



Impactos COVID-19

06/10/2023

Grupo 42

Iris Oliveira

Igor do Nascimento Alves

Jordan Ferreira

Salomão Oliveira

Sumário

Sumário.....	1
Relatório de negócios.....	2
Objetivos.....	2
Questões Seleccionadas.....	2
Tabela de Questões Saúde.....	2
Tabela de Questões Socioeconômicas:.....	3
Análises.....	4
I. Home office impediu os sintomas do covid ?.....	4
II. Home Office através dos meses.....	5
III. Pandemia através do tempo.....	6
IV. Impacto no emprego pós pandemia.....	7
Conclusão.....	8
Fontes.....	8
Relatório técnico.....	9
Etapas do projeto.....	9
I. Tratando os dados.....	9
II. Explorando os dados.....	9
III. Insights com os dados.....	9
Tecnologias.....	10
I. Big Query (GCP).....	10
II. Script Python.....	10
III. Jupyter notebook.....	10
IV. Power BI.....	10
Conclusão técnica.....	10
Códigos Fontes.....	11

Relatório de negócios

Durante a pandemia do COVID-19 diversas ações tiveram consequências na saúde e na economia pública, este relatório busca entender os impactos e guiar futuras decisões em uma futura pandemia.

Como um grande hospital, temos vozes com as decisões de governos¹ e somos diretamente afetadas por eles, então esse relatório busca apresentar dados e orientar futuras decisões frente a uma pandemia.

¹Esse relatório não tem viés político e não toma partido em relação a isso.

Objetivos

1. Identificar os principais grupos de pessoas afetados pela pandemia
2. Sugerir estratégias para mitigar os impactos da pandemia

Questões Seleccionadas

Na base de dados, existem cerca de 145 colunas, depois de analisar o dicionário de dados, foram seleccionadas 22 questões que poderiam representar as informações buscadas.

Tabela de Questões Saúde

Questões utilizadas para identificar possíveis pessoas com sintomas do Covid.

Nº	Questão	Código
1	Na semana passada teve febre?	B0011
2	Na semana passada teve tosse?	B0012
3	Na semana passada teve dificuldade para respirar?	B0014
4	Na semana passada teve dor de cabeça?	B0015
5	Na semana passada teve dor de garganta?	B0013
6	Na semana passada teve nariz entupido ou escorrendo?	B0018
7	Na semana passada teve fadiga?	B0019
8	Na semana passada teve dor no peito?	B0016

9	Na semana passada teve náusea?	B0017
10	Na semana passada teve perda de cheiro ou sabor?	B00111
11	Ao procurar o hospital, teve que ficar internado por um dia ou mais?	B005

Tabela de Questões Socioeconômicas:

Questões buscando identificar o perfil econômico das pessoas afetadas pelo Covid

Nº	Questão	Código
12	Escolaridade	A005
13	Local que buscou atendimento foi posto de saúde/Unidade básica de saúde /Equipe de Saúde da Família?	B0041
14	Local que buscou atendimento foi ambulatório ou consultório privado?	B0044
15	Local que buscou atendimento foi pronto socorro do SUS/UPA?	B0042
16	Local que buscou atendimento foi hospital do SUS?	B0043
17	Durante a internação, foi sedado, entubado e colocado em respiração artificial com ventilador?	B006
18	Tem algum plano de saúde médico, seja particular, de empresa ou de órgão público?	B007
19	Na semana passada, o(a) Sr(a) estava em trabalho remoto (home office ou teletrabalho)?	C013
20	Seguro desemprego	D0061
21	Auxílios emergenciais relacionados ao coronavírus	D0051
22	Tem carteira de trabalho assinada ou é funcionário público estatutário?	C007B

Essas foram as colunas selecionadas, para entender o perfil das pessoas que responderam o questionário, identificando se estavam trabalhando em casa e se pegaram covid.

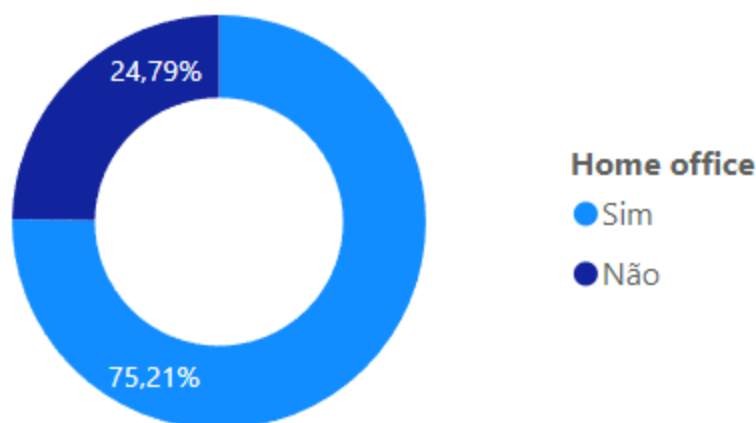
Análises

Foram criadas visualizações buscando entender o que os dados nos contam em relação a pandemia e as questões selecionadas do questionário. Um ponto de atenção: muitos dados da nossa base estavam nulos e o dicionário classificou esses dados como *Não informado*. Isso dificultou a análise gerando preocupação quanto a precisão que teremos nas análises.

I. Home office impediu os sintomas do covid ?

As pessoas que apresentaram sintomas estavam de home office, um resultado inesperado já que o home office deveria mitigar a pandemia. Aqui entramos no campo da especulação do motivo que pode ser que essas pessoas não deixaram de sair mesmo trabalhando em casa ou outras pessoas que moram na mesma casa não trabalhavam em home office. (Em uma segunda análise podemos pegar mais questões do questionário e descobrir isso)

Teve Sintomas de COVID



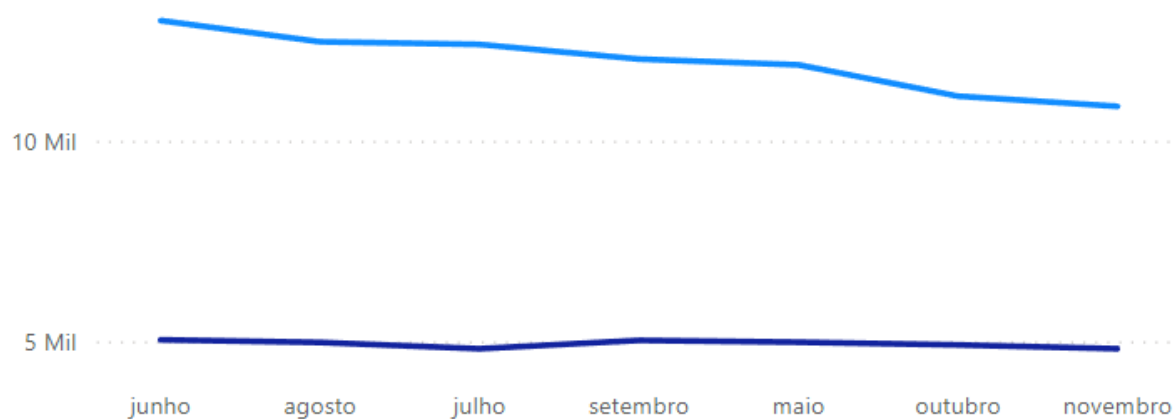
Fonte: Autor

II. Home Office através dos meses

Buscamos entender também se entre os entrevistados existiu um aumento de pessoas que passaram para o modelo home office de trabalho ao longo dos meses. E não foi o que identificamos, na verdade as pessoas começaram a voltar ao trabalho presencial.

Homeoffice

● Não ● Sim



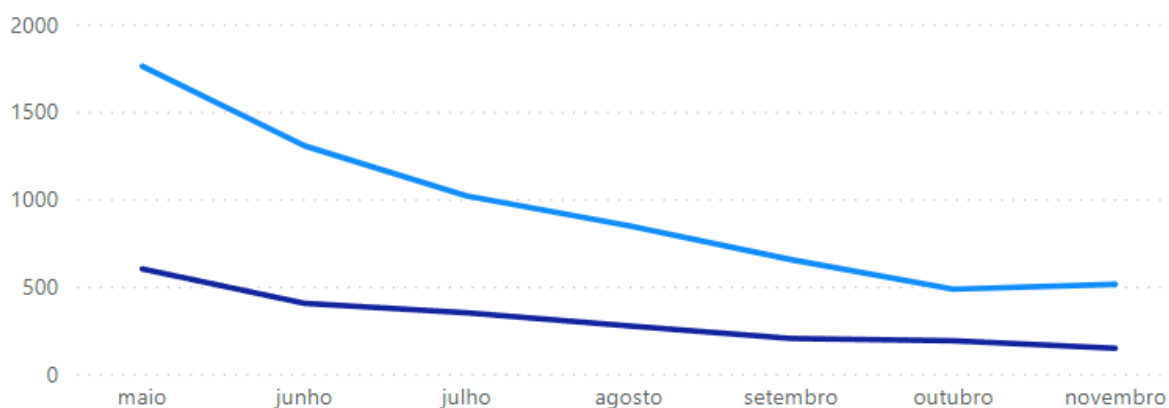
Fonte: Autor

III. Pandemia através do tempo

Outra análise foi se os casos foram aumentando e com isso foi impactado em relação a pessoa trabalhar em home office ou não. Aqui notamos uma queda nas pessoas com sintomas em ambos os casos. Parece ser um comportamento da pandemia e não relacionado ao modelo de trabalho.

Teve Sintomas de COVID

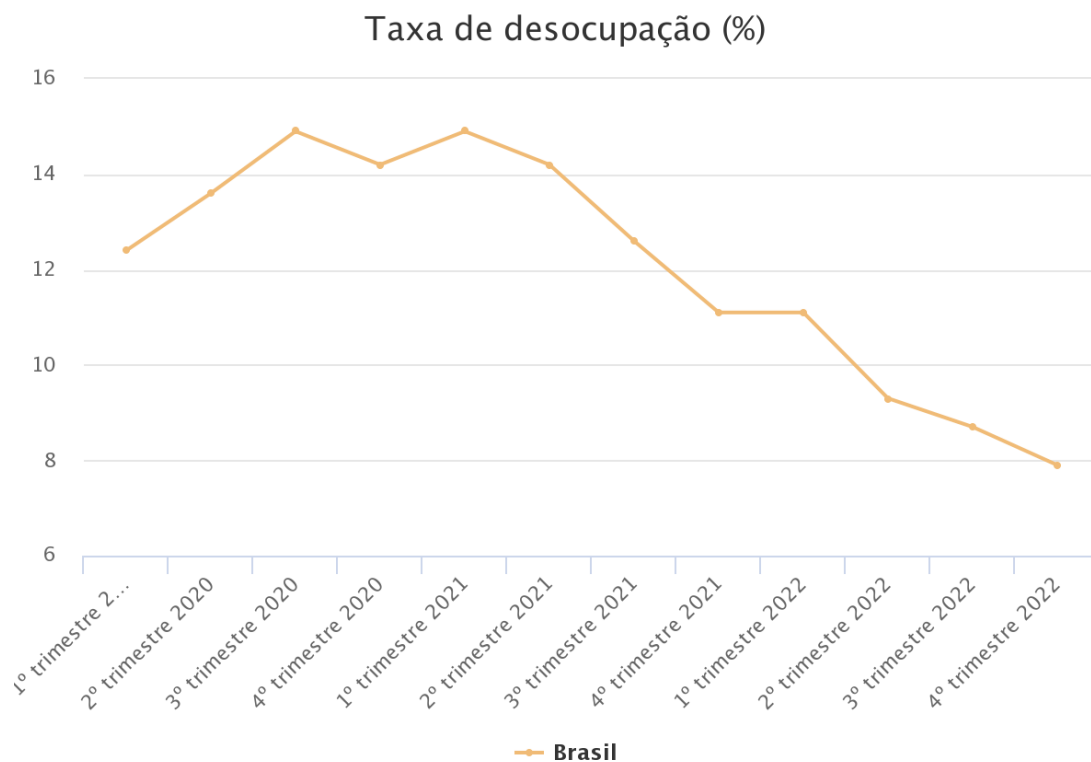
Home Office ● Não ● Sim



Fonte: Autor

IV. Impacto no emprego pós pandemia

Buscando análises externas feita pelo IBGE encontramos informações de aumento no desemprego durante a pandemia, então pessoas não empregadas durante esse período não entrariam na nossa análise de home office ou não, podendo ter um grande impacto no número de pessoas com covid.



Fonte: IBGE – Pesquisa Nacional por Amostra de Domicílios Contínua trimestral

Conclusão

Diante dos dados apresentados, é fundamental que o hospital esteja preparado para tomar medidas eficazes em caso de um novo surto de COVID-19. A análise demonstra que o comportamento da população é influenciado por uma variedade de fatores, incluindo medidas de saúde pública, características econômicas e atitudes individuais. Portanto, o planejamento deve ser flexível e adaptável, levando em consideração não apenas as características clínicas dos sintomas, mas também a dinâmica social e econômica.

Além disso, é importante considerar a qualidade dos dados, uma vez que muitos deles estavam nulos ou classificados como "Não informado". Isso destaca a necessidade de coletar informações mais precisas em futuras pesquisas para melhor orientar as ações do hospital em resposta a uma possível nova pandemia.

Em resumo, a análise dos dados do PNAD-COVID-19 do IBGE fornece informações valiosas para o planejamento e resposta a surtos de COVID-19, destacando a importância de abordar não apenas os aspectos clínicos, mas também os fatores sociais e econômicos que influenciam o comportamento da população.

Fontes

1. <https://covid19.ibge.gov.br/pnad-covid/>
2. <https://www.ibge.gov.br/explica/desemprego.php>
3. <https://www.ibge.gov.br/estatisticas/multidominio/condicoes-de-vida-desigualdade-e-pobreza.html>
4. [Taxa média de desemprego cai a 9,3% em 2022, menor patamar desde 2015 | Agência de Notícias](#)

Relatório técnico

Para produzir o relatório de negócios seguimos diversas etapas técnicas aqui descritas e no final uma conclusão sobre deficiência identificadas no processo e os produtos resultados deste projeto.

Etapas do projeto

I. Tratando os dados

Inicialmente o projeto foi carregado na GCP e com o auxílio da ferramenta BigQuery conseguimos explorar nossos dados e decidir os tratamentos iniciais.

II. Explorando os dados

Depois carregamos os dados loucamente e os acessamos via SQL. Essa etapa envolveu entender quais colunas estavam presentes em todas as tabelas.

Inicialmente tínhamos a maior tabela com 148 colunas e a menor com 114.

Com isso rodamos um código para identificar as colunas em comum.

Com as colunas identificadas podemos juntar todas as tabelas em um DataFrame só com 114 colunas e 2.650.459 linhas.

Por último passamos para uma etapa manual de identificar as descrições das colunas e entender quais faziam sentido para a nossa análise. Chegando 24 colunas (temos a coluna do mês da pesquisa e coluna que combina as informações de sintomas de Covid)

III. Insights com os dados

Com a fonte de dados tratada e reduzida passamos para uma ferramenta de análise onde criamos diversos gráficos e selecionamos os que traziam informações sobre a análise que buscamos.

Tecnologias

I. Big Query (GCP)

Utilizar uma solução na nuvem foi essencial na parte inicial do projeto já que contamos com um volume alto de dados. Utilizando SQL conseguimos processar nossos dados e entendê-los sem a lentidão que encontramos ao rodar os códigos locais.

II. Script Python

Utilizar um script python para carregar os dados em um banco on-premise, isso permitiu trabalhar com os dados localmente.

III. Jupyter notebook

Com um Jupyter notebook terminamos de tratar e reduzir os dados. Aqui encontramos lentidão por limitações da máquina em relação ao volume de dados.

IV. Power BI

Com a base reduzida podemos criar visualizações importantes da análise na ferramenta de BI.

Conclusão técnica

Com este projeto, alcançamos resultados significativos, incluindo a criação de um painel de controle que acompanha as principais métricas para analisar o impacto do home office na contaminação pelo COVID-19. Além disso, estabelecemos uma base de dados e um pipeline de tratamento de dados, culminando na elaboração de um relatório de negócios baseado nas informações encontradas.

Durante o processo, identificamos dois pontos críticos. Primeiramente, devido ao volume considerável de dados, tornou-se essencial utilizar ferramentas de nuvem para armazenar e processar os dados, especialmente antes da etapa de tratamento. Isso evitou qualquer lentidão no processo. Em segundo lugar, destacamos a importância de criar um catálogo de dados estruturado. No projeto, notamos que a documentação dos dados estava apresentada em uma planilha Excel, o que dificultou a análise. A criação de um dicionário de dados bem organizado é fundamental para orientar futuros projetos.

Este projeto foi um exemplo de integração de diversas habilidades, incluindo o uso do BigQuery e SQL para manipulação de dados, programação em Python e Jupyter Notebook para o tratamento de dados, visualização de dados e o Power BI para a criação do painel de controle. Além disso, a colaboração de uma equipe multidisciplinar foi crucial para o seu sucesso.



Neste projeto, desenvolvemos um painel de controle para monitorar métricas relacionadas ao impacto do home office na contaminação por COVID-19. Além disso, criamos uma base de dados e um pipeline de tratamento de dados, resultando em um relatório de negócios.

Identificamos dois pontos críticos:

1. A necessidade de usar ferramentas de nuvem devido ao grande volume de dados, especialmente antes do tratamento, para evitar lentidão.
2. A importância de criar um catálogo de dados bem organizado em vez de usar uma planilha Excel para facilitar análises futuras.

Este projeto envolveu habilidades em BigQuery, SQL, Python, Jupyter Notebook, visualização de dados e Power BI, sendo executado por uma equipe multidisciplinar.

Códigos Fontes

1. https://github.com/IgorNascAlves/Tech_Challenge_Fase03