Gabriel Martins Danilo Francisco Pires

Pós Tech - 4DTAT

Link do repositório GitHub: https://github.com/machinegab/pos-data-analytics-3-big-data

Caracterização

Requisitos:

- Utilização de no máximo 20 questionamentos realizados na pesquisa
- Utilizar 3 meses para construção da solução
- Caracterização dos sintomas clínicos da população
- Comportamento da população na época da COVID-19
- Características econômicas da Sociedade

Objetivo: Seu objetivo será trazer uma breve análise dessas informações, como foi a organização do banco, as perguntas selecionadas para a resposta do problema e quais seriam as principais ações que o hospital deverá tomar em caso de um novo surto de COVID-19.

Organização da base: os dados selecionados são dos meses de maio a julho de 2020 e o critério para seleção das variáveis passou por diversas categorias dentre as perguntas disponíveis. Nesse caso, buscamos selecionar variáveis que explicassem o comportamento da população nos aspectos educacionais, profissionais, de renda e de saúde (obs: a coluna "horas semanais" foi descartada pois não apresentava valores válidos).

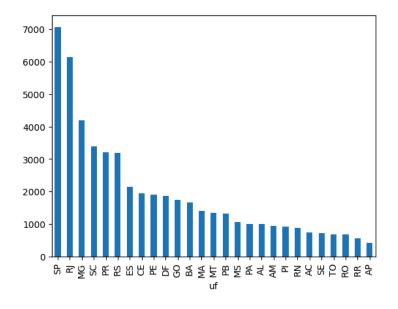
variável	descrição	categoria
mes		
sigla_uf		
a003	sexo	perfil
a004	raça ou cor	perfil
v1022	situação domicilio	perfil
a005	escolaridade	educação
a006	frequenta escola	educação
a006b	aulas presenciais	educação
c007d	atividade da empresa	profissão
c008	horas semanais	profissão
c013	home office	profissão
c01011	faixa rendimento	renda
c01012	valor em dinheiro	renda
b0011	sintoma - febre	saúde
b0012	sintoma - tosse	saúde

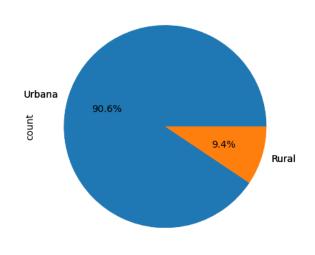
b0019	sintoma - fadiga	saúde
b002	procurou estabelecimento de saúde por conta de sintoma?	saúde
b009b	resultado exame cotonete na boca ou nariz	saúde
b009d	resultado exame coleta sangue furo no dedo	saúde
b009f	resultado exame sangue da veia	saúde
b005	internação >= 1 dia	saúde
b011	medidas de isolamento	saúde

Arquitetura: os dados foram pré tratados no Google Big Query consumindo do banco de dados **pnad covid** e o resultado foi exportado para o Google Colab.

Análise dos dados

A base contém 52.144 linhas com os resultados dos respondentes. Nesse recorte, **51,9**% dos respondentes foram homens e **48,1**% foram mulheres, mostrando equilíbrio entre os entrevistados. Os respondentes são, predominantemente, das regiões sudeste e sul do país e vivem, em grande maioria, nas regiões urbanas.





Distribuição da escolaridade

Quando voltamos a análise para um olhar de escolaridade, nota-se que a maioria dos respondentes (37,8%) possuem ensino **superior completo**, seguido por **Pós-graduação, mestrado ou doutorado** e **Médio completo** (ambos com 18,5%). Os **não instruídos** representam 0,7% da base de dados.

A pandemia de COVID-19 potencializou as desigualdades, com estudantes de baixa renda sofrendo mais devido à falta de acesso à tecnologia. O desempenho acadêmico caiu, em diversas disciplinas, e muitos alunos abandonaram a escola. O ensino remoto

	proportion
escolaridade	
Superior completo	37.8%
Pós-graduação, mestrado ou doutorado	18.5%
Médio completo	18.5%
Superior incompleto	9.3%
Fundamental incompleto	8.3%
Médio incompleto	3.8%
Fundamental completa	3.0%
Sem instrução	0.7%

se expandiu, mas nem sempre manteve a qualidade, e cursos práticos foram amplamente afetados. O impacto pode ser duradouro, especialmente na ampliação da desigualdade de oportunidades no futuro do mercado de trabalho.

Renda

A análise de **renda média por grau de escolaridade** mostra maior concentração de renda em níveis mais elevados (pós-graduação, mestrado ou doutorado) e menor em níveis como "sem instrução". Ao voltar o olhar para o impacto da pandemia na educação e analisarmos

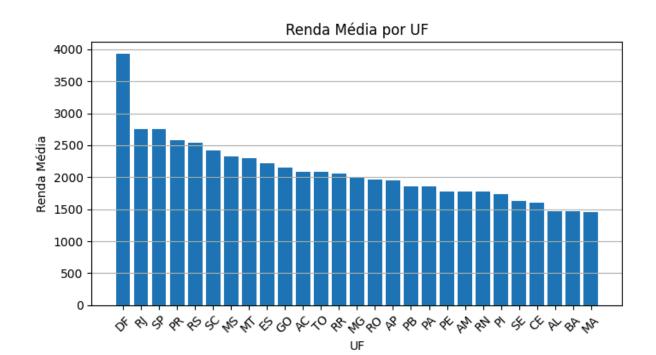
Renda média por escolaridade 6000 5000 4000 3000 2000 1000 0 -undamental incompleto Pós-graduação, mestrado ou doutorado Superior completo Superior incompleto Médio completo Fundamental completa Médio incompleto sem instrução escolaridade

com a renda média por grau de escolaridade, podemos considerar o seguinte:

Desigualdade destacada: Pessoas com maior escolaridade provavelmente sofreram menos impactos financeiros durante a pandemia, já que muitas puderam trabalhar remotamente em setores menos afetados. Isso está alinhado com a expectativa de que esses grupos têm maior renda média.

Trabalhadores com menos escolaridade: Aqueles em níveis de escolaridade mais baixos (como "sem instrução" ou "fundamental incompleto") geralmente ocupam trabalhos que não permitem o home office e que foram mais afetados pelas medidas de isolamento. Com isso, a perda de renda ou emprego foi mais severa para esse grupo.

O resultado pode indicar uma maior resiliência financeira para pessoas com escolaridade mais alta durante a pandemia, enquanto grupos de escolaridade mais baixa enfrentaram desafios econômicos mais acentuados. Isso reforça a desigualdade no impacto da crise.



Desigualdade regional: O Distrito Federal (DF) tem uma renda média significativamente mais alta em comparação com outros estados, o que pode ser explicado pela concentração de órgãos governamentais e empregos públicos na capital. Estados como Rio de Janeiro (RJ), São Paulo (SP) e Paraná (PR) também apresentam rendas mais elevadas, em parte devido à maior industrialização e nível de urbanização.

Disparidade Norte-Nordeste vs. Centro-Sul: Observa-se uma clara tendência de que estados das regiões Norte e Nordeste possuem as menores rendas médias. Isso reflete em resultados históricos no desenvolvimento econômico e no acesso a oportunidades.

Impacto das atividades econômicas: Estados com economias mais diversificadas e setores industriais mais avançados (como SP e RJ) tendem a ter uma renda média maior. Já estados dependentes de atividades agropecuárias ou de setores menos lucrativos apresentam uma renda média mais baixa.

Concentração de riqueza: A análise reforça a concentração de renda em regiões mais desenvolvidas do país, enquanto estados menos desenvolvidos continuam com dificuldades econômicas, perpetuando as desigualdades regionais.

Atividade profissional



Educação tem uma grande proporção, o que indica uma preocupação maior com o impacto da pandemia nesse setor. Isso também reflete a grande quantidade de pessoas que foram impactadas diretamente, especialmente pela adoção do ensino remoto.

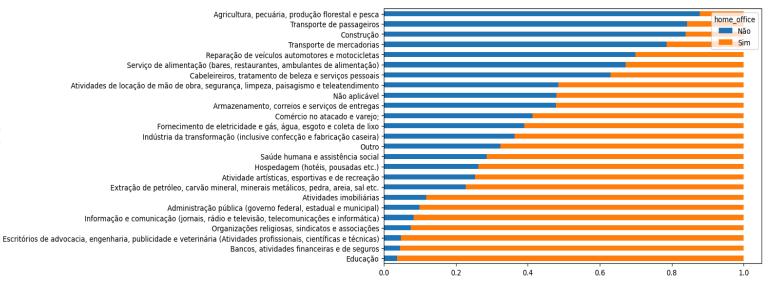
Administração Pública e setores profissionais como engenharia, advocacia, publicidade e veterinária também têm representatividade, que pode refletir maior facilidade desses profissionais de trabalhar remotamente, ou dificuldade continuar de

atividades essenciais durante a pandemia.

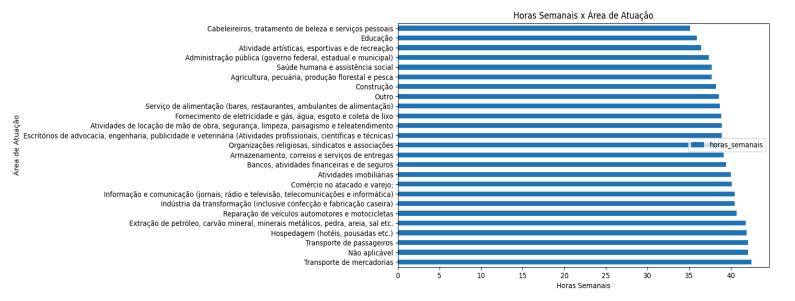
Setores de menor contagem, como **Hospedagem** e **Armazenamento**, provavelmente representam aqueles mais afetados por restrições operacionais severas, ou simplesmente por uma menor proporção de trabalhadores disponíveis para a pesquisa.

Serviços essenciais como **saúde**, **alimentação**, **transportes e segurança** também têm uma amostra considerável, destacando sua importância durante a crise de saúde pública.

Durante a pandemia, o trabalho home office cresceu rapidamente em setores que podiam operar a distância, como tecnologia e educação. Porém, áreas como construção, transporte e serviços enfrentaram dificuldades para adotar esse modelo, mantendo maior parte de suas atividades presenciais, o que aumentou o risco de exposição ao vírus.



A pandemia deu luz a uma divisão entre os setores que conseguiram adotar o home office e aqueles que, pelo perfil de suas atividades, não tiveram essa opção. As áreas que exigem trabalho manual ou presencial, como agricultura, transporte e serviços pessoais, ficaram mais expostas ao vírus e às dificuldades econômicas, enquanto setores mais qualificados, como os de educação, finanças e advocacia, conseguiram se adaptar melhor ao contexto da pandemia com o uso de tecnologias digitais e o trabalho remoto.



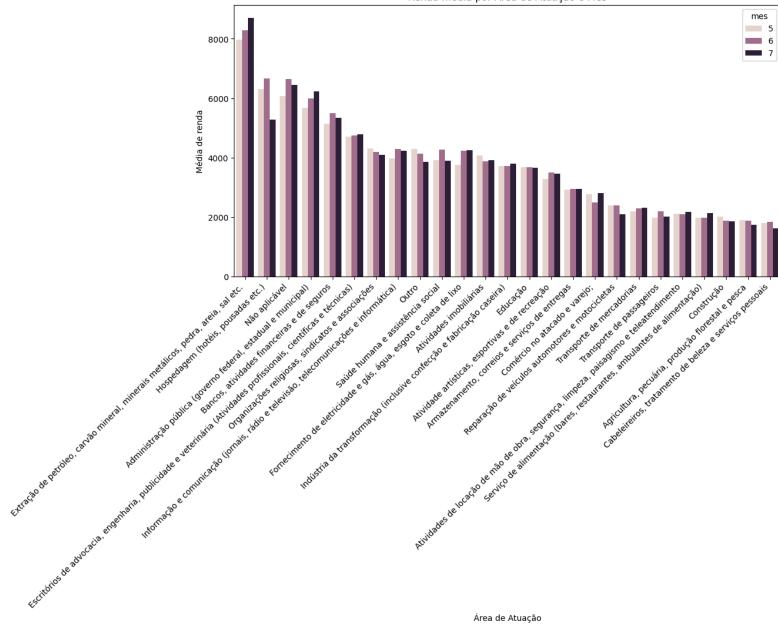
Os dados de renda média por área de atuação mostram que a capacidade de adaptação ao home office foi muito importante durante a pandemia. Setores que dependem de presencial e aqueles com maior flexibilidade para trabalho remoto não tiveram grandes variações na remuneração. A pandemia manteve as desigualdades existentes entre setores bem remunerados e aqueles com rendas mais baixas, com impactos discretos na capacidade de adaptação ao novo comportamento.

Disparidade salarial significativa: A maior renda média está associada à extração de petróleo, carvão mineral e minerais metálicos, com valores mais elevados do que outras áreas. Esse setor, especializado e de risco, justifica a maior salário, mas também tem menor número de trabalhadores, o que contribui para essa média mais alta.

Setores de menor renda: Cabeleireiros, agricultura, construção e serviços de alimentação são as áreas com menor renda média. Esses setores foram impactados durante a pandemia, principalmente por exigirem trabalho presencial e estarem associados a atividades que enfrentaram restrições ou diminuição na demanda.

Setores intermediários: Setores como educação, saúde, segurança e atividades imobiliárias têm uma renda média intermediária, indicando uma compensação moderada pelo tipo de trabalho. Essas áreas também sofreram diferentes graus de impacto durante a pandemia, especialmente na educação, que precisou se adaptar ao ensino remoto.





Setores de maior rendimento: Além da extração de petróleo, bancos, advocacia, engenharia, publicidade e administração pública são as áreas com maior renda média. Esses setores, principalmente bancos e advocacia, têm uma alta demanda por qualificação e conseguiram manter grande parte de suas operações remotamente durante a pandemia, protegendo seus profissionais tanto da exposição ao vírus quanto da perda de renda.

Impactos da pandemia:

• Exposição ao vírus e a renda: Áreas de menor renda, como agricultura e serviços de alimentação, têm pouca flexibilidade para home office, aumentando a exposição ao vírus. Já áreas como bancos e advocacia, que oferecem rendas mais altas,

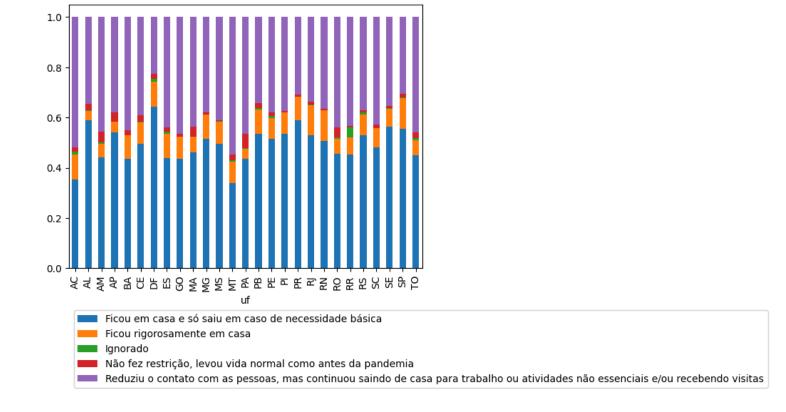
- migraram para o trabalho remoto, mantendo segurança financeira e proteção à saúde.
- Desigualdade de recuperação: Os setores com maior flexibilidade para trabalho remoto e maior renda média não apenas enfrentaram menos impacto, mas também se recuperaram mais rapidamente. Enquanto isso, setores de baixa renda, expostos e sem a possibilidade de transição digital, enfrentaram dificuldades maiores tanto em termos econômicos quanto de saúde pública.

Comportamento e saúde

Os primeiros meses após a chegada do vírus ao Brasil foram cercados de muitas incertezas, desinformação e falta de senso de urgência para uma doença que já era tratada com sua devida importância em diversos países do mundo. Pessoas influentes e pessoas com cargos importantes à frente do país demonstravam pouca ou nenhuma preocupação com os impactos de certas posturas e suas consequências em relação à disseminação do vírus.

Influenciados ou não por esse cenário, os respondentes da pesquisa adotaram um comportamento de média e baixa exposição onde apenas 1,6% (281) das pessoas responderam que **não fizeram nenhum tipo de restrição, levando a vida normalmente**, enquanto 51,1% (8.797) **ficaram em casa e só saíram em caso de necessidade básica**, 37,6% (6.483) **reduziram o contato com as pessoas, mas continuaram saindo de casa para trabalho ou atividades não essenciais e recebendo visitas** e apenas 9,3% (1.596) **ficaram rigorosamente em casa**.

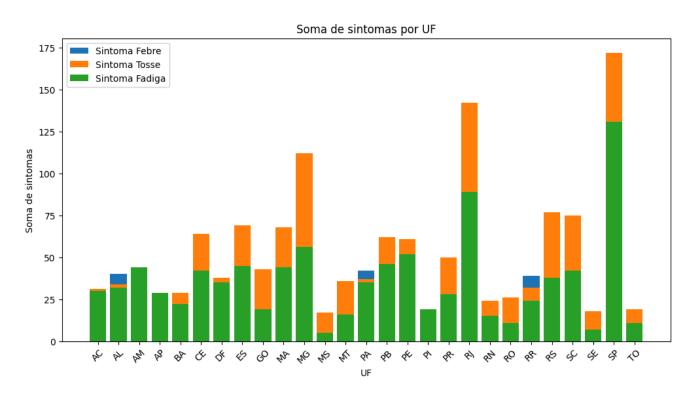
Numa análise por estado, o gráfico revela que as regiões Norte e Nordeste, especialmente os estados do Amazonas (AM) e Pará (PA), apresentaram uma maior incidência de registros vermelhos, indicando que mais pessoas nesses estados levaram uma vida normal, sem restrições, durante a pandemia. Por outro lado, a região Sudeste, com destaque para o Distrito Federal (DF) e Minas Gerais (MG), teve uma maior proporção de pessoas que ficaram rigorosamente em casa (barras laranjas), refletindo um comportamento mais restritivo em relação à mobilidade.



Para complementar a análise dos impactos da exposição, selecionamos os sintomas mais comuns relacionados à COVID-19 de acordo com os órgãos de saúde: **tosse, febre e fadiga**. Além disso, analisamos a taxa de procura de atendimento especializado, resultados de três tipos de exames (exame de sangue no dedo, exame de cotonete e exame de sangue na veia) e ocorrência de internação maior que um dia:

- De todos os respondentes, 4,4% relataram que sentiram pelo menos um ou mais sintomas escolhidos. O sintoma mais recorrente foi tosse, seguido por fadiga e depois por febre.
- Dos respondentes que sentiram algum tipo de sintoma, 24,8% procuraram algum tipo de atendimento especializado. Quando comparamos com os respondentes que não sentiram esses sintomas, a procura por atendimento foi de 0,5%
- Daqueles que relataram algum tipo de sintoma, 2,8% receberam resultado positivo para COVID-19, enquanto apenas 0,5% receberam positivo entre aqueles que não sentiram sintomas.
 - o 4,4% dos que sentiram febre testaram positivo
 - 4,2% dos que sentiram fadiga testaram positivo
 - 3,2% dos que sentiram tosse testaram positivo

O gráfico mostra a soma de relatos dos sintomas (febre, tosse e fadiga) por estado (UF):



São Paulo (SP) tem o maior número total de sintomas relatados, seguido por Rio de Janeiro (RJ) e Minas Gerais (MG). Em todo o Brasil, o sintoma mais relatado foi a fadiga...

Isso sugere que a distribuição de sintomas não foi homogênea no país, com alguns estados apresentando um maior impacto durante a pandemia. O volume de habitantes pode influenciar diretamente a soma de sintomas relatados em cada unidade federativa (UF). Estados com maior população, como **São Paulo (SP)** e **Minas Gerais (MG)**, tendem a ter

mais pessoas infectadas, o que aumenta a probabilidade de maior número de sintomas registrados. Essa relação é esperada, já que regiões com mais habitantes apresentam uma maior exposição ao vírus e maior circulação de pessoas, aumentando o risco de transmissão. No entanto, para verificar se a relação é de fato proporcional à população, podemos analisar os dados pela quantidade de habitantes em cada estado. Podemos entender se o número de sintomas por habitante é maior em alguns estados do que em outros.

Com base na tabela populacional por estado divulgada em 28/06/2023 pelo IBGE, podemos calcular a taxa de sintomas por habitante por unidade federativa usando o seguinte cálculo:

Taxa de sintomas por habitante = Soma de sintomas no estado / População

O cálculo mostra a relação entre a soma de sintomas e a população para cada unidade federativa, calculando a "taxa de sintomas por habitante". Os estados com a maior taxa de sintomas por habitante são:

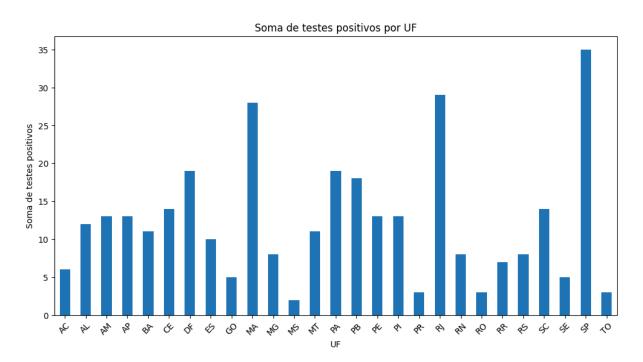
- 1. Roraima (RR): 0.000079 sintomas por habitante.
- 2. Acre (AC): 0.000075 sintomas por habitante.
- 3. Amapá (AP): 0.000059 sintomas por habitante.
- 4. Espírito Santo (ES): 0.000030 sintomas por habitante.
- 5. Paraíba (PB): 0.000026 sintomas por habitante.

Já os estados com as menores taxas são:

- 1. Bahia (BA): 0.000004 sintomas por habitante.
- 2. Minas Gerais (MG): 0.000004 sintomas por habitante.
- 3. São Paulo (SP): 0.000006 sintomas por habitante.

Isso indica que estados com menor população, como Roraima e Acre, apresentam uma taxa de sintomas por habitante mais elevada alta, enquanto estados mais populosos como São Paulo e Minas Gerais possuem taxas mais baixas, o que pode ser influenciado pelo grande número de habitantes diluindo os sintomas registrados.

Ao realizarmos essa análise olhando para os testes positivos, vemos um comportamento um pouco diferente em relação aos sintomas:



São Paulo continua sendo o estado com maior número de casos positivos, de novo seguido por Rio de Janeiro, mas agora com o Maranhão logo em seguida. Porém, para trazer um contexto proporcional à população, precisamos realizar novamente o cálculo feito para os sintomas. Com o cálculo realizado, se observa o seguinte:

Os estados com a maior taxa de infectados por habitante são:

- 1. Amapá (AP): 0.0000177 infectados por habitante.
- 2. Roraima (RO): 0.0000109 infectados por habitante.
- 3. Acre (AC): 0.0000072 infectados por habitante.
- 4. **Distrito Federal (DF)**: 0.0000067 infectados por habitante.
- 5. Paraíba (PB): 0.0000045 infectados por habitante.

Já os estados com as menores taxas são:

- 1. Mato Grosso do Sul (MS): 0.00000072 infectados por habitante.
- 2. Goiás (GO): 0.0000007 infectados por habitante.
- 3. Paraná (PR): 0.00000026 infectados por habitante.

Podemos ver que os estados do norte lideram o quadro de infectados por estado. Também não é possível identificar uma relação com o número de habitantes, pois se esperava que estados mais populosos apresentassem uma taxa proporcional menor do que estados com menor número de habitantes.

É interessante ressaltar que a tabela de dados possui 52.144 linhas de informação e representa 0,026% da população brasileira de 2022.

Conclusão e prospecção futura

A pandemia de COVID-19 trouxe muitos desafios para nossa sociedade em todos os aspectos imagináveis. Tivemos que enfrentar muitas barreiras no combate ao vírus, a desinformação espalhou muitas inverdades sobre a doença, atrapalhando o processo de conscientização. Muitas pessoas tiveram suas rotinas de trabalho extremamente alteradas por conta de novos modelos de trabalho e restrições aplicadas, e, nessas restrições, houve desafios no engajamento acerca da importância do distanciamento social. O atendimento aos infectados e realizações de testes em massa foram mal planejados, o que contribuiu ainda mais na disseminação do vírus.

Muitas reflexões foram tiradas desse período pandêmico, e é obrigação de órgãos públicos e de saúde estarem melhor preparados para uma outra eventual pandemia, abaixo estão algumas sugestões de possíveis ações que o hospital, tanto como ambiente de saúde mas também como agente conscientizador, pode tomar para mitigar possíveis novos impactos. Essas ações buscam não apenas diminuir o impacto de um novo surto, mas também reduzir as desigualdades identificadas durante a análise, garantindo uma resposta mais organizada e eficiente:

- 1. **Reforçar campanhas de vacinação**: É importante garantir que a vacinação chegue a todos, ainda mais aos grupos mais vulneráveis e às regiões onde a adesão ao isolamento foi menor, como Norte e Nordeste. Promover campanhas educativas focadas na importância da vacinação pode aumentar a proteção da população contra novos surtos.
- 2. Fortalecer medidas de isolamento social: Reforçar as medidas de isolamento e incentivar a adesão às recomendações, sobretudo entre os grupos de maior risco. A comunicação sobre a importância do distanciamento social deve ser intensificada, principalmente em áreas com alta exposição ao vírus e onde se verificou uma adesão menor ao isolamento durante o surto anterior.
- 3. Aumento da capacidade hospitalar e recursos médicos: Ampliar a capacidade de atendimento hospitalar para casos graves, especialmente com mais leitos e recursos de emergência, será essencial para evitar a quebra do sistema de saúde. Garantir disponibilidade de testes e equipamentos de suporte, como ventiladores, pode contribuir para melhorar a resposta em caso de nova pandemia.
- 4. **Melhorar atendimento domiciliar e telemedicina**: Expandir o uso de telemedicina e atendimentos domiciliares pode ajudar a diminuir a sobrecarga dos hospitais. A adoção de consultas virtuais para acompanhar de perto os sintomas e oferecer ajuda médica remota deve ser uma prioridade, permitindo um acompanhamento seguro e de qualidade.
- 5. Priorizar assistência às populações vulneráveis: Grupos de menor escolaridade e trabalhadores de setores que não permitem home office foram bastante impactados, tanto pela exposição ao vírus quanto pela diminuição de renda. O foco deve ser em apoiar essas áreas com recursos que possam reduzir a exposição e oferecer suporte econômico.
- 6. Monitoramento constante dos indicadores de saúde: Manter um sistema de monitoramento contínuo das taxas de sintomas e casos positivos é interessante para antecipar medidas preventivas. Os dados podem ser usados para criar alertas de saúde e definir ações em tempo hábil, garantindo uma resposta rápida às mudanças nos indicadores analisados.

- 7. Oferecer suporte psicológico e financeiro: As pessoas de áreas essenciais que não tiveram opção de realizar home office precisam de suporte para minimizar os riscos à saúde e o impacto econômico. Isso inclui suporte psicológico, que veio à tona nos anos de pandemia, já que muitos enfrentaram maior exposição ao risco, e auxílio financeiro para assegurar sua proteção e bem-estar.
- 8. Campanhas educativas e comunicação direta: A comunicação deve ser clara e focada em conscientizar sobre a importância das medidas preventivas. Campanhas educativas focadas no público em geral, utilizando linguagem de fácil entendimento e adaptada aos contextos regionais, podem contribuir para aumentar a adesão às recomendações de saúde.