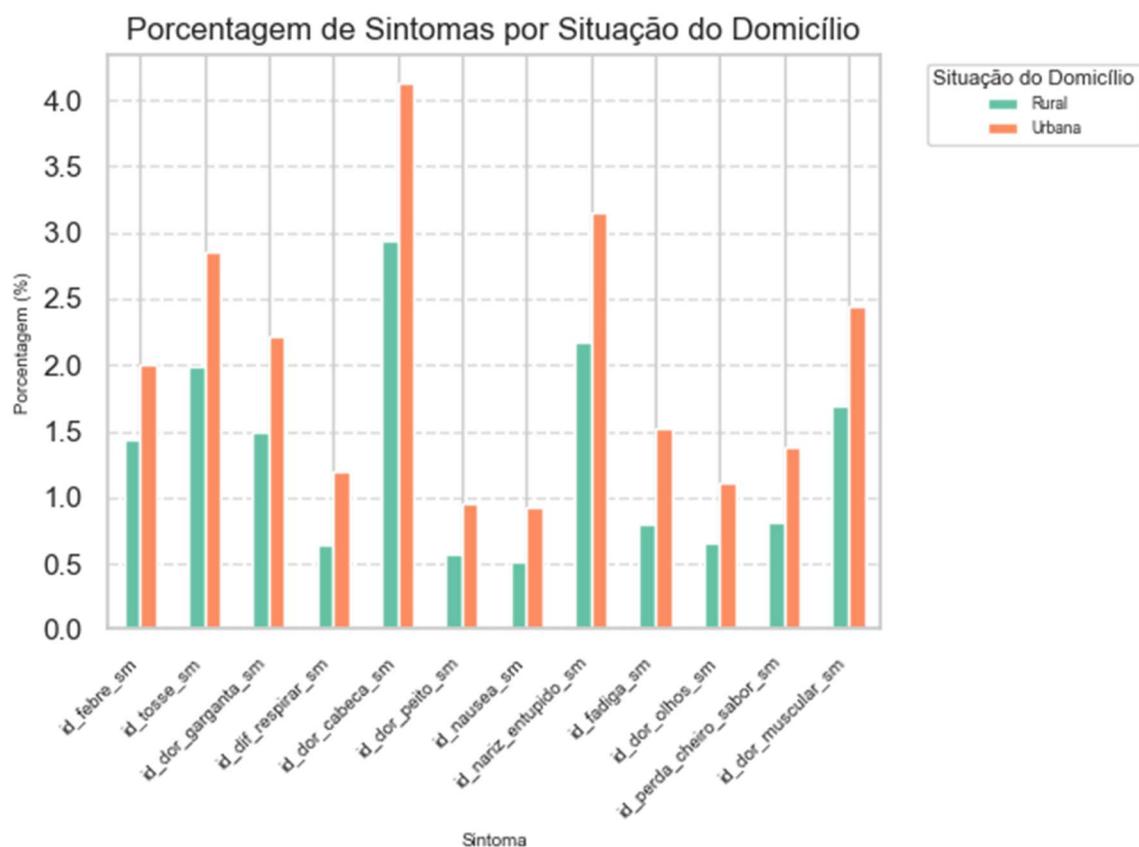
A análise dos dados mostra que há uma relação entre o nível de escolaridade e a percepção de certos sintomas.

- Dor de cabeça é mais prevalente entre indivíduos com níveis de escolaridade mais altos, como pós-graduação, mestrado e doutorado. Isso pode sugerir que pessoas mais educadas têm maior consciência e relatam mais sintomas, possivelmente devido a pressões acadêmicas e profissionais.

- Fadiga e perda de cheiro/sabor também são reportados com mais frequência entre aqueles com níveis educacionais mais altos. Isso pode indicar que esses indivíduos têm uma maior capacidade de reconhecer e relatar sintomas relacionados à saúde.
- Em contraste, pessoas sem instrução tendem a relatar menos sintomas, o que pode refletir uma combinação de fatores, incluindo subnotificação e diferenças no acesso ao sistema de saúde, que afetam a percepção e o diagnóstico.
- Para febre e tosse, a distribuição dos sintomas é mais uniforme entre os diferentes níveis de escolaridade, indicando que esses sintomas podem ser mais universalmente reconhecidos e não estão tão relacionados à educação.

Essas observações sugerem que a escolaridade pode influenciar a percepção e o relato de sintomas, indicando a necessidade de considerar esses fatores na abordagem de saúde pública e na comunicação sobre doenças.

- **Porcentagem de sintomas por situação do domicílio:**



Conclusão de sintomas por situação do domicílio:

A análise indica que a população urbana tende a relatar uma maior prevalência de sintomas em relação à população rural. Fatores como densidade populacional, poluição do ar, condições de trabalho e estresse podem estar contribuindo para essa diferença. A população rural, por outro lado, apresenta menores taxas de sintomas, o que pode estar relacionado ao ambiente de vida mais espacoso e natural.

Esses dados sugerem a necessidade de políticas de saúde pública direcionadas para áreas urbanas, onde os sintomas parecem ser mais prevalentes, especialmente aqueles relacionados a problemas respiratórios e dor física.

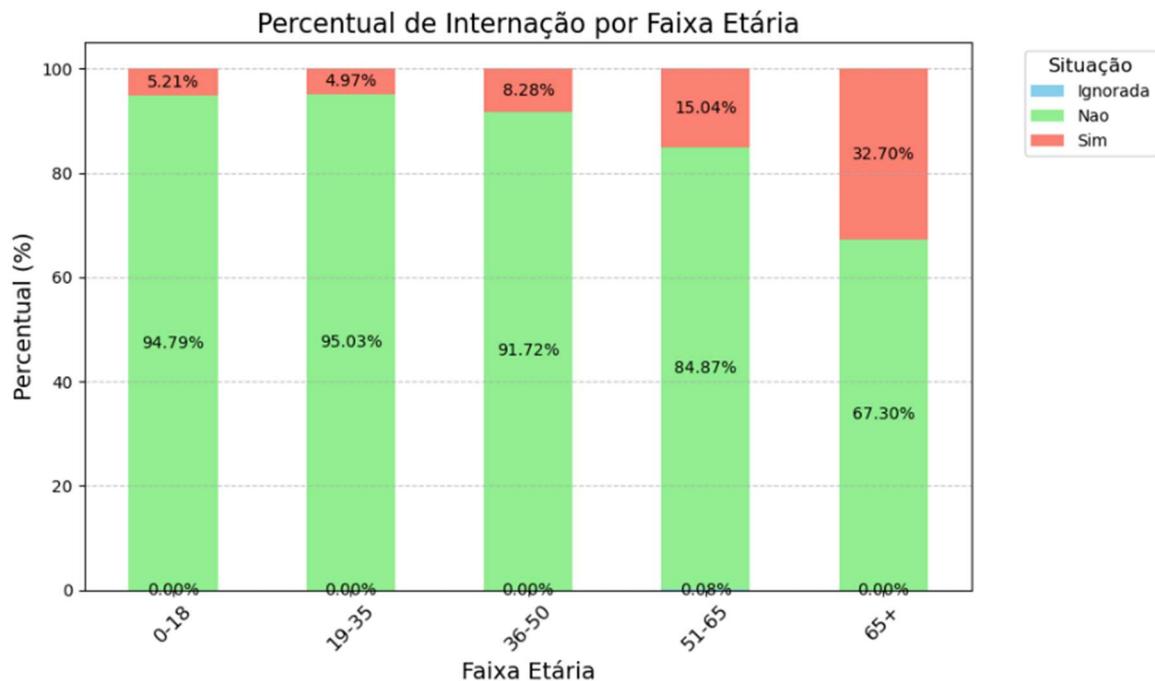
- **Pacientes com Covid-19 que permaneceram internados por mais de 1(um) dia**

```
Percentual de id_internado_1_mais:
id_internado_1_mais
Nao Aplicavel    99.446419
Nao                0.491414
Sim                 0.062077
Ignorada            0.000090
Name: proportion, dtype: float64
```

Ao visualizar o resultado do total de pacientes que ficaram internados por mais de um dia, constatamos que 99.46% não foram informados na base de dados. Para uma análise mais precisa deste critério, serão desconsiderados os resultados classificados como Nao Aplicavel para uma equidade da base.

```
Percentual de id_internado_1_mais (excluindo 'Nao Aplicavel'):
id_internado_1_mais
Nao      88.770053
Sim      11.213742
Ignorada 0.016205
Name: proportion, dtype: float64
```

- **Análise por Faixa Etária:**



Resumo da Análise Exploratória da Faixa Etária do Paciente:

O gráfico “Percentual de Internação por Faixa Etária” mostra a distribuição das internações hospitalares em diferentes faixas etárias. Aqui estão os principais pontos observados:

Faixa Etária 0-19 Anos:

- Não Internados: 94.79%
- Internados: 5.21%
- Ignorados: Pequena porcentagem não especificada

Faixa Etária 20-59 Anos:

- Não Internados: 95.03%

- Internados: 4.97%
- Ignorados: Pequena porcentagem não especificada

Faixa Etária 60-79 Anos:

- Não Internados: 91.72%
- Internados: 8.28%
- Ignorados: Pequena porcentagem não especificada

Faixa Etária 80+ Anos:

- Não Internados: 84.87%
- Internados: 15.13%
- Ignorados: Pequena porcentagem não especificada

Observações: Tendência de Internação: A taxa de internação aumenta significativamente com a idade. Enquanto as faixas etárias mais jovens (0-19 e 20-59 anos) têm uma baixa porcentagem de internações, as faixas etárias mais velhas (60-79 e 80+ anos) mostram um aumento notável.

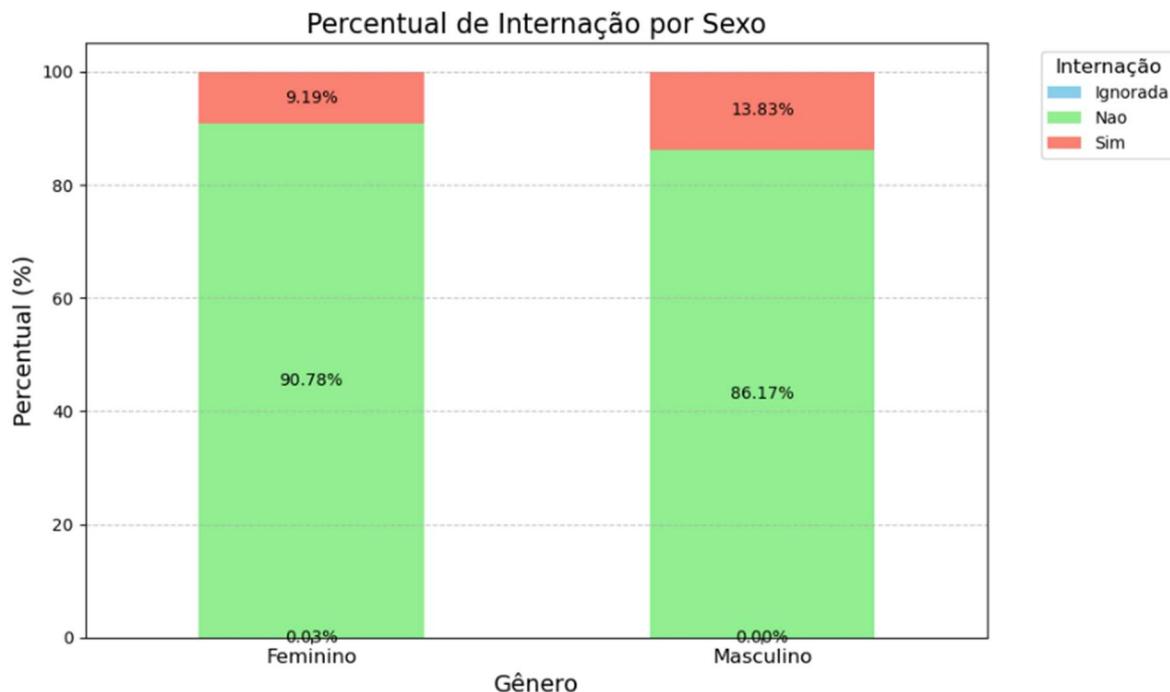
Faixa Etária 80+ Anos: Esta faixa etária apresenta a maior taxa de internação (15.13%), indicando uma maior necessidade de cuidados hospitalares entre os idosos.

Faixa Etária 60-79 Anos: Também apresenta uma taxa de internação relativamente alta (8.28%), sugerindo que a necessidade de internação começa a aumentar significativamente a partir dos 60 anos.

Conclusão:

A análise sugere que a idade é um fator importante na probabilidade de internação hospitalar. As faixas etárias mais avançadas (60+ anos) têm uma maior necessidade de cuidados hospitalares, o que pode ser crucial para o planejamento de recursos de saúde e alocação de serviços médicos.

• Análise por Gênero dos Pacientes:



Resumo da Análise Exploratória Por Gênero do Paciente:

O gráfico “Percentual de Internação por Sexo” mostra a distribuição das internações hospitalares entre os sexos feminino e masculino. Aqui estão os principais pontos observados:

Sexo Feminino:

Internados: 90.78%

Não Internados: 9.19%

Ignorados: Pequena porcentagem não especificada

Sexo Masculino:

Internados: 86.17%

Não Internados: 13.83%

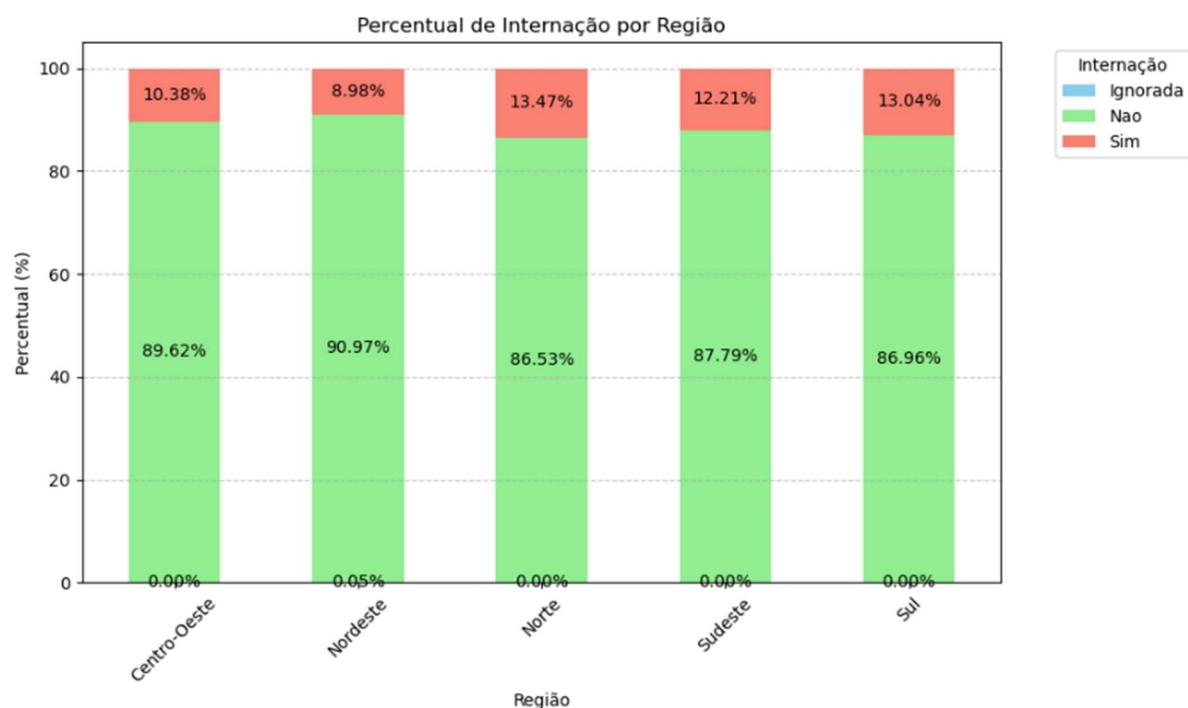
Ignorados: Pequena porcentagem não especificada

Observações: Tendência de Internação: A taxa de internação é alta para ambos os sexos, mas é ligeiramente maior para o sexo feminino (90.78%) em comparação ao sexo masculino (86.17%). Diferença entre os Sexos: A diferença na taxa de internação entre os sexos pode ser influenciada por vários fatores, incluindo condições de saúde específicas de cada sexo, acesso a cuidados de saúde, e comportamentos relacionados à busca por atendimento médico.

Conclusão:

A análise sugere que há uma alta taxa de internação hospitalar tanto para homens quanto para mulheres, com uma leve predominância de internações entre as mulheres. Esses dados podem ser úteis para entender melhor as necessidades de saúde de diferentes grupos e para planejar recursos e serviços de saúde de maneira mais eficaz.

- **Análise por Região:**

**Resumo da Análise Exploratória Por Região:**

O gráfico “Percentual de Internação por Região” mostra a distribuição das internações hospitalares em diferentes regiões do Brasil. Aqui estão os principais pontos observados:

Centro-Oeste:

- Internados: 10.38%
- Não Internados: 89.62%

Nordeste:

- Internados: 8.98%

- Não Internados: 90.97%

Norte:

- Internados: 13.47%
- Não Internados: 86.53%

Sudeste:

- Internados: 12.21%
- Não Internados: 87.79%

Sul:

- Internados: 13.04%
- Não Internados: 86.96%

Observações: Tendência de Internação: A taxa de internação varia entre as regiões, com o Norte (13.47%) e o Sul (13.04%) apresentando as maiores taxas de internação. Diferença Regional: As regiões Centro-Oeste (10.38%) e Nordeste (8.98%) têm taxas de internação mais baixas em comparação com as outras regiões.

Conclusão:

A análise sugere que há variações regionais significativas nas taxas de internação hospitalar no Brasil. Essas diferenças podem ser influenciadas por fatores como acesso a serviços de saúde, condições socioeconômicas e prevalência de doenças específicas em cada região. Esses dados são importantes para o planejamento de recursos de saúde e a alocação de serviços médicos de maneira mais eficaz.

- **Pacientes com CODIV19 que foram sedados ou entubados ou colocado em respirador artificial**

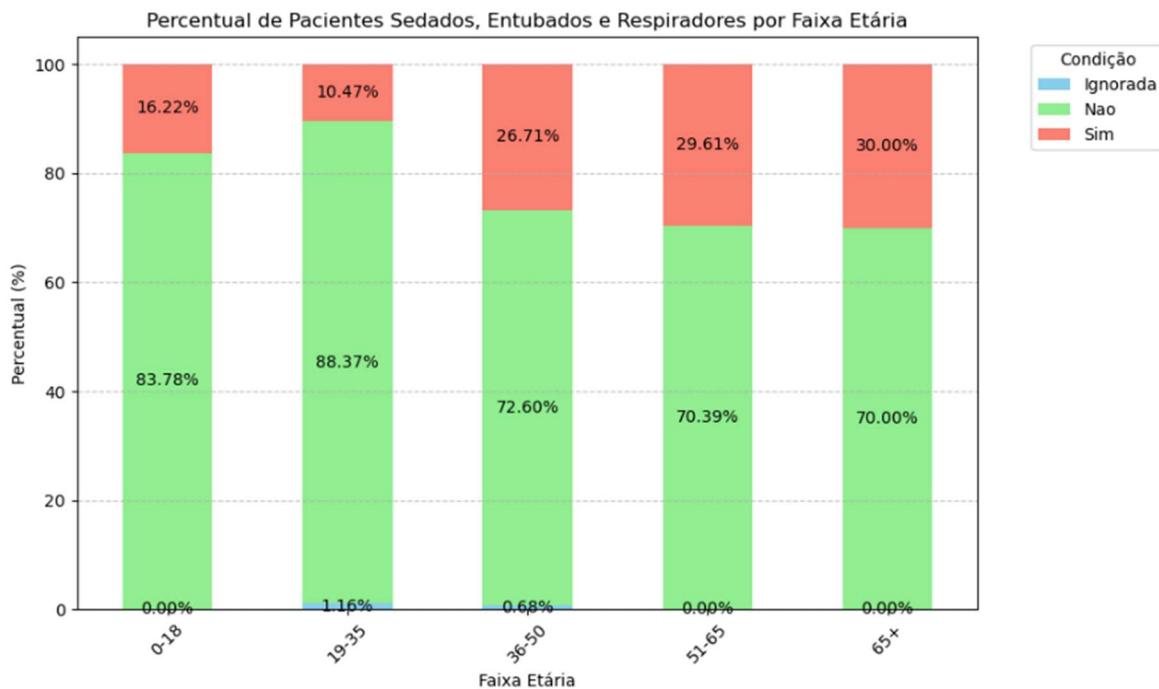
- Durante a internação, foi sedado, entubado e colocado em respiração artificial com ventilador

```
Percentual de id_sedado_entubado:
id_sedado_entubado
Nao Aplicavel    99.937923
Nao              0.045750
Sim              0.016147
Ignorada        0.000179
Name: proportion, dtype: float64
```

Ao visualizar o resultado do total de pacientes que ficaram internados por mais de um dia, constatamos que 99.93% não foram informados na base de dados. Para uma análise mais precisa deste critério, serão desconsiderados os resultados classificados como Nao Aplicavel para uma equidade da base.

```
Percentual de id_sedado_entubado (excluindo 'Nao Aplicavel'):
id_sedado_entubado
Nao          73.699422
Sim          26.011561
Ignorada    0.289017
Name: proportion, dtype: float64
```

- Análise de Pacientes Entubados por Faixa Etária:**



Resumo da Análise Exploratória Por Faixa Etária:

O gráfico “Percentual Pacientes Sedados, Entubados, Respirador por Faixa Etária” mostra a distribuição dos pacientes que estão sedados e entubados com respirador por diferentes faixas etárias. Aqui estão os principais pontos observados:

Faixa Etária 0-9:

- Não Sedados/Entubados: 83.78%
- Sedados/Entubados: 16.22%
- Ignorada: Muito pequeno, quase insignificante

Faixa Etária 10-19:

- Não Sedados/Entubados: 88.37%
- Sedados/Entubados: 10.47%
- Ignorada: 1.16%

Faixa Etária 20-39:

- Não Sedados/Entubados: 72.60%
- Sedados/Entubados: 26.71%
- Ignorada: 0.68%

Faixa Etária 40-59:

- Não Sedados/Entubados: 70.39%
- Sedados/Entubados: 29.61%
- Ignorada: Muito pequeno, quase insignificante

Faixa Etária 60+:

- Não Sedados/Entubados: 70.00%
- Sedados/Entubados: 29.61%
- Ignorada: Muito pequeno, quase insignificante

Observações:

Tendência Geral: A maioria dos pacientes em todas as faixas etárias não está sedada ou entubada.

Faixa Etária 20-39: Apresenta um aumento significativo no percentual de pacientes sedados e entubados (26.71%). Faixa Etária 60+: Também tem um percentual relativamente alto de pacientes sedados e entubados (29.61%).

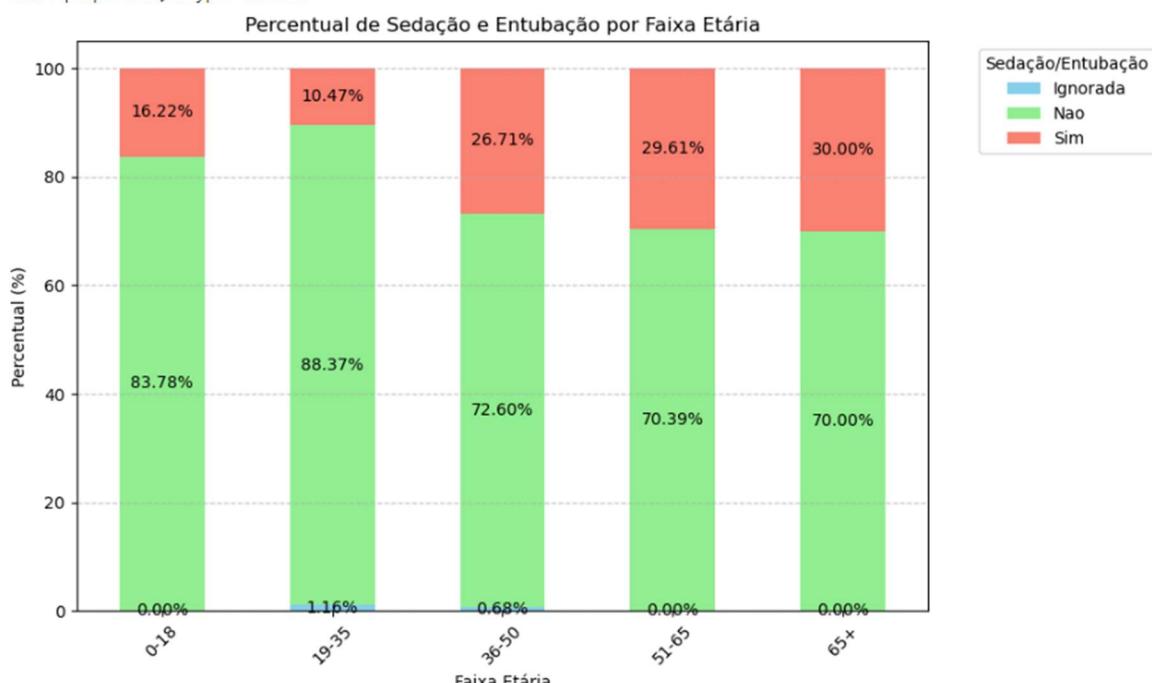
Conclusão:

Os dados indicam que, embora a maioria dos pacientes não necessite de sedação ou entubação, há um aumento notável nesses procedimentos em faixas etárias mais altas, especialmente a partir dos 20 anos. Isso pode refletir a gravidade das condições médicas que afetam essas faixas etárias.

- Análise de Pacientes Entubados por Gênero:**

Percentual de id_sedado_entubado (excluindo 'Nao Aplicavel'):

```
id_sedado_entubado
Nao      73.699422
Sim     26.011561
Ignorada 0.289017
Name: proportion, dtype: float64
```



Resumo da Análise Exploratória Por Faixa Etária:

O gráfico “Percentual de Sedação e Entubação por Faixa Etária” mostra a distribuição dos pacientes que estão sedados e entubados por diferentes faixas etárias.

Faixa Etária 0-19:

- Não Sedados/Entubados: 83.78%
- Sedados/Entubados: 16.22%
- Ignorada: Muito pequeno, quase insignificante

Faixa Etária 20-39:

- Não Sedados/Entubados: 72.60%
- Sedados/Entubados: 26.71%
- Ignorada: 0.68%

Faixa Etária 40-59:

- Não Sedados/Entubados: 70.39%
- Sedados/Entubados: 29.61%
- Ignorada: Muito pequeno, quase insignificante

Faixa Etária 60-79:

- Não Sedados/Entubados: 70.00%

- Sedados/Entubados: 29.61%
- Ignorada: Muito pequeno, quase insignificante

Faixa Etária 80+:

- Não Sedados/Entubados: 70.00%
- Sedados/Entubados: 29.61%
- Ignorada: Muito pequeno, quase insignificante

Observações:

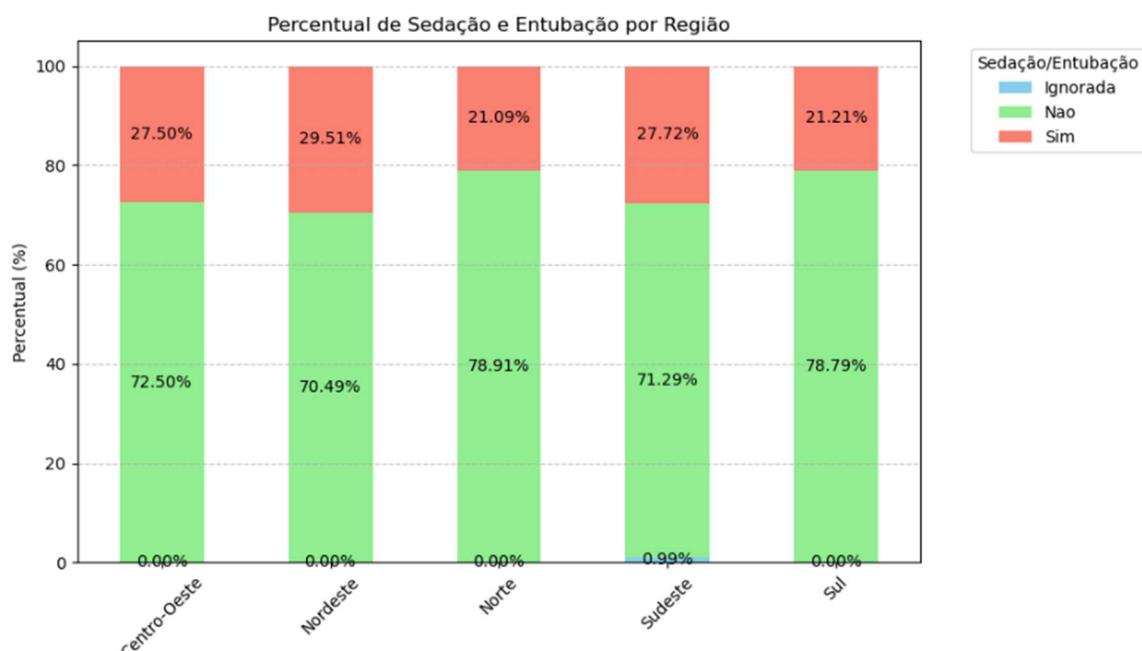
Tendência Geral: A maioria dos pacientes em todas as faixas etárias não está sedada ou entubada.

Faixa Etária 20-39: Apresenta um aumento significativo no percentual de pacientes sedados e entubados (26.71%). Faixa Etária 60-79 e 80+: Também têm percentuais relativamente altos de pacientes sedados e entubados (29.61%).

Conclusão:

Os dados indicam que, embora a maioria dos pacientes não necessite de sedação ou entubação, há um aumento notável nesses procedimentos em faixas etárias mais altas, especialmente a partir dos 20 anos. Isso pode refletir a gravidade das condições médicas que afetam essas faixas etárias.

• Análise de Pacientes Entubados por Região:



Resumo da Análise Exploratória Por Região:

O gráfico “Percentual de Sedação e Intubação por Região” mostra a distribuição dos pacientes que estão sedados e entubados por diferentes regiões do Brasil.

Centro-oeste:

- Não Sedados/Entubados: Maioria
- Sedados/Entubados: Menor percentual
- Ignorada: Muito pequeno, quase insignificante

Nordeste:

- Não Sedados/Entubados: Maioria
- Sedados/Entubados: Menor percentual
- Ignorada: Muito pequeno, quase insignificante

Norte:

- Não Sedados/Entubados: Maioria
- Sedados/Entubados: Menor percentual
- Ignorada: Muito pequeno, quase insignificante

Sudeste:

- Não Sedados/Entubados: Maioria
- Sedados/Entubados: Menor percentual
- Ignorada: Muito pequeno, quase insignificante

Sul:

- Não Sedados/Entubados: Maioria
- Sedados/Entubados: Menor percentual
- Ignorada: Muito pequeno, quase insignificante

Observações:

Tendência Geral: A maioria dos pacientes em todas as regiões não está sedada ou entubada.

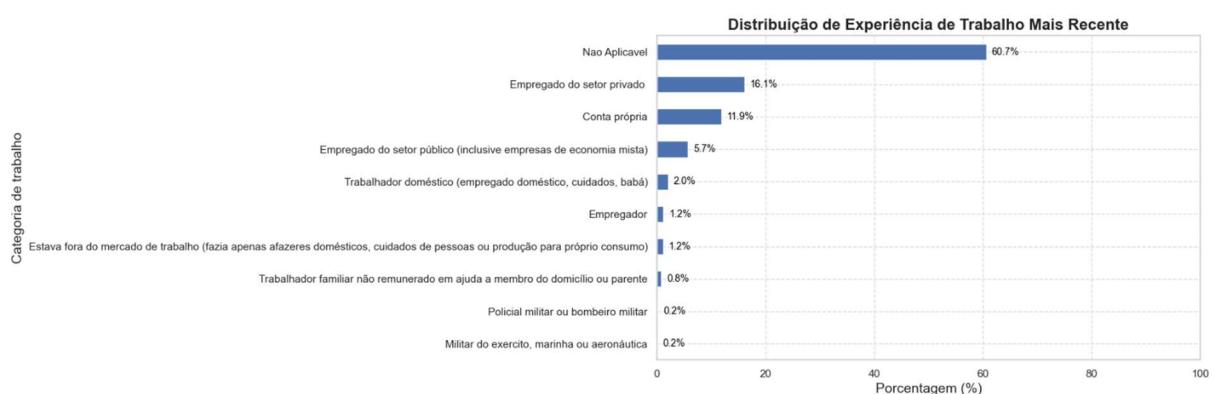
Percentual de Sedação e Entubação: Relativamente baixo em todas as regiões, com pequenas variações.

Conclusão:

Os dados indicam que, independentemente da região, a maioria dos pacientes não necessita de sedação ou entubação. As variações entre as regiões são mínimas, sugerindo uma consistência nas práticas médicas ou nas condições dos pacientes.

4. Análise Exploratória: Características de trabalho das pessoas

- *No trabalho (único ou principal) que tinha nessa semana, era*



A análise mostra que 60,7% dos respondentes classificaram sua situação de trabalho como "Não Aplicável", possivelmente refletindo desemprego, afastamento ou incerteza no mercado de trabalho devido à COVID-19.

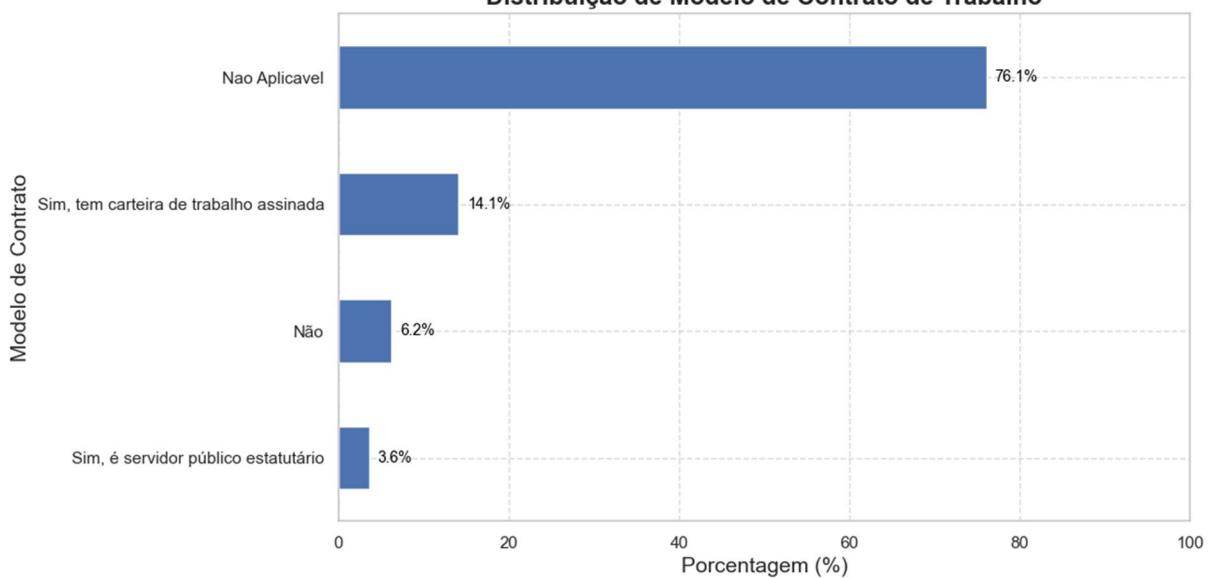
Entre os que declararam ter experiência recente, 16,1% trabalham no setor privado, um grupo potencialmente afetado por mudanças como home office ou redução de jornada. Além disso, 11,9% são trabalhadores por conta própria, uma categoria vulnerável às flutuações econômicas e queda na demanda durante a pandemia.

O setor público (5,7%) e trabalhadores domésticos (2%) também aparecem, mas em menor número, enquanto categorias como empregadores (1,2%) e trabalhadores familiares não remunerados (0,8%) indicam impactos econômicos distintos.

Essa distribuição revela como a pandemia afetou o mercado de trabalho de maneira desigual, destacando a importância de políticas de apoio para diferentes grupos.

- Tem carteira de trabalho assinada ou é funcionário público estatutário?**

Distribuição de Modelo de Contrato de Trabalho

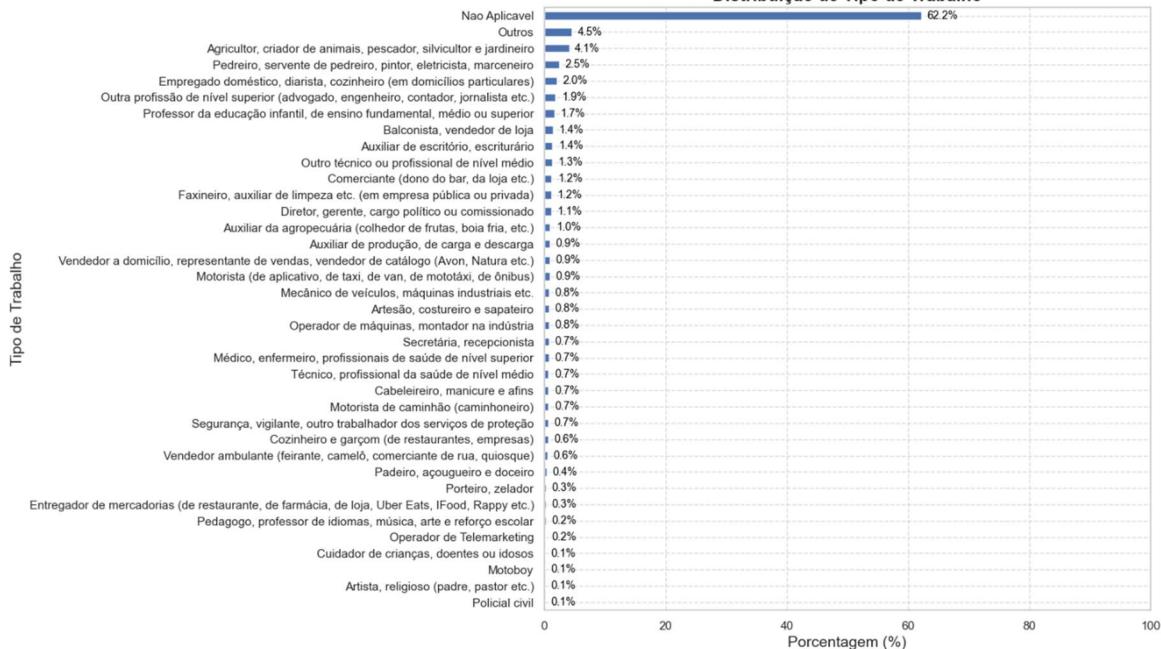


Entre os respondentes que indicaram um modelo de contrato, 14,1% têm carteira de trabalho assinada, o que sugere um grupo com maior estabilidade formal no mercado de trabalho durante o período pandêmico. 6,2% dos respondentes afirmaram não ter contrato formal, o que pode incluir trabalhadores informais ou autônomos, mais vulneráveis aos impactos da pandemia.

Por fim, 3,6% são servidores públicos estatutários, um grupo com maior proteção e estabilidade no emprego, mesmo diante das adversidades econômicas.

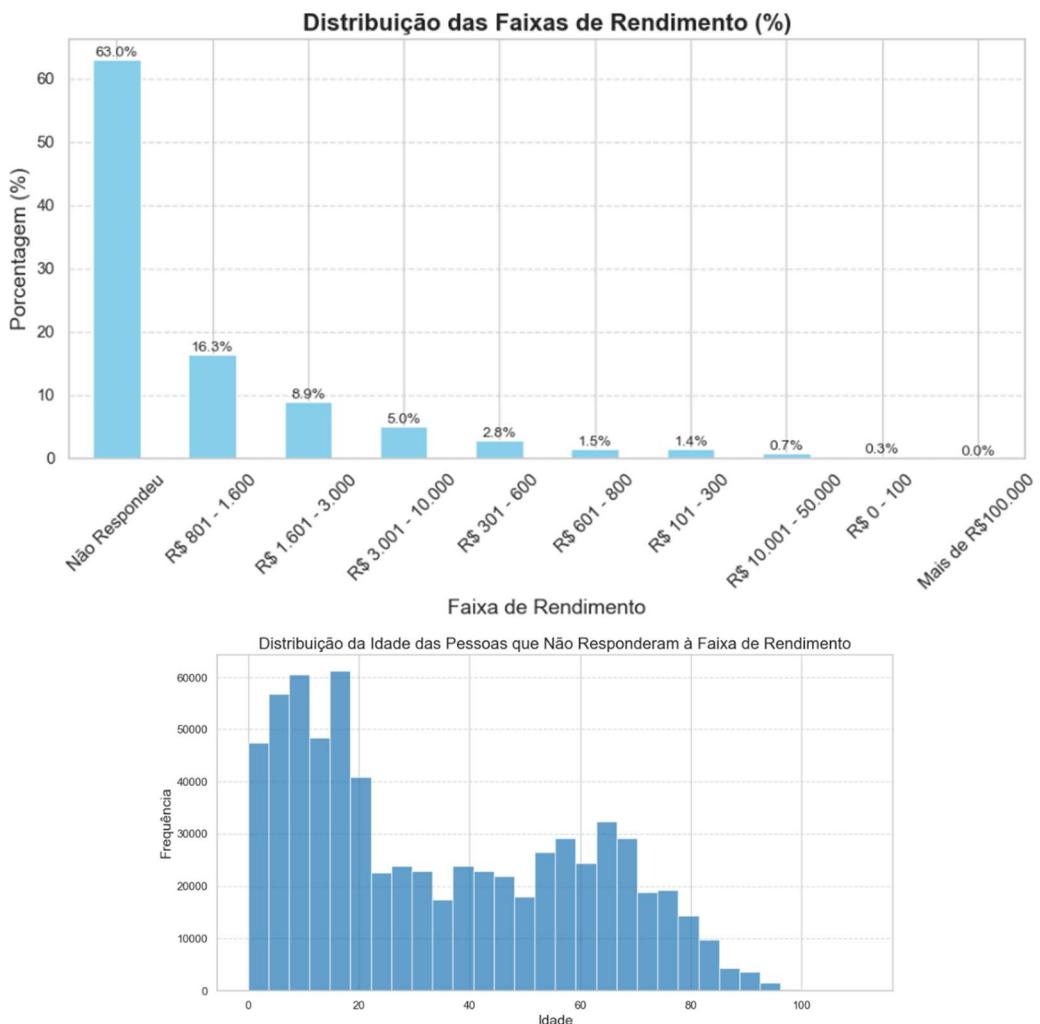
- Cargo ou função, realiza no seu trabalho**

Distribuição de Tipo de Trabalho



Analisando a profissão dos respondentes, pode-se observar que profissões com menor probabilidade de vínculo trabalhistas e com mais riscos a incertezas.

- **Número da faixa do rendimento**



A análise dos dados revela insights significativos sobre a faixa de rendimento e a idade dos indivíduos, especialmente aqueles que não responderam à pesquisa sobre sua situação financeira.

- Aproximadamente 63% dos entrevistados não responderam à pergunta sobre a faixa de rendimento. Essa alta taxa de não resposta levanta questões sobre a representatividade dos dados coletados. É importante considerar se a falta de resposta está relacionada a fatores socioeconômicos, como a falta de entendimento sobre a própria situação financeira ou a relutância em compartilhar informações pessoais.

- A mediana da idade das pessoas que não responderam é de 28 anos, com uma média de 34 anos. O fato de que uma parte significativa desses indivíduos é jovem (com muitos abaixo de 20 anos) sugere que eles podem não ter experiência suficiente com o mercado de trabalho ou compreensão da renda familiar. Essa faixa etária é frequentemente associada a jovens em fase de formação educacional, o que pode resultar em uma percepção distorcida sobre sua própria situação financeira.

- Entre os indivíduos que responderam à pesquisa, a faixa de rendimento que vai de R\$801 a R\$3.000 representa aproximadamente 25% dos respondentes, enquanto a faixa de R\$301 a R\$800 (que é inferior ao salário mínimo) abrange quase 6%. Esses dados sugerem que uma parte considerável da população analisada enfrenta dificuldades financeiras, refletindo uma realidade econômica desafiadora.

- Relação entre Faixa de Rendimento e Idade dos Moradores

```
id_faixa_rendimento
Mais de R$100.000      49.573770
Não Respondeu          34.322787
R$ 0 - 100              41.061264
R$ 1.601 - 3.000        41.791392
R$ 10.001 - 50.000       47.690139
R$ 101 - 300             40.454974
R$ 3.001 - 10.000        44.005169
R$ 301 - 600              39.281476
R$ 601 - 800             39.753680
R$ 801 - 1.600            39.136995
Name: idade_morador, dtype: float64
```

Faixa de Rendimento Alta: "Mais de R\$100.000" (Média de idade: 49.57)

- A média de idade relativamente alta nessa faixa sugere que pessoas com rendimentos mais elevados podem ter alcançado esses níveis de renda após anos de experiência no mercado de trabalho. Isso pode indicar que a riqueza tende a se acumular ao longo do tempo e que é comum que pessoas mais velhas ocupem posições de maior rendimento.

Faixas de Rendimento Baixo: Faixas: "R\$ 0 - 100" (41.06), "R\$ 1.601 - 3.000" (41.79), e "R\$ 301 - 600" (39.28)

- A média de idade nessas faixas é próxima e um pouco mais baixa do que na faixa superior. Isso pode indicar que as pessoas mais jovens estão mais propensas a estarem em faixas de rendimento mais baixas, talvez devido à falta de experiência ou à presença de novos trabalhadores no mercado.

Faixa: "Não Respondeu" (34.32)

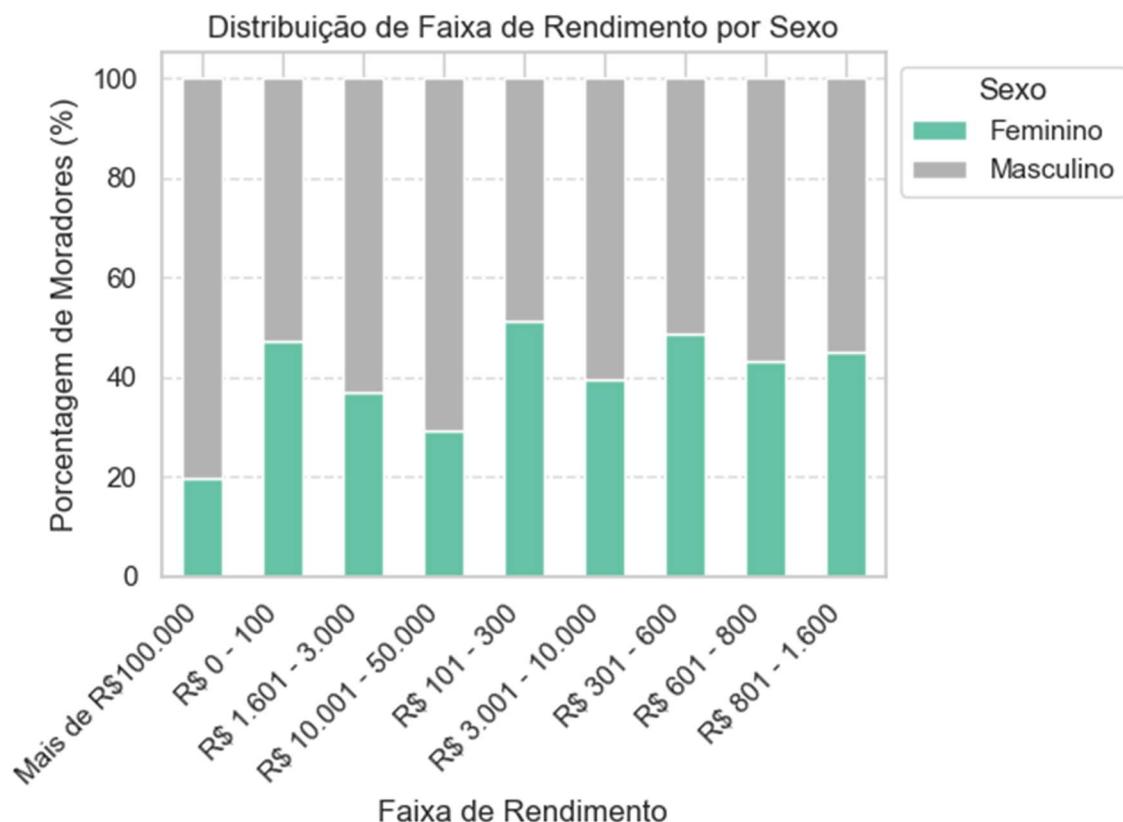
- A média de idade é significativamente mais baixa nesta categoria. Isso pode indicar que pessoas mais jovens são menos propensas a fornecer informações sobre sua situação financeira, ou que podem não ter uma renda definida, como estudantes ou jovens em início de carreira.

Faixas Intermediárias: Faixas: "R\$ 10.001 - 50.000" (47.69), "R\$ 3.001 - 10.000" (44.01), e "R\$ 601 - 800" (39.75)

- A média de idade varia entre essas faixas, sugerindo que a renda intermediária pode ser ocupada por uma variedade de idades. A faixa "R\$ 10.001 - 50.000" tem uma média mais alta, o que pode indicar uma acumulação de rendimento à medida que as pessoas progridem em suas carreiras.

Concluímos que a análise da média de idade por faixa de rendimento sugere que a idade pode ser um fator relevante na renda, com pessoas mais velhas geralmente tendendo a ter rendimentos mais altos. Além disso, a educação e a experiência de trabalho são prováveis determinantes. Esse padrão pode ser explorado em análises mais profundas para entender melhor as desigualdades de renda.

- Relação entre Faixa de Rendimento e Sexo



- Faixa de Rendimento Mais de R\$100.000: Uma porcentagem significativa de 80,33% dos indivíduos nessa faixa são masculinos. Isso pode indicar uma predominância de homens em posições de maior renda, possivelmente refletindo desigualdades de gênero no mercado de trabalho.

- Faixa Não Respondeu: Apesar de ser um valor elevado (57,58% para mulheres), a alta taxa de não resposta sugere que muitos indivíduos, principalmente mulheres, podem estar relutantes em compartilhar informações sobre sua renda. Isso pode refletir questões culturais ou de privacidade.

- Renda Baixa (R\$ 0 - 100 e R\$ 301 - 600): As porcentagens são relativamente equilibradas entre os sexos, mas ainda assim, as mulheres têm uma leve maioria nas faixas de renda muito baixa. Isso pode sugerir que as mulheres estão mais representadas em situações de vulnerabilidade econômica.

- Faixa R\$ 1.601 - 3.000: Essa faixa apresenta uma diferença notável, com 63,23% dos indivíduos sendo masculinos. Isso pode indicar que os homens estão mais propensos a estar em rendas moderadas.

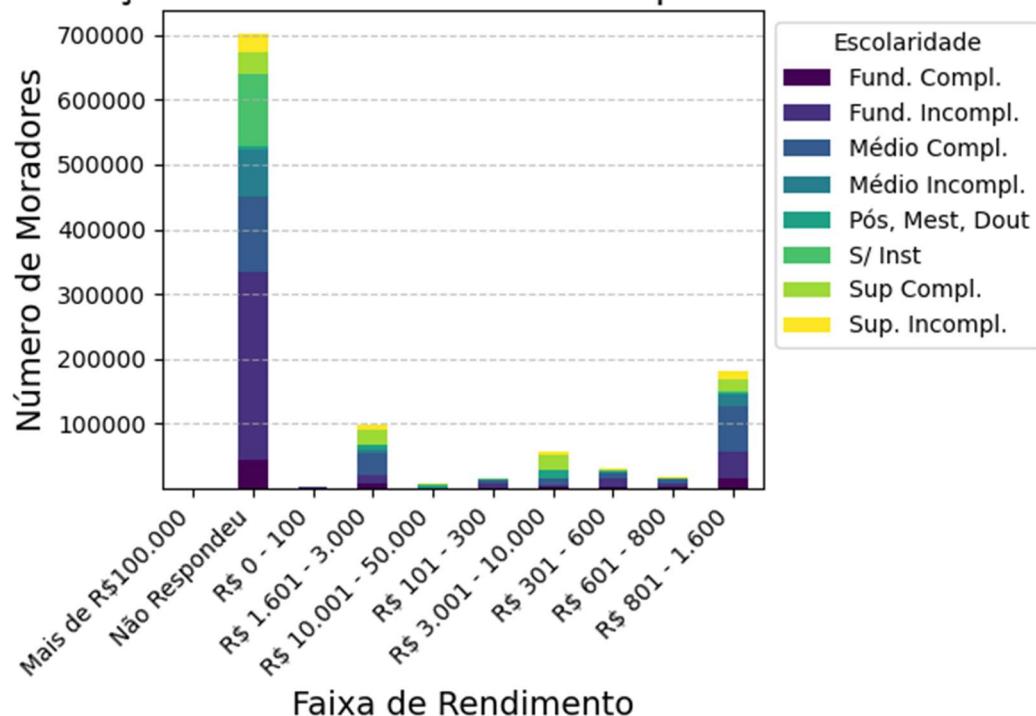
- Faixas de Rendimento R\$ 10.001 - 50.000 e R\$ 3.001 - 10.000: Ambas as faixas mostram uma predominância masculina, o que pode sugerir uma tendência de homens ocupando cargos de maior rendimento e responsabilidade, refletindo desigualdades estruturais no acesso a oportunidades.

- Faixa R\$ 101 - 300 e R\$ 601 - 800: Aqui, os números estão mais equilibrados, mas ainda há uma leve vantagem masculina. Isso pode indicar que, embora ambos os sexos compartilhem experiências de baixa renda, os homens ainda têm uma representação ligeiramente maior.

Conclusões Gerais: As análises sugerem uma desigualdade de gênero clara em relação às rendas mais altas, com uma maior representação masculina nas faixas de renda alta. E a alta taxa de não resposta, especialmente entre mulheres, indica uma possível dificuldade em acessar ou relatar informações de rendimento, o que pode ser um indicador de vulnerabilidade econômica. Esses dados podem ser fundamentais para direcionar políticas públicas focadas na igualdade de gênero no mercado de trabalho e na inclusão econômica de grupos vulneráveis, especialmente mulheres.

- Relação entre Faixa de Rendimento e Escolaridade

Distribuição de Faixa de Rendimento por Escolaridade



O gráfico fornece uma visão abrangente da quantidade de moradores em diferentes faixas de rendimento, segmentada por nível de escolaridade. A seguir, são destacados alguns pontos importantes da análise:

- Alta Taxa de Não Resposta: A categoria "Não Respondeu" é predominante em várias faixas de rendimento, especialmente em "Fund. Incompl." (289.538), indicando uma possível falta de compreensão ou disposição para fornecer informações sobre a renda. Isso pode refletir uma dificuldade de comunicação ou educação financeira entre os indivíduos.

- Educação e Renda: A faixa de rendimento "Mais de R\$100.000" possui apenas 29 registros em "Sup. Compl." (Superior Completo), sugerindo que a alta escolaridade está associada a rendimentos mais elevados. No entanto, o número ainda é relativamente baixo, o que indica que a maioria da população com educação superior não necessariamente alcança esses rendimentos elevados.

- Por outro lado, "Médio Compl." e "Fund. Compl." mostram números consideráveis em faixas de rendimento mais baixas, sugerindo que, mesmo com uma educação formal, muitos ainda estão inseridos em faixas de rendimento limitadas.

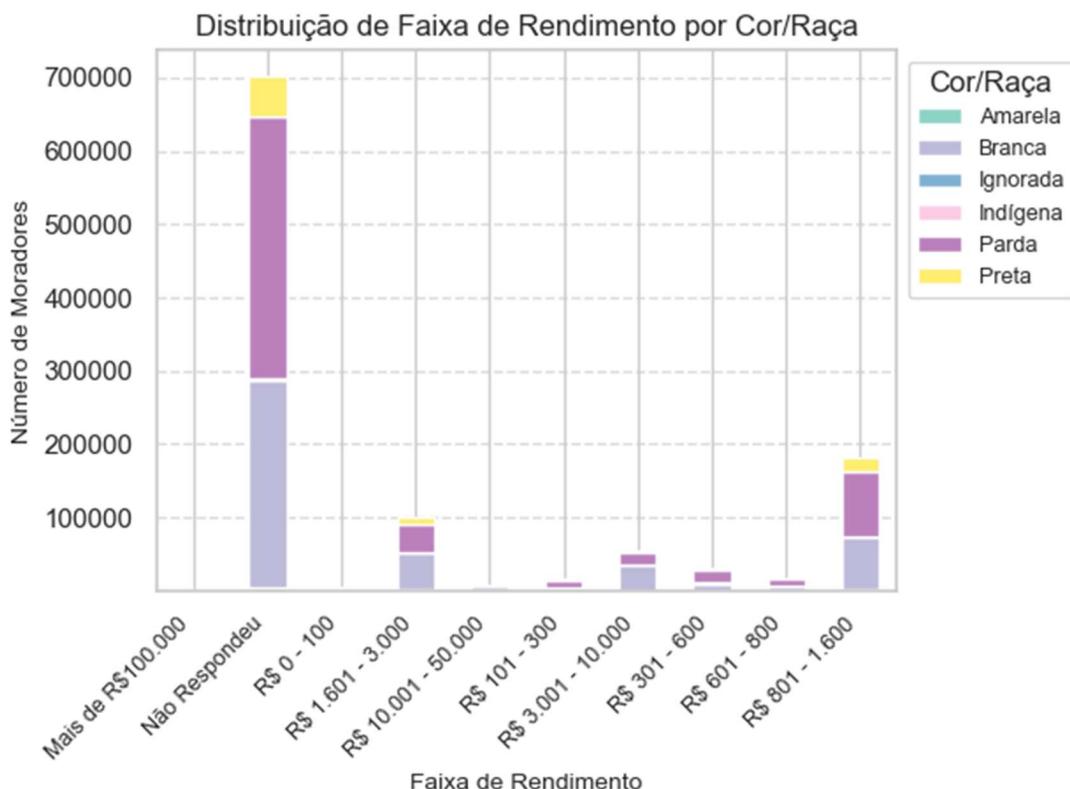
- Rendimentos Baixos: As faixas de rendimento de R\$0 a R\$800 mostraram uma grande quantidade de pessoas com "Fund. Incompl." (1.590 em R\$0-100 e 11.702 em R\$301-600). Isso sugere uma correlação entre a baixa escolaridade e a inserção em faixas de rendimento muito baixas, o que pode indicar a necessidade de intervenções sociais e educacionais.

A faixa de R\$801 a R\$1.600 também tem uma quantidade significativa de moradores com "Médio Compl." (71.113), revelando que muitos com educação média ainda não conseguem alcançar uma estabilidade financeira adequada.

- Desigualdade de Gênero: Embora a tabela não forneça dados diretos sobre gênero, a observação da renda em relação à escolaridade sugere que podem existir desigualdades de gênero nas distribuições de rendimento, especialmente nas faixas mais altas. Análises adicionais que considerem o sexo podem revelar disparidades significativas.

- Educação Pós-Graduada: Para a faixa "Pós, Mest, Dout", os números, embora respeitáveis em algumas faixas, como R\$3.001 a R\$10.000 (11.121), ainda são modestos em comparação com a alta formação acadêmica. Isso pode indicar que a valorização da educação superior não se traduz proporcionalmente em rendimentos mais altos, sugerindo a necessidade de revisão das políticas de valorização profissional.

- **Relação entre Faixa de Rendimento e Cor/Raça**



O grafico fornece informações sobre a quantidade de moradores em diferentes faixas de rendimento, segmentada por cor ou raça. A seguir, são destacados os principais pontos da análise:

- Desigualdade Racial em Faixas de Rendimento: A análise indica uma desigualdade racial clara nas faixas de rendimento. O grupo "Branca" apresenta números altos em todas as categorias de rendimento, enquanto os "Pretos" e "Pardos" estão mais concentrados em faixas de rendimento mais baixas, como "R\$ 0 - 100" (2.244 para Pardos e 377 para Pretos) e "R\$ 301 - 600" (17.677 para Pardos e 3.174 para Pretos). Isso evidencia a necessidade de políticas públicas que abordem a desigualdade econômica racial.

- Concentração em Baixas Faixas de Rendimento: Tanto os "Pretos" quanto os "Pardos" dominam as faixas mais baixas de rendimento. Por exemplo, na faixa de "R\$ 0 - 100", a quantidade de moradores "Pardos" é de 2.244, enquanto que o número de "Branco" é muito menor (973). Isso sugere que as comunidades negras e pardas enfrentam desafios econômicos mais significativos.

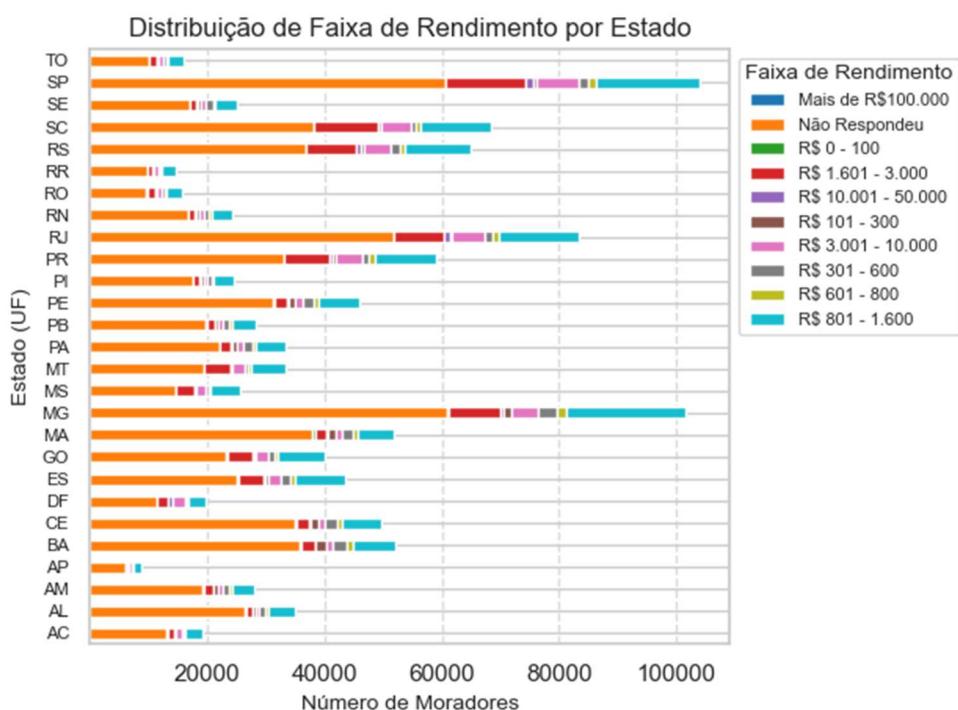
- Distribuição em Faixas de Rendimento Altas: Em contraste, a faixa "Mais de R\$100.000" apresenta números muito baixos em todos os grupos raciais, com apenas 1 para "Amarela", 42 para "Branca", 7 para "Parda" e 11 para "Preta". Isso sugere que, independentemente da cor/race, a elite econômica é muito reduzida, indicando uma concentração de riqueza nas mãos de poucos.

- Evidências de Exclusão: A faixa de "R\$ 10.001 - 50.000" tem apenas 1.771 "Pardos" e 295 "Pretos", o que reflete uma exclusão significativa de grupos raciais em posições econômicas intermediárias. Essa situação pode estar ligada a fatores estruturais, como acesso desigual à educação e oportunidades de emprego.

- População Indígena: A quantidade de pessoas da raça "Indígena" é visivelmente menor em comparação aos outros grupos, mas sua presença em várias faixas de rendimento, incluindo "Não Respondeu" (2.872) e "R\$ 0 - 100" (26), destaca uma população que também pode enfrentar desafios específicos.

Conclusão: A análise evidencia desigualdades significativas nas condições socioeconômicas da população, com uma clara divisão racial nas faixas de rendimento. A predominância de respostas não fornecidas indica uma necessidade de aprimorar a coleta de dados e a confiança da população em compartilhar informações. Para enfrentar essas desigualdades, recomenda-se que as políticas públicas sejam direcionadas para melhorar a educação, e a inclusão econômica dos grupos racialmente marginalizados.

- **Relação entre Faixa de Rendimento e Localização (Estado/UF)**



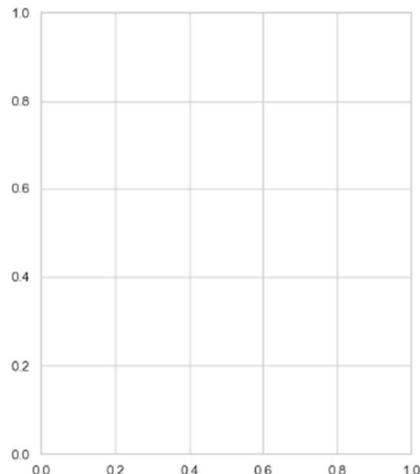
O gráfico apresentado mostra a distribuição da faixa de rendimento das pessoas por estado (UF). Algumas observações relevantes:

- Não Responderam: Uma grande parte da população não respondeu à questão de faixa de rendimento, o que pode indicar uma falta de confiança ou interesse em compartilhar essa informação. Em muitos estados, essa categoria representa a maior parte dos moradores.

- Faixas de Rendimento: As faixas de rendimento mais baixas (como R\$ 0 - 100 e R\$ 101 - 300) têm uma quantidade significativa de pessoas, principalmente em estados como BA, CE e MA, indicando uma realidade econômica desafiadora em algumas regiões.

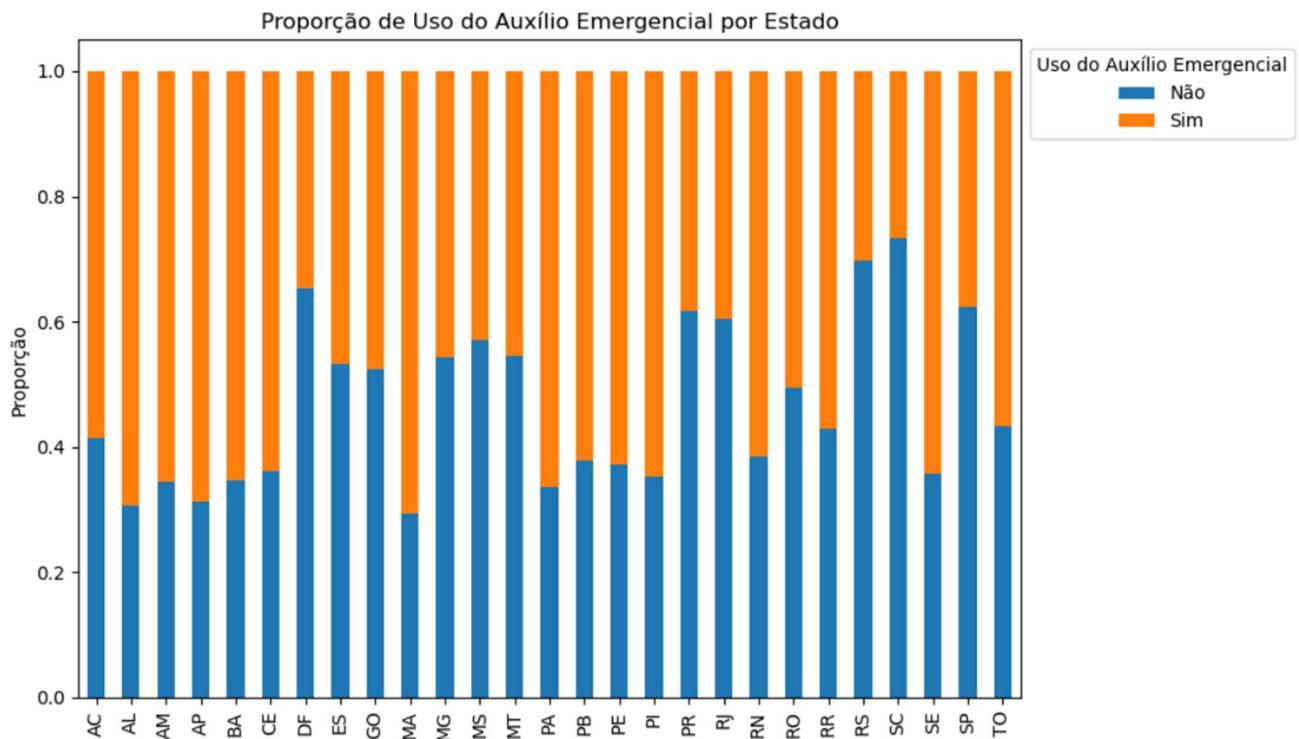
- Rendimento Alto: A faixa "Mais de R\$100.000" é bem menos representativa, com poucos registros em todos os estados, o que sugere que altos rendimentos são uma exceção na população analisada.

Distribuição Regional: É notável que os estados do Nordeste, como AL, BA e CE, apresentam números elevados de pessoas na categoria "Não Respondeu", bem como em faixas de rendimento muito baixas, o que pode refletir as desigualdades regionais no Brasil.



5. Análise Exploratória: Rendimento de outras fontes dos moradores

- Análise dos Auxílios emergenciais relacionados ao coronavírus**



- Variação Siginificativa por Estado:

Observa-se uma variação significativa na proporção de pessoas que utilizaram o auxílio emergencial entre os estados. Alguns estados, como AC (Acre), AL (Alagoas) e AP (Amapá), têm uma alta proporção de pessoas que usaram o auxílio (representado pela cor laranja). Já em outros estados, como DF (Distrito Federal), RS (Rio Grande do Sul) e SC (Santa Catarina), a maioria das pessoas não usou o auxílio (representado pela cor azul).

- Regiões Norte e Nordeste com Maior Uso:

Em estados do Norte e Nordeste (ex.: AC, AL, AM, CE), há uma clara predominância de pessoas que utilizaram o auxílio emergencial, sugerindo que essas regiões podem ter enfrentado uma maior dependência desse benefício durante o período analisado.

- Menor Dependência no Sul e Sudeste:

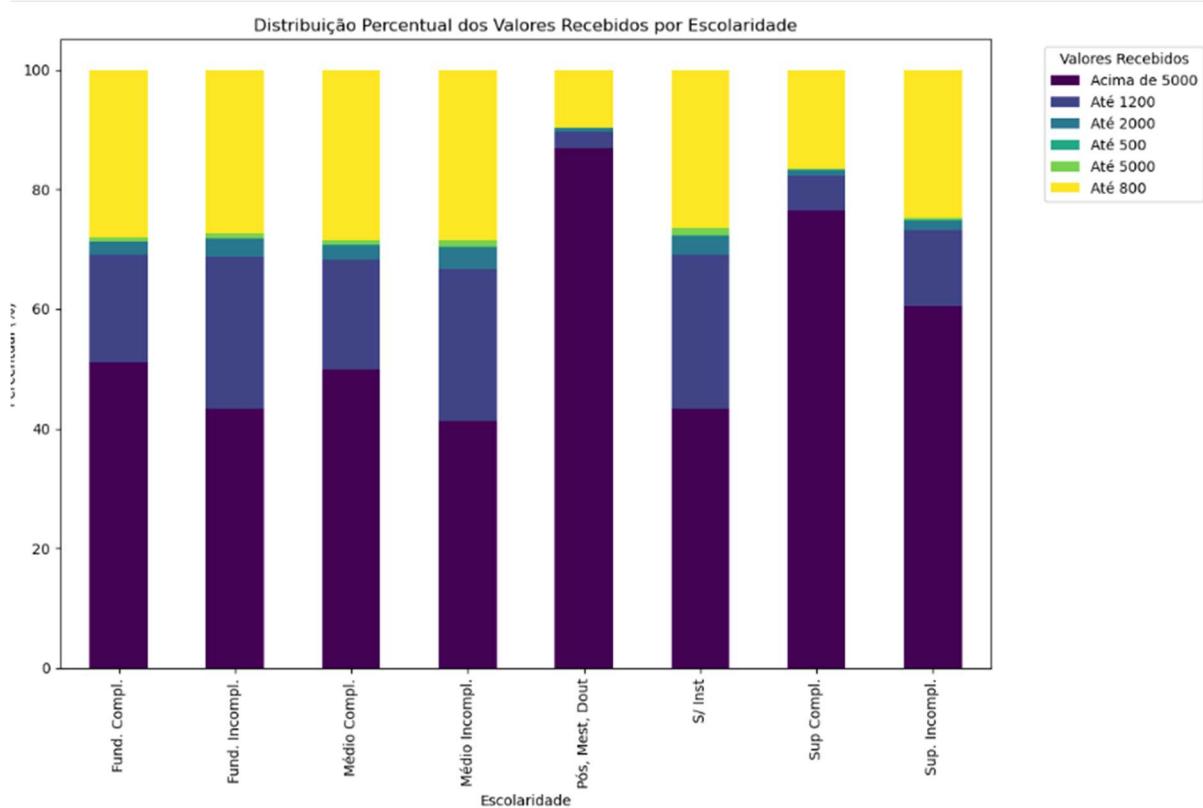
No Sul e Sudeste (ex.: SP, RS, SC), observa-se uma maior proporção de pessoas que não utilizaram o auxílio emergencial. Isso pode estar relacionado a fatores econômicos regionais, como a presença de mais oportunidades de emprego formal ou melhores condições socioeconômicas.

- Análise de Estados Específicos:

O Distrito Federal (DF) se destaca como o estado com a maior proporção de pessoas que não utilizaram o auxílio emergencial, o que pode ser explicado pelo perfil econômico da capital federal, com altos índices de servidores públicos e trabalhadores formais. Em contraste, estados como MA (Maranhão) e PI (Piauí) têm uma alta proporção de pessoas que dependiam do auxílio.

A análise sobre este tópico sugere que o impacto econômico da pandemia foi mais severo em certas regiões, principalmente no Norte e Nordeste do Brasil, onde o auxílio emergencial desempenhou um papel crucial no sustento das famílias. Em outras regiões, como no Sul e Sudeste, as proporções indicam uma menor dependência desse benefício.

- Análise da somatória de valores recebidos por Nível de Escolaridade**



- Distribuição Percentual dos Valores Recebidos por Escolaridade:**

Análise do Impacto do Nível de Escolaridade:

Fundamental Completo: A maior parte dos valores recebidos está na faixa “Acima de 5000” (aproximadamente 51%), seguida por “Até 800” (28%) e “Até 1200” (18%).

Fundamental Incompleto: Similar ao Fundamental Completo, a maior parte dos valores está na faixa “Acima de 5000” (43%), com uma distribuição significativa também em “Até 800” (27%) e “Até 1200” (25%).

Médio Completo: A distribuição é bastante similar ao Fundamental Completo, com “Acima de 5000” representando cerca de 50% e “Até 800” com 28%. Médio Incompleto: A maior parte dos valores está na faixa “Acima de 5000” (41%), seguida por “Até 800” (29%) e “Até 1200” (25%).

Superior Completo: A maior parte dos valores recebidos está na faixa “Acima de 5000” (76%), indicando uma tendência de maiores valores recebidos com maior nível de escolaridade.

Superior Incompleto: A distribuição é mais equilibrada, com “Acima de 5000” representando 60% e “Até 1200” com 13%.

Pós-graduação Completa: A maior parte dos valores está na faixa “Acima de 5000” (86%), indicando que indivíduos com pós-graduação tendem a receber valores mais altos.

Pós-graduação Incompleta: A distribuição é similar, com “Acima de 5000” representando a maior parte (86%).

Sem Instrução: A maior parte dos valores recebidos está na faixa “Acima de 5000” (43%), seguida por “Até 800” (26%) e “Até 1200” (25%).

Tendência Geral:

Há uma clara tendência de que níveis mais altos de escolaridade estão associados a uma maior proporção de valores recebidos na faixa “Acima de 5000”.

Distribuição Equilibrada:

Níveis de escolaridade mais baixos (Fundamental e Médio) mostram uma distribuição mais equilibrada entre as faixas de valores recebidos.

Impacto da Escolaridade:

A escolaridade parece ter um impacto significativo na distribuição dos valores recebidos, com níveis mais altos de escolaridade correlacionando-se com valores recebidos mais altos.

- **Distribuição Percentual dos Valores Recebidos por Escolaridade**

Análise do Impacto do Nível de Escolaridade:

Fundamental Completo: A maior parte dos valores recebidos está na faixa “Acima de 5000” (aproximadamente 51%), seguida por “Até 800” (28%) e “Até 1200” (18%).

Fundamental Incompleto: Similar ao Fundamental Completo, a maior parte dos valores está na faixa “Acima de 5000” (43%), com uma distribuição significativa também em “Até 800” (27%) e “Até 1200” (25%).

Médio Completo: A distribuição é bastante similar ao Fundamental Completo, com “Acima de 5000” representando cerca de 50% e “Até 800” com 28%. Médio Incompleto: A maior parte dos valores está na faixa “Acima de 5000” (41%), seguida por “Até 800” (29%) e “Até 1200” (25%).

Superior Completo: A maior parte dos valores recebidos está na faixa “Acima de 5000” (76%), indicando uma tendência de maiores valores recebidos com maior nível de escolaridade.

Superior Incompleto: A distribuição é mais equilibrada, com “Acima de 5000” representando 60% e “Até 1200” com 13%.

Pós-graduação Completa: A maior parte dos valores está na faixa “Acima de 5000” (86%), indicando que indivíduos com pós-graduação tendem a receber valores mais altos.

Pós-graduação Incompleta: A distribuição é similar, com “Acima de 5000” representando a maior parte (86%).

Sem Instrução: A maior parte dos valores recebidos está na faixa “Acima de 5000” (43%), seguida por “Até 800” (26%) e “Até 1200” (25%).

Tendência Geral:

Há uma clara tendência de que níveis mais altos de escolaridade estão associados a uma maior proporção de valores recebidos na faixa “Acima de 5000”.

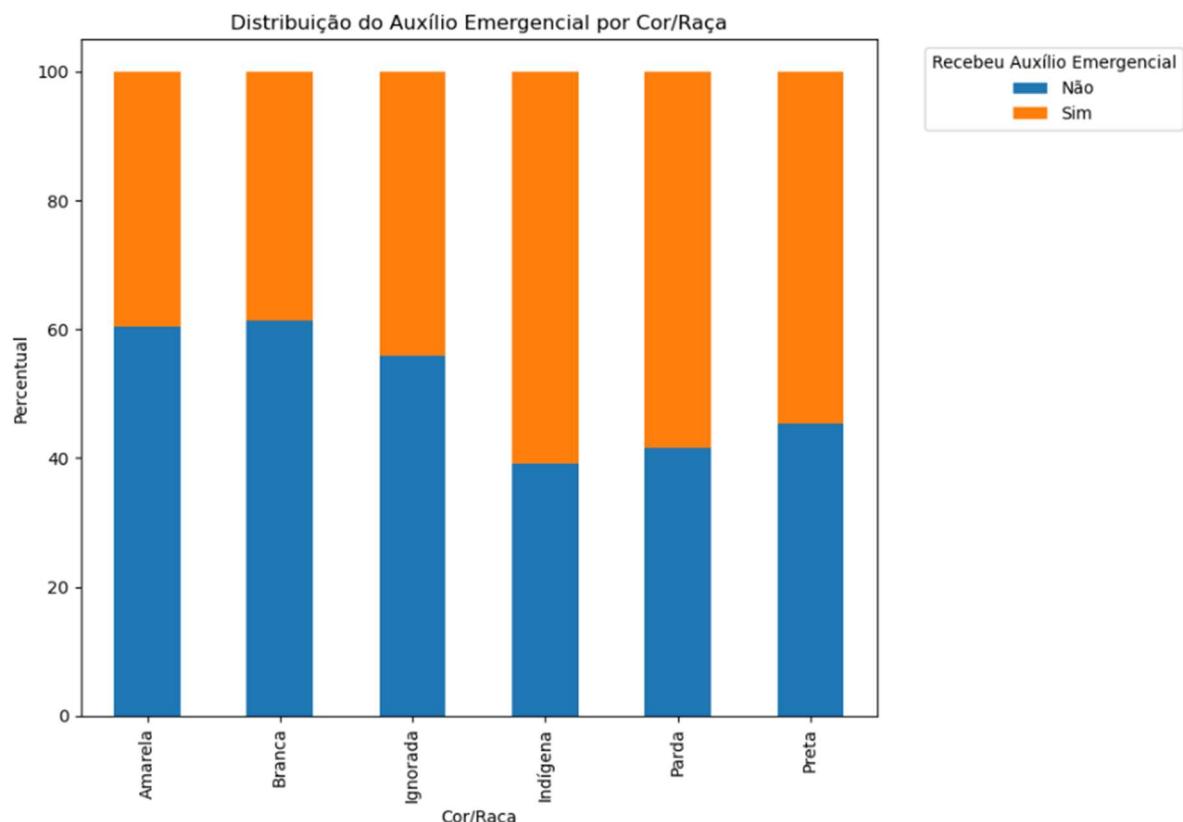
Distribuição Equilibrada:

Níveis de escolaridade mais baixos (Fundamental e Médio) mostram uma distribuição mais equilibrada entre as faixas de valores recebidos.

Impacto da Escolaridade:

A escolaridade parece ter um impacto significativo na distribuição dos valores recebidos, com níveis mais altos de escolaridade correlacionando-se com valores recebidos mais altos.

- **Análise da somatória de valores recebidos por Faixa Etária e Cor/Raça**



- Distribuição do Auxílio Emergencial por cor de raça:

Amarela:

Aproximadamente 60% das pessoas não receberam auxílio emergencial..

Branca:

A distribuição é semelhante à da cor amarela, com cerca de 61%.

Indígena:

A maioria das pessoas indígenas (cerca de 61%) recebeu auxílio emergencial.

Parda:

Cerca de 58% das pessoas pardas receberam auxílio.

Preta:

Aproximadamente 55% das pessoas pretas receberam auxílio.

Observações Gerais:

As categorias “Indígena” e “Parda” têm as maiores porcentagens de recebimento de auxílio emergencial, com 61% e 58%, respectivamente.

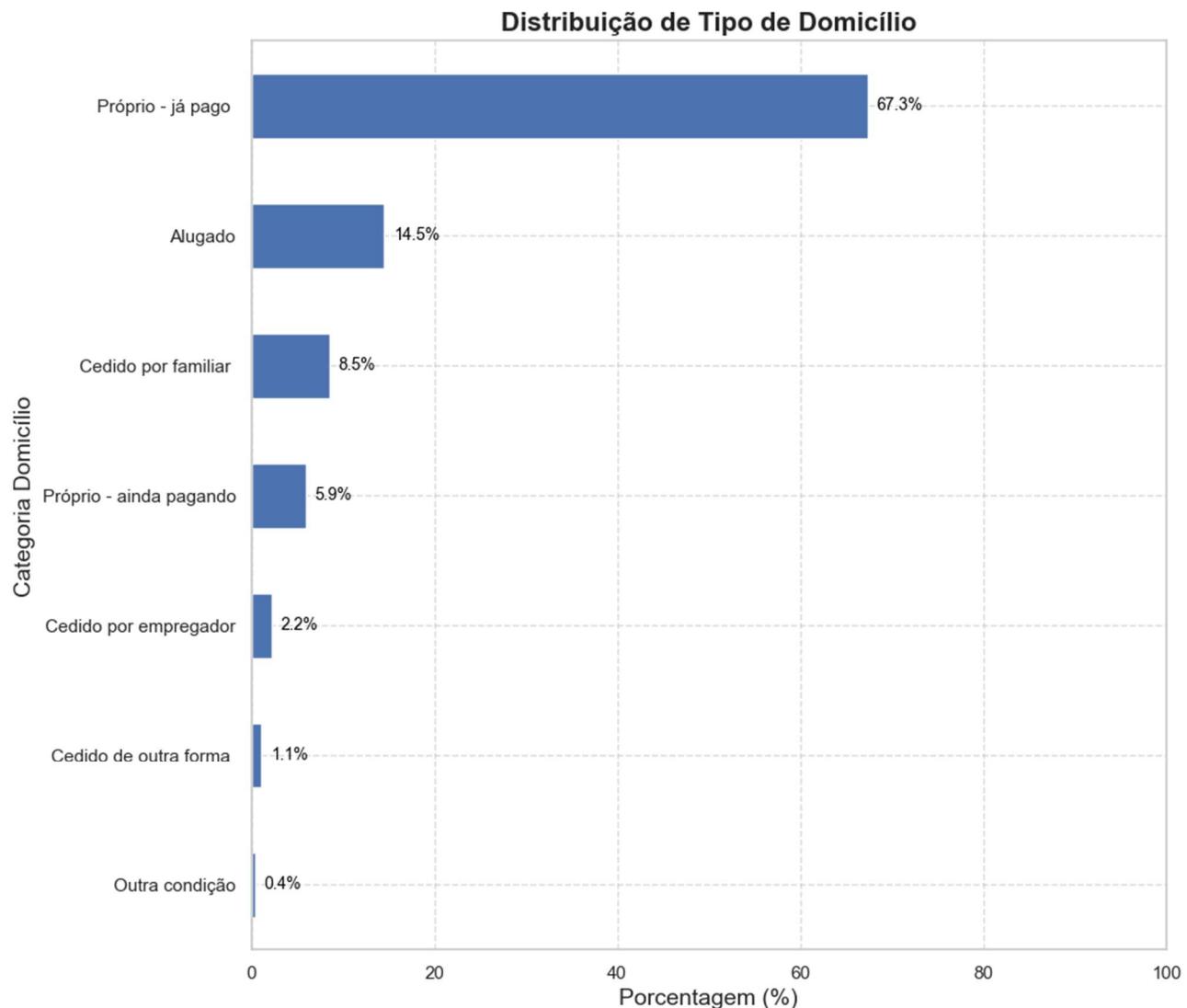
As categorias “Amarela” e “Branca” têm as menores porcentagens de recebimento de auxílio, com cerca de 40% e 39%, respectivamente.

Desigualdade Racial:

A análise sugere que há uma variação significativa na distribuição do auxílio emergencial entre diferentes grupos raciais. Grupos como os indígenas e pardos parecem ter uma maior necessidade ou acesso ao auxílio emergencial.

6. Análise Exploratória: Características da habitação

- **Análise Domiciliar**



Analizando o tipo de moradia, entre os respondentes mais de 65% tem casa própria quitada, o que traz segurança para grande parcela da amostra.

Porém ao analisar moradia própria mais ainda pagando e pessoas que moram de aluguel temos uma fatia de 20,4%, que preocupa em momento de incerteza de pandemias.

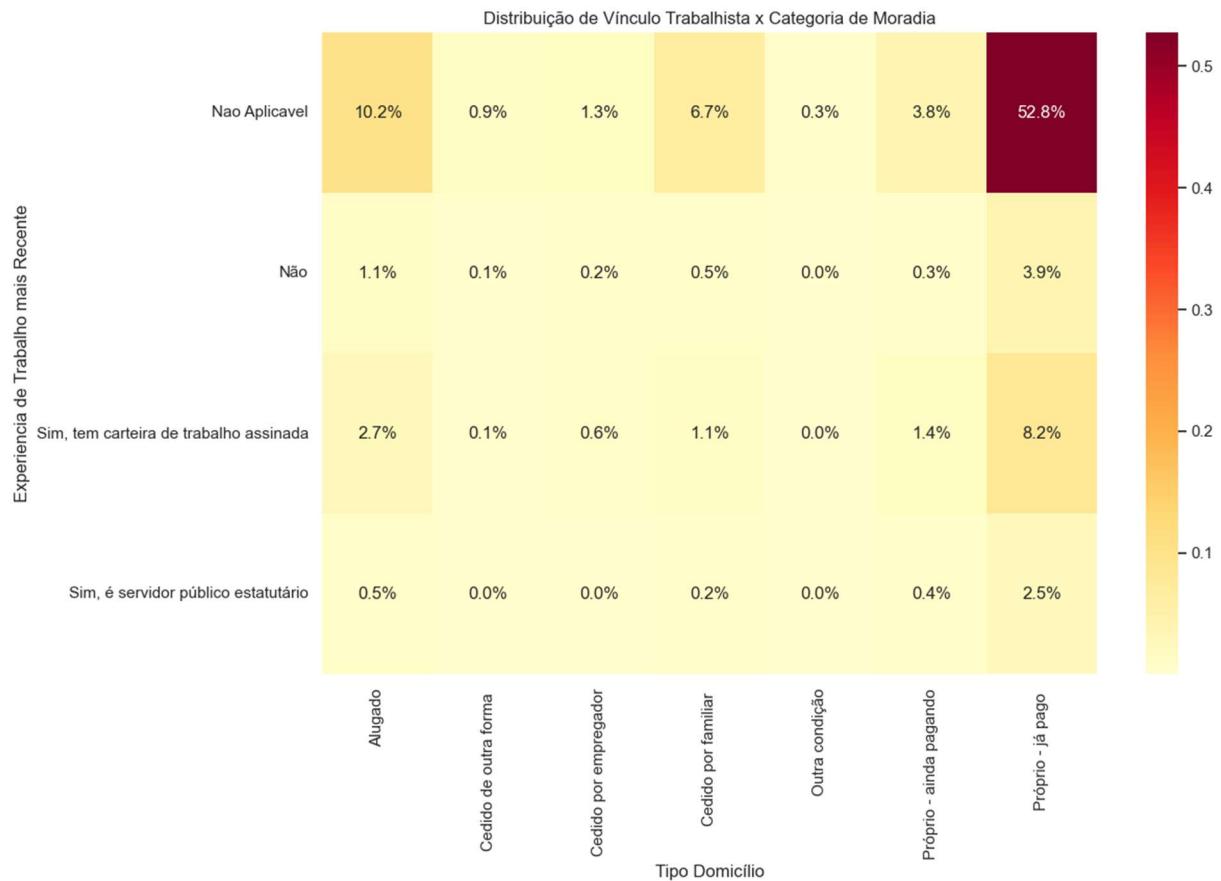
Cruzando informações de pessoas que optaram por não responder:

- 53% tem moradia própria paga;
- 10% mora de aluguel;
- 7% moram com familiares

Analizando pessoas que trabalham de carteira assinada:

- 8% tem moradia própria
- 3% mora de aluguel

Pode-se concluir que apesar de uma grande amostra ter casa própria, temos um percentual relevante de pessoas que podem ser impactadas com efeitos da pandemia.



7. Conclusão

Conforme as análises, a pandemia da COVID-19 evidenciou diversos aspectos críticos do comportamento populacional, da saúde e da economia do Brasil. Os principais pontos a serem destacados incluem:

1. Perfil Demográfico e Socioeconômico:

- As características demográficas indicam que os grupos raciais Pardo e Branco têm maior representação em praticamente todos os níveis de escolaridade, mas o ensino fundamental incompleto ainda é predominante.
- As faixas de menor escolaridade estão associadas a condições econômicas mais precárias, evidenciando desigualdades socioeconômicas.

2. Sintomas da COVID-19:

- A maioria da população não apresentou sintomas clássicos, como febre e tosse. Sintomas como dor de cabeça, muscular e perda de cheiro ou sabor foram relatados em menor frequência.
- Observou-se uma redução gradual na prevalência de sintomas de maio a julho de 2020, indicando um possível declínio nos casos ou adaptação ao contexto pandêmico.

3. Impacto por Faixa Etária e Região:

- Adultos de 19-60 anos foram os mais afetados em termos de sintomas relatados, enquanto a população idosa apresentou uma maior taxa de internação e necessidade de ventilação mecânica.
- As regiões Norte e Nordeste registraram maiores prevalências de sintomas, internações e uso do auxílio emergencial, reforçando a vulnerabilidade dessas áreas durante a pandemia.

4. Desigualdades Sociais e Econômicas:

- A análise da renda por faixa etária, gênero e raça evidenciou disparidades significativas. Homens tendem a ocupar as faixas de maior rendimento, enquanto mulheres e grupos raciais como Pretos e Pardos se concentram nas faixas de menor rendimento.
- A dependência do auxílio emergencial foi maior em regiões mais carentes e entre indivíduos com menor escolaridade, evidenciando a importância desse benefício em momentos de crise.

5. Recomendações:

- É crucial que políticas públicas sejam direcionadas para a inclusão econômica e social de grupos mais vulneráveis, com foco em educação, saúde e geração de empregos.
- O fortalecimento do sistema de saúde, principalmente em regiões com maior incidência de sintomas e internações, será essencial para lidar com futuros surtos.

Essas análises destacam fragilidades e desigualdades presentes na sociedade brasileira, apontando para a necessidade de uma abordagem multidisciplinar e inclusiva no enfrentamento de futuras crises sanitárias e socioeconômicas.