

PNAD COVID

Pós Tech Data Analytics Turma 9DTAT - Fiap

Grupo 205

Geovana dos Santos Ferreira - rm364998

Outubro/2025

Introdução	3
Arquitetura dos dados	4
Etapas realizadas:	4
Variáveis analisadas	5
Caracterização dos sintomas clínicos da população	5
Comportamento da População na Pandemia	8
Características Sócio-Demográficas	11
Conclusão Geral	14

Introdução

A pandemia de COVID-19 trouxe desafios sem precedentes para a saúde pública em todo o mundo. No Brasil, a rápida disseminação do vírus destacou a importância de decisões estratégicas baseadas em dados para minimizar impactos sobre a população e o sistema de saúde.

Nesse contexto, o presente projeto foi desenvolvido como parte do Tech Challenge, com o objetivo de aplicar conhecimentos adquiridos ao longo da fase em Data Analytics para analisar dados populacionais e gerar insights para planejamento hospitalar.

O hospital, que contrata este projeto, busca entender três dimensões principais do comportamento da população durante a pandemia:

1. Sintomas clínicos: identificação dos sintomas mais frequentes, distribuição entre diferentes grupos da população e evolução ao longo do tempo.
2. Comportamento social: adesão às medidas preventivas, isolamento social, afastamento do trabalho e padrões de circulação da população.
3. Aspectos econômicos: análise de renda, ocupação e vulnerabilidades socioeconômicas que possam afetar o risco e o acesso a cuidados de saúde.

Para garantir que a análise fosse estratégica e focada nas necessidades do hospital, foram selecionados no máximo 20 questionamentos-chave da pesquisa PNAD-COVID-19, priorizando variáveis que permitem observar sintomas clínicos, comportamento social e características econômicas da população.

Além disso, a análise considerou três meses representativos da pesquisa, permitindo avaliar tendências, padrões temporais e comportamentos emergentes durante a pandemia.

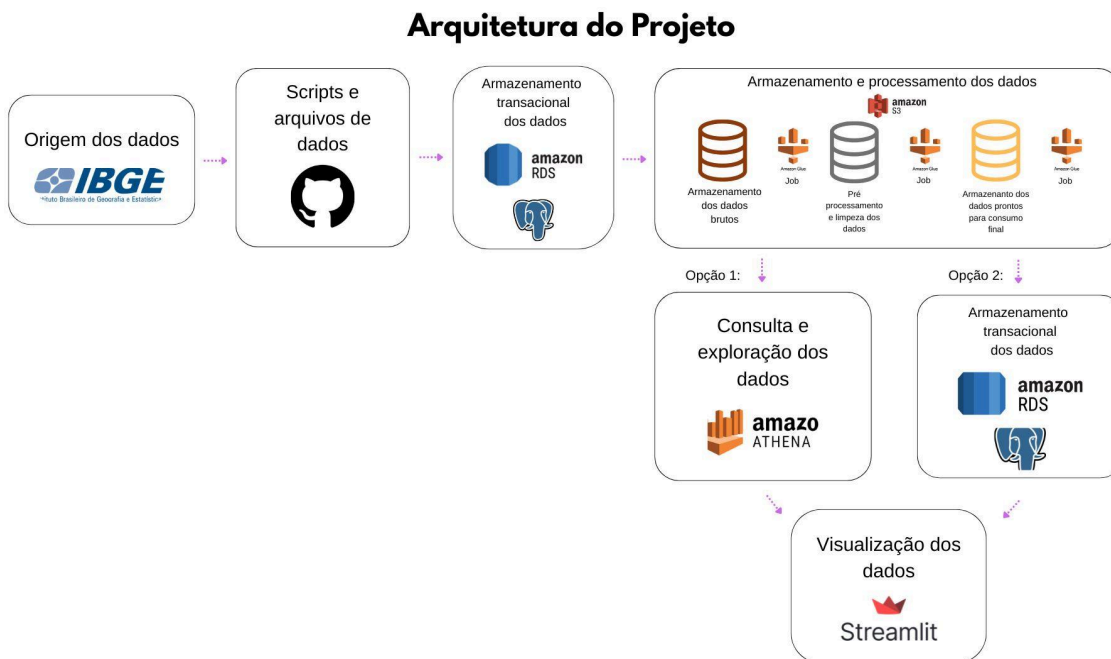
O relatório que se segue detalha a metodologia de coleta e organização dos dados, análise exploratória, visualizações, modelagem preditiva e recomendações estratégicas, fornecendo um panorama completo que serve de base para decisões futuras do hospital em caso de novos surtos de COVID-19.

Arquitetura dos dados

Fonte dos dados: PNAD-COVID19 IBGE.

Período analisado: De Setembro à Novembro de 2020 (3 meses)

Ferramentas utilizada: Github, VsCode, AWS RDS (Postgresql), AWS S3, AWS Glue, Python, SQL, Streamlit



Etapas realizadas:

1. Identificação dos arquivos necessários para utilização.
2. Upload dos arquivos no repositório do projeto no Github.
3. Criação do banco de dado na AWS, buckets no s3 e parametrizações no Glue.
4. Validações das conexões acima.
5. Leitura e iteração dos arquivos no Git via requests (últimos 3 meses).
6. Criação das subpastas (Bronze, Silver e Gold) no s3.
7. Ingestão dos arquivos do git no RDS.
8. Conexão e ingestão dos dados do RDS para a camada Bronze no s3.
9. Conexão e ingestão dos dados da Bronze para a Silver no s3 (Seleção das variáveis).
10. Conexão e ingestão dos dados da Silver para a Gold no s3 (Enriquecimento e Transformações aplicadas nas variáveis).
11. Opção 1: Ingestão da tabela da pasta Gold para o RDS (Análise dos dados por RDS e publicação de dataviz via streamlit - Disponibilidade 24/7) - **OPÇÃO ESCOLHIDA**.
12. Opção 2: Criação do banco de dados e crawlers no Glue (Análise dos dados pelo Athena e publicação de dataviz via streamlit - Disponibilidade apenas quando o console AWS estivesse conectado).

Variáveis analisadas

- Características clínicas dos sintomas;
- Características da população;
- Características da sociedade;

A partir dessas condições, trouxemos uma breve análise dessas informações, explicitando o modo como organizamos o banco de dados, o porquê da seleção das melhores perguntas que trariam as melhores respostas e, por fim, as ações mais efetivas que o hospital deverá tomar em caso de um novo surto de COVID-19.

Caracterização dos sintomas clínicos da população

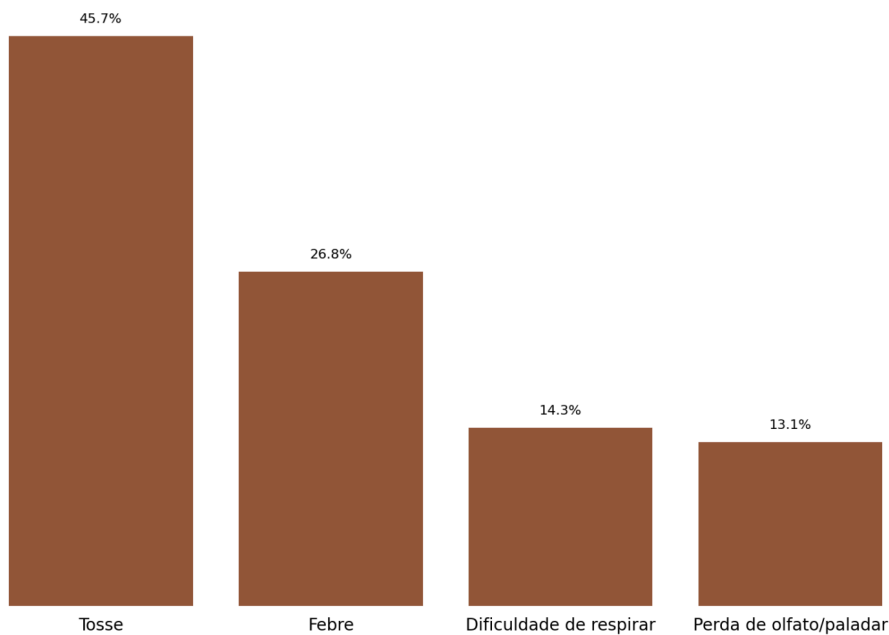
Sintomas, diagnóstico e tratamento:

A maioria dos entrevistados com sintomas relatou tosse (45,7%) e febre (28,8%), enquanto sintomas como dificuldade de respirar e perda de olfato/paladar foram menos frequentes. Mesmo com esses sintomas, a maior parte das pessoas não buscou auxílio médico. Dos que procuraram atendimento e foram diagnosticados, o teste de Swab foi o mais utilizado para confirmação (47,6%), enquanto o de sangue do dedo foi o menos comum (24,2%). A necessidade de internação (2,2%) ou sedação (0,5%) foi a minoria dos casos.

Há uma clara discrepância entre a ocorrência de sintomas e a busca por ajuda profissional. A maioria da população se auto diagnosticou e tratou os sintomas em casa, com febre e tosse sendo as manifestações mais comuns. A baixa taxa de internação e sedação sugere que os casos foram majoritariamente leves ou moderados, ou que a busca por tratamento ocorreu em estágios avançados, o que não é possível saber com os dados.

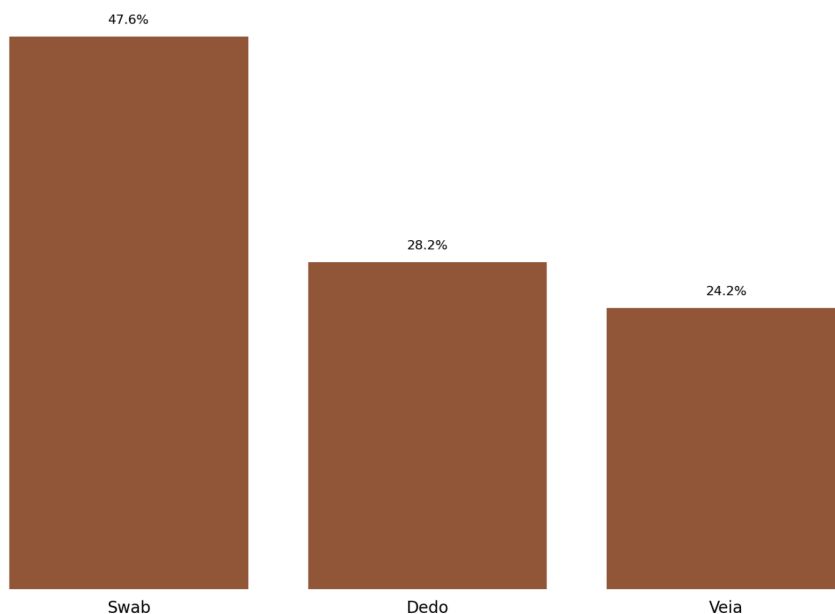
Veja as análises gráficas abaixo.

Distribuição dos sintomas mais comuns (em %)

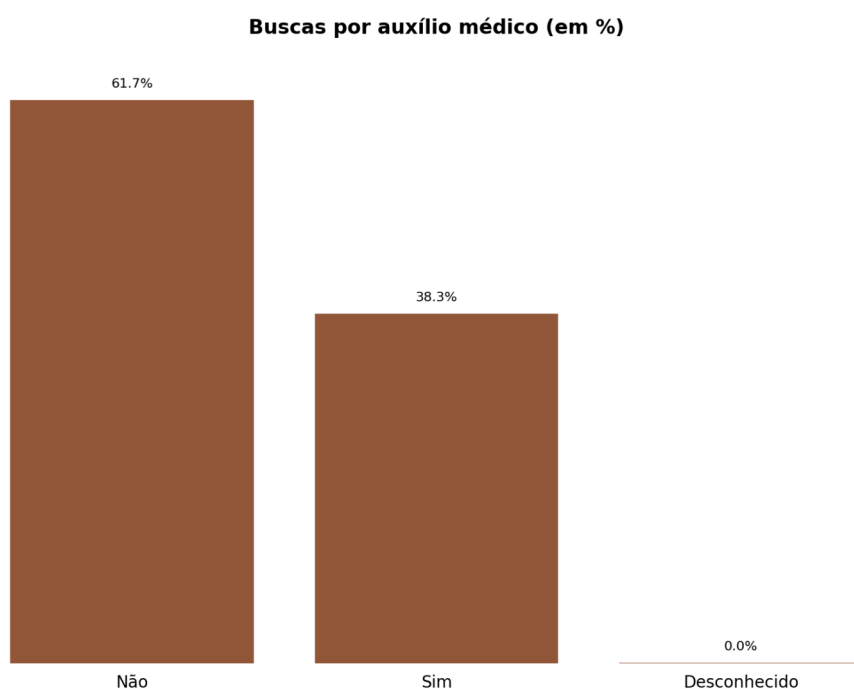


Tosse e febre foram os sintomas predominantes entre os entrevistados que declararam ter sentido algum tipo de sintoma. Isso sugere que, na época da pesquisa (2020), esses eram os sintomas mais evidentes ou facilmente identificáveis pela população.

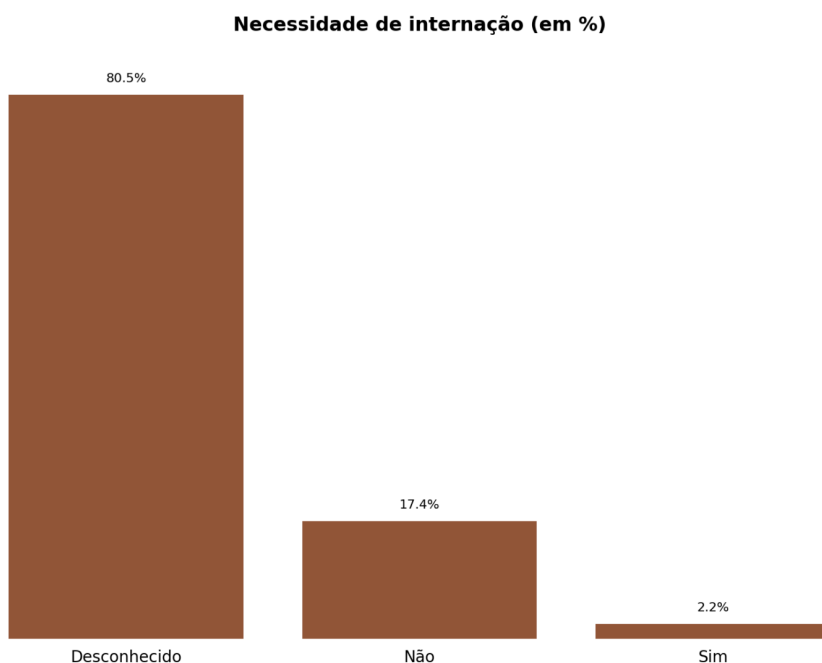
Resultados positivos por tipo de teste (em %)



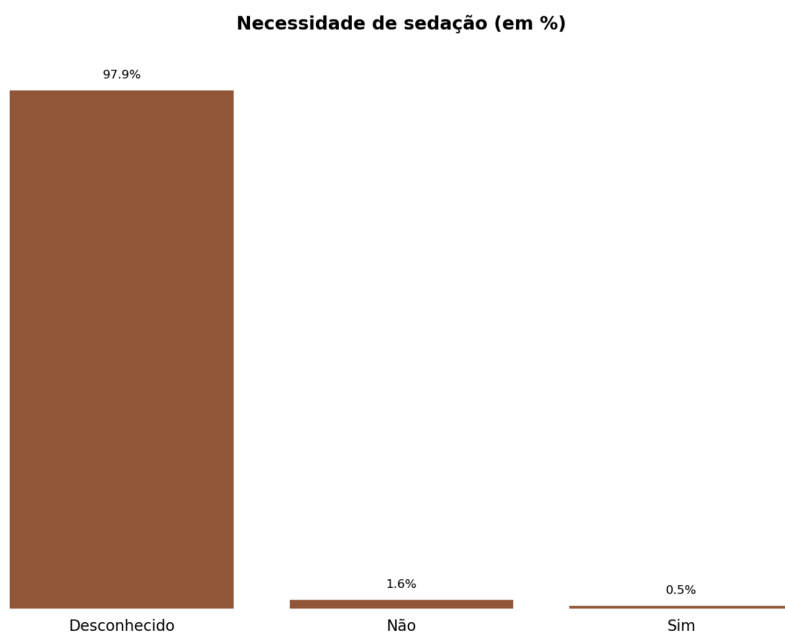
O teste de Swab foi o método mais comum para detectar casos positivos, o que era consistente com as práticas de testagem da época. A menor quantidade de positivos nos testes rápidos (dedo e veia) pode refletir tanto uma menor frequência de uso quanto uma menor sensibilidade desses testes na fase inicial da doença.



Uma parcela significativa da população com sintomas não procurou atendimento médico. Isso pode ter ocorrido por diversos motivos, como medo, dificuldade de acesso aos serviços de saúde, ou por considerarem os sintomas leves o suficiente para serem tratados em casa.



Embora uma grande parte da população tenha sentido sintomas, a doença não evoluiu para um quadro grave que exigisse hospitalização para a maioria dos casos.



A necessidade de sedação (normalmente associada a casos graves de UTI) foi um evento raro, reforçando a conclusão de que a maioria dos casos da doença não progrediu para um estado crítico.

Comportamento da População na Pandemia

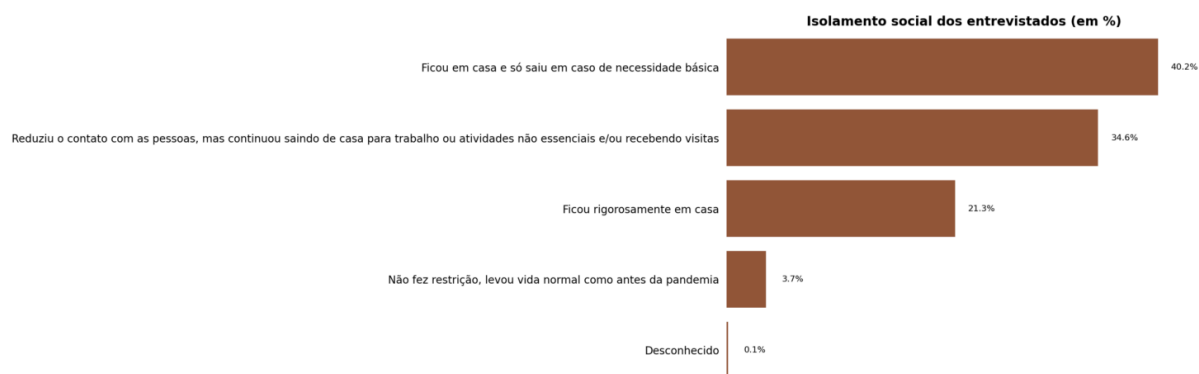
Isolamento social e Renda:

Apenas 40,2% da população com sintomas ficou rigorosamente em casa. A maioria dos entrevistados residia em casa própria (52,2%), o que sugere um certo nível de estabilidade, mas não impediu a exposição.

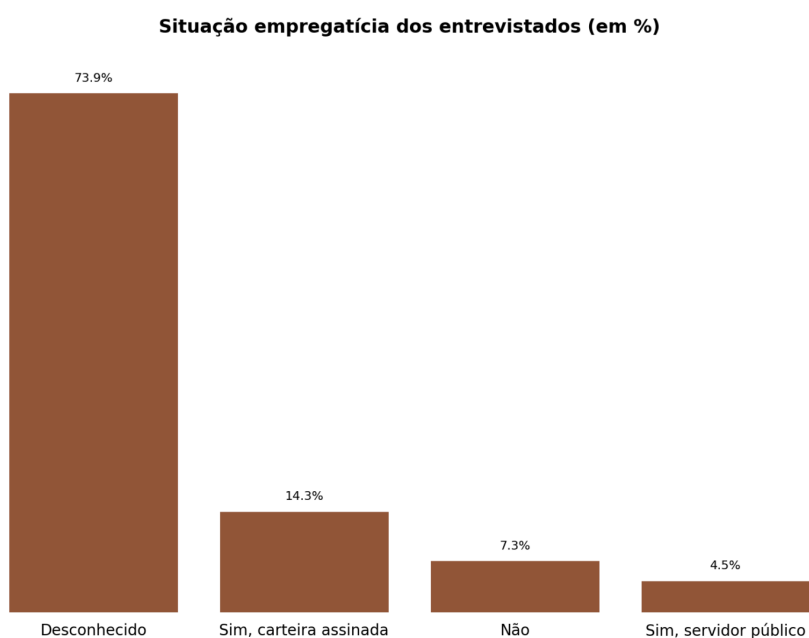
A faixa de renda mais comum entre os entrevistados com sintomas desconsiderando as respostas desconhecidas, é a de 800 - 1.600 (17,1%), que seria classificada como renda baixa, o que justifica os 50,6% da população que reportou ter recebido auxílio emergencial, mesmo que maioria dos entrevistados esteja trabalhando atualmente.

O isolamento social rigoroso não foi uma opção para a maioria dos entrevistados. A baixa escolaridade e as faixas de renda mostram uma vulnerabilidade social e econômica na amostra, o que, somado à baixa utilização do auxílio emergencial, sugere que as necessidades financeiras foram um fator decisivo para a manutenção de atividades fora de casa, mesmo com sintomas.

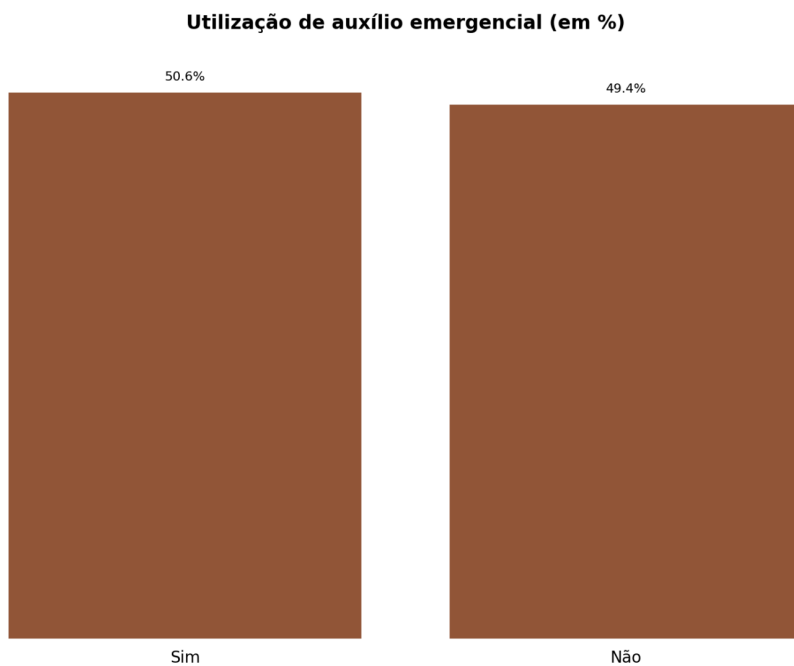
Veja as análises gráficas abaixo.



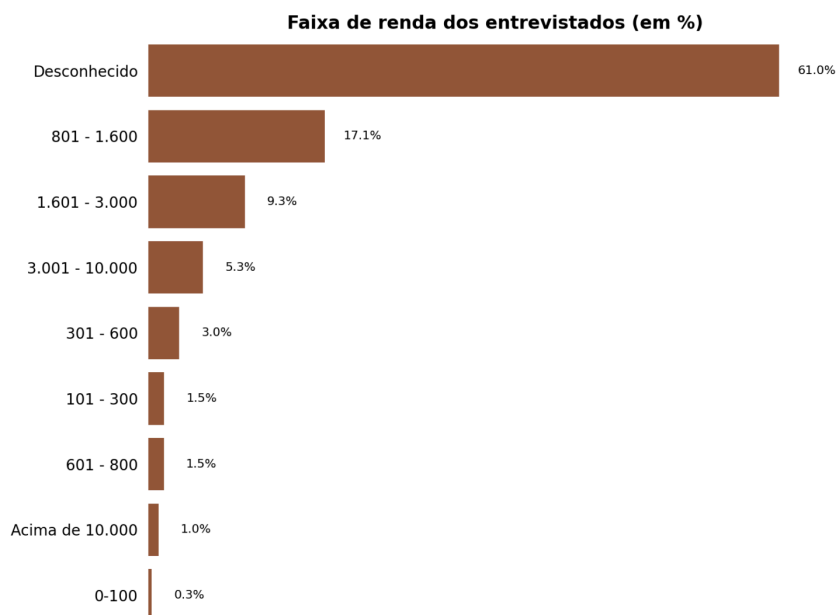
O isolamento social, embora adotado por muitos, não foi rigoroso para a maioria dos entrevistados com sintomas, que precisaram manter atividades fora de casa. Isso indica um desafio na implementação de medidas de distanciamento social.



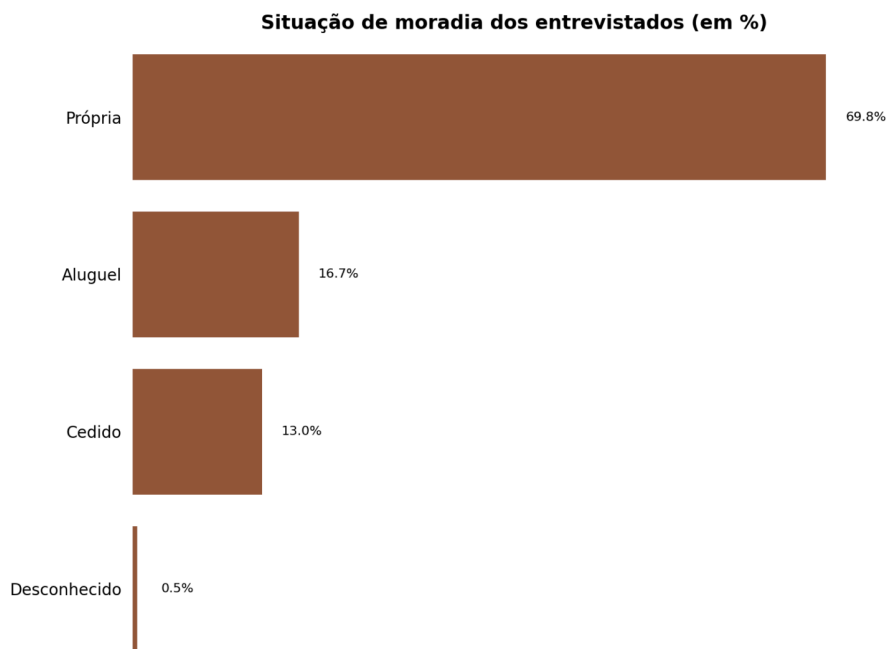
A maioria das pessoas com sintomas na amostra não estava empregada no momento da pesquisa. Isso pode indicar uma vulnerabilidade econômica ou que a doença afetou desproporcionalmente grupos desempregados.



O auxílio emergencial foi uma ferramenta importante para uma parte da população, mas não alcançou a maioria dos entrevistados que relataram sintomas.



A maioria dos entrevistados com sintomas que tinham alguma renda se encontrava em faixas salariais baixas, indicando que a população com menos recursos pode ter sido mais afetada pela doença ou teve mais dificuldade em se proteger.



A maioria da população que reportou sintomas tinham moradia própria, o que pode ter proporcionado uma estabilidade maior durante a pandemia, mas não foi suficiente para evitar a exposição à doença.

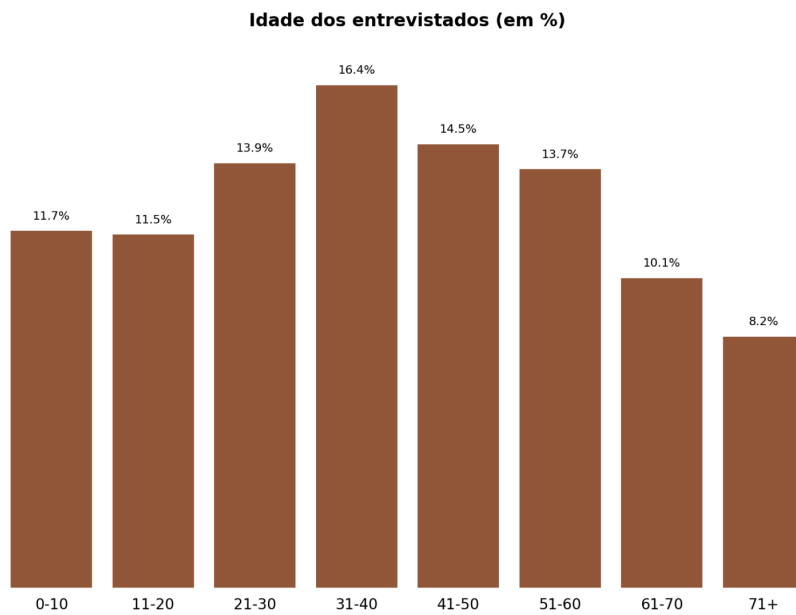
Características Sócio-Demográficas

Idade, Gênero e Escolaridade:

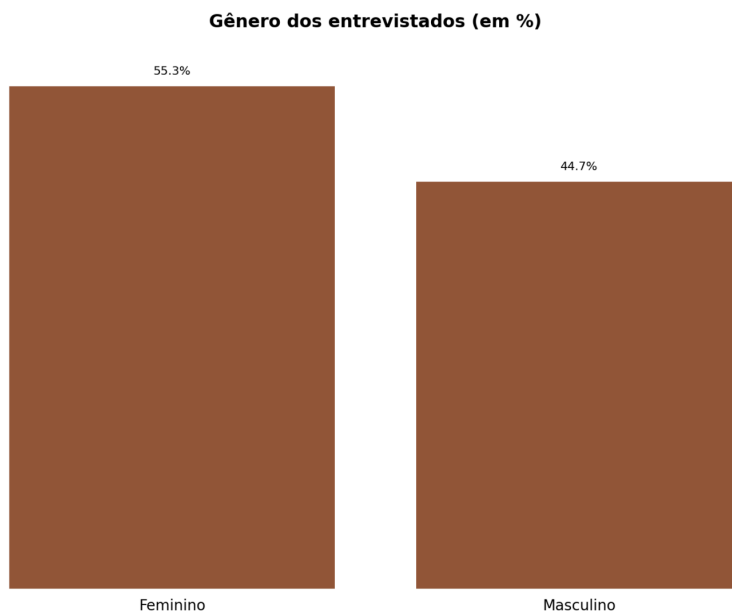
As faixas etárias mais afetadas foram as de 31-40 anos (16,4%), 41-50 anos (14,5%) e 21-30 anos (13,9%). Os dados mostram que a incidência de sintomas foi ligeiramente maior em mulheres (55,3%) do que em homens (44,7%). Em relação à escolaridade, a maioria das pessoas que sentiram sintomas tinham até o Ensino Fundamental incompleto (31.6%).

As pessoas em idade economicamente ativa foram as que mais reportaram sintomas. Isso pode ser um reflexo de uma maior exposição social e profissional, já que esses grupos não puderam aderir completamente ao isolamento social.

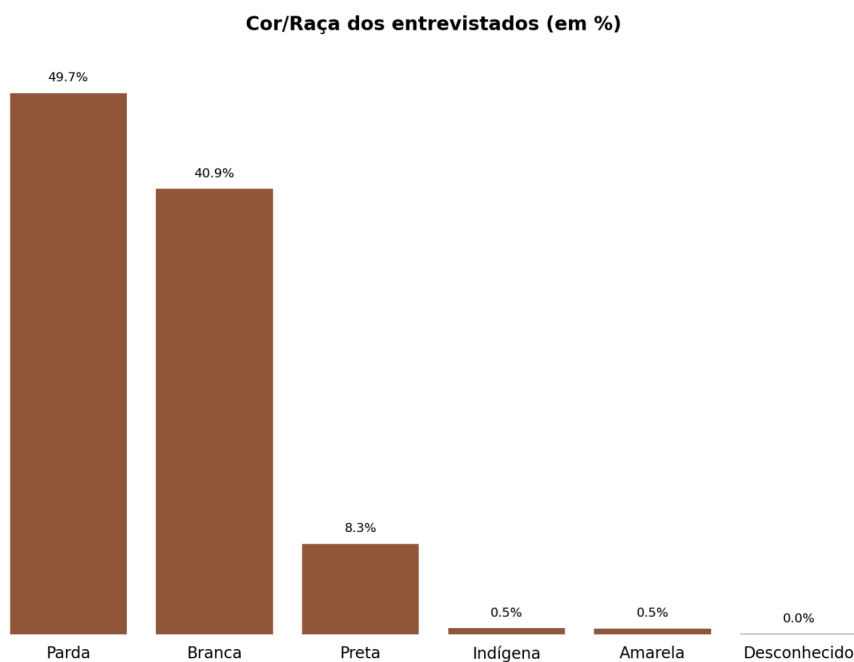
Veja as análises gráficas abaixo.



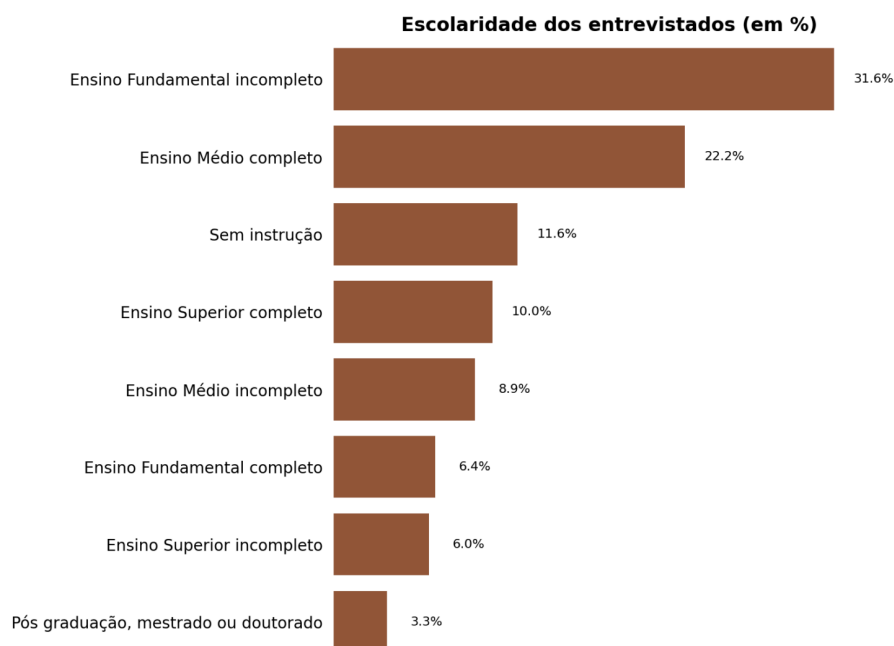
As faixas etárias mais jovens e economicamente ativas (20-50 anos) foram as que mais reportaram sintomas na pesquisa, possivelmente devido à maior exposição social e profissional.



A doença afetou homens e mulheres de forma bastante equilibrada em termos de relatos de sintomas.



A população Parda foi a mais representada na amostra de pessoas com sintomas. Isso pode refletir a distribuição demográfica do Brasil ou uma maior incidência de relatos de sintomas nesse grupo.



A maior parte dos entrevistados com sintomas tinha um nível de escolaridade médio a baixo, o que pode estar correlacionado com as faixas de renda e as atividades profissionais mais expostas.

Conclusão Geral

A análise do questionário revela que a população que experimentou sintomas de COVID-19 em 2020 era majoritariamente composta por pessoas em idade economicamente ativa (entre 21 e 50 anos), com uma incidência levemente maior em mulheres. A maioria desses indivíduos se autodeclarou Parda e possuía até o Ensino Médio completo, com renda baixa ou desempregada. A tosse e a febre foram os sintomas mais comuns, mas, notavelmente, uma grande parte dessa população não buscou auxílio médico, e os casos que evoluíram para internação ou sedação foram minoria. Essa falta de busca por atendimento, aliada a fatores socioeconômicos, sugere que o acesso à saúde foi um desafio significativo.

Ações Estratégicas para o Hospital em um Novo Surto

Com base nessas conclusões, um hospital pode adotar as seguintes ações para se preparar para um novo surto de forma mais eficaz:

1. Otimização dos Protocolos de Triagem e Testagem:

- Foco nos sintomas mais comuns: Aprimorar os protocolos de triagem para priorizar a triagem de pacientes com tosse e febre.
- Incentivo à testagem: Fortalecer a comunicação sobre a importância da testagem, especialmente com o método Swab.

2. Expansão da Acessibilidade ao Atendimento:

- Canais alternativos: Desenvolver e promover canais de telemedicina para pacientes com sintomas leves a moderados, a fim de aumentar a busca por auxílio e aliviar a pressão sobre as emergências.

3. Comunicação Direcionada e Educação em Saúde:

- Campanhas focadas: Criar campanhas de comunicação direcionadas às faixas etárias de 21 a 50 anos, à população Parda e a indivíduos com menor nível de escolaridade.
- Linguagem clara: As campanhas devem usar uma linguagem simples e acessível para explicar a importância do isolamento social, da testagem e da busca por atendimento médico.

A arquitetura em conjunto com as análises e dataviz deste estudo estão disponíveis nos links abaixo:

- [Github](#)
- [Streamlit](#)