

FIAP

GIULIA FELIX DA SILVA

RM 357020

Data Analytics

PÓS TECH - 6DTAT

São Paulo

12/2024

Sumário

Introdução	3
O PNAD	4
Recuperação dos dados.....	4
Exploração de dados.....	4
Visualização de dados e impactos	5
Plano de ação	15
Referencias	16

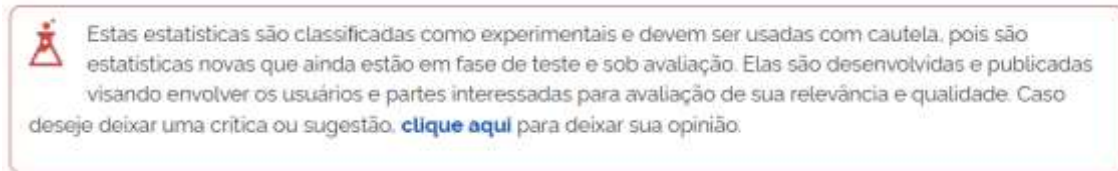
Introdução

O objetivo deste projeto é analisar os dados que foram levantados através da PNAD (pesquisa nacional por amostra de domicílios) durante o período da pandemia pelo instituto brasileiro de geografia e estatística (IBGE).

A partir desta análise espera-se a criação de um plano de ação para os hospitais caso uma nova pandemia ocorra.

O PNAD

Segundo o site oficial do IBGE o PNAD tem como objetivo “*estimar o número de pessoas com sintomas referidos associados à síndrome gripal e monitorar os impactos da pandemia da COVID-19 no mercado de trabalho brasileiro.*” Além disso, a pesquisa conta com dados experimentais conforme a observação abaixo:



Recuperação dos dados

Os dados utilizados nesta pesquisa, foram retirados da plataforma do IBGE, são referentes aos meses de julho, agosto e setembro.

A base possui um volume consideravelmente expressivo e visto a necessidade do uso de banco de dados em nuvem, optei por subir os arquivos na GCP, mais especificamente no BigQuery.

Os arquivos foram importados em uma instância do serviço após uma análise prévia sobre o conteúdo via Excel, com foco em garantir a integridade do arquivo “a olho nu”.

Exploração de dados

Compreender a base é um fator importante para que possamos estruturar de forma adequada nossas análises e sugestões com relação ao plano de ação para os hospitais, para isso contei com um dicionário de dados que me deixou claro o objetivo de cada campo

Além disso, através de alguns comandos no banco eu consegui detalhar melhor os dados inicialmente.

Etapa 01

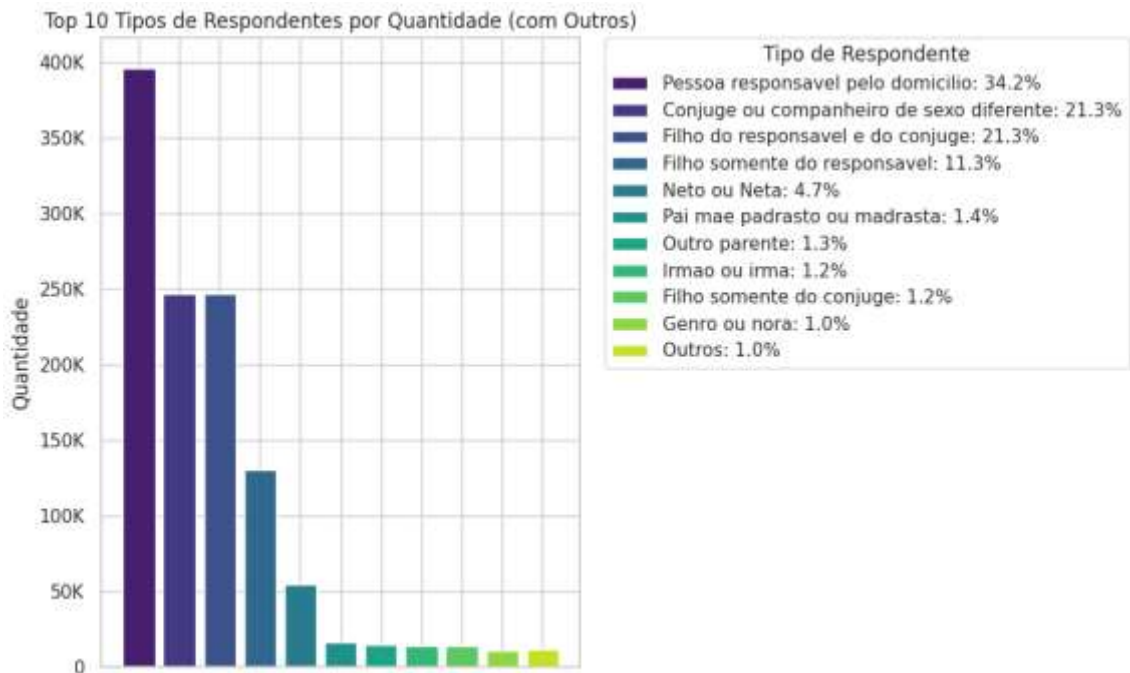
- Garantia de que os dados foram imputados na tabela de forma conforme esperado.
- Conhecendo os tipos de dados
- Garantindo que o mês está igualmente o identificador numérico do mesmo
- Validação da idade mínima e máxima contida na base
- Verificação da renda das pessoas com foco em identificar outliers

Além destes itens que foram importantes para que eu garantisse conhecimento inicial, o processo de consulta no banco e busca dos identificadores dos dados no material de apoio me fizeram ir se habituando e conhecendo os demais dados.

Visualização de dados e impactos

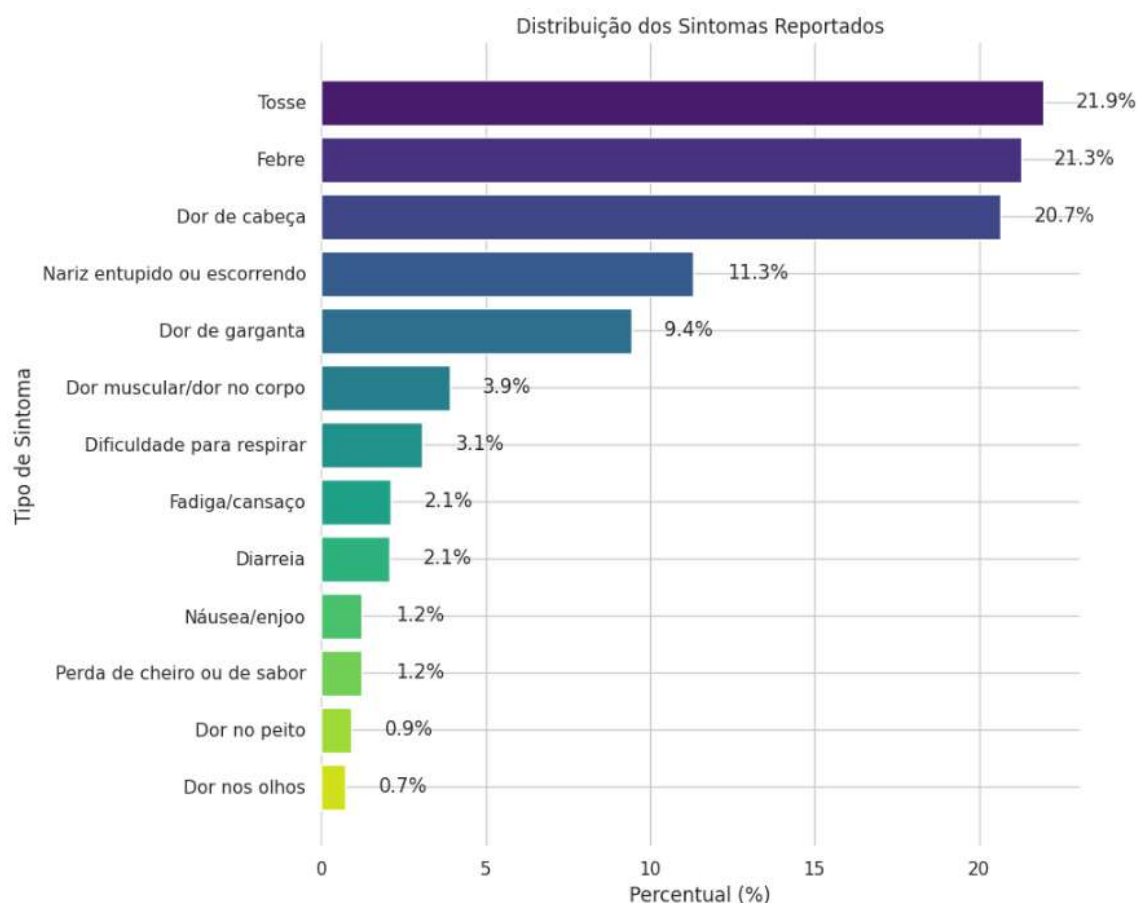
Nesta pesquisa foram entrevistadas 1.157.984 pessoas, considerando o período de julho de 2020 a setembro de 2020

A pesquisa do IBGE foi realizada via telefone e em 34,2% dos casos conseguimos falar com os responsáveis da residência.



A grande maioria das pessoas que foram entrevistadas não sentiram nenhum sintoma do Covid-19, analisando os dados temos somente 5,60% de indivíduos com algum desconforto.

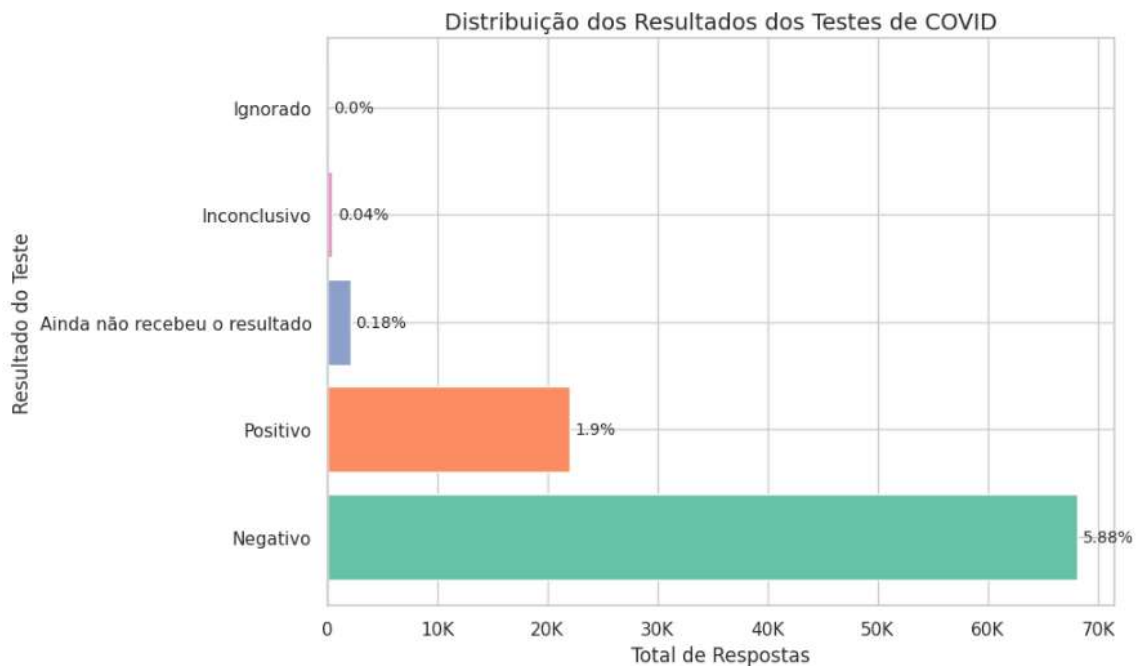
Dentre os sintomas mais comuns, podemos destacar: Tosse, febre e dor de cabeça como sendo os principais.



A população mesmo com sintomas não foi tão aderente na busca de testes somente 8,16% dos entrevistados que sentiram algo desconforto, realizaram algum teste para saber se estava infectado pelo Corona vírus. Os testes mais realizados foram do cotonete, exame de sangue via furo no dedo e coleta de sangue via veia do braço.

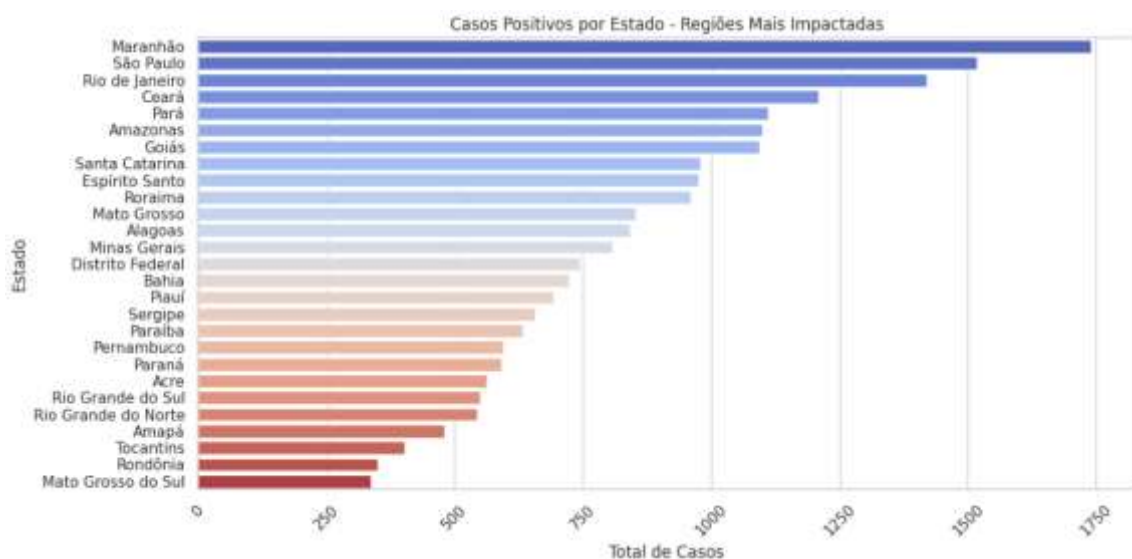
Realizou o teste	QTD	%
Não	1058382	91,4%
Sim	94517	8,2%
Ignorado	5085	44%

É importante que a partir dos sintomas o teste seja feito para que diagnosticado o problema, o tratamento e as medidas preventivas sejam passadas através da orientação médica.



Dentre as pessoas que buscaram fazer o teste temos 22485 com resultado positivo e 69108 com resultado negativo. Os demais resultados não serão tratados aqui visto sua relevância ao atual cenário.

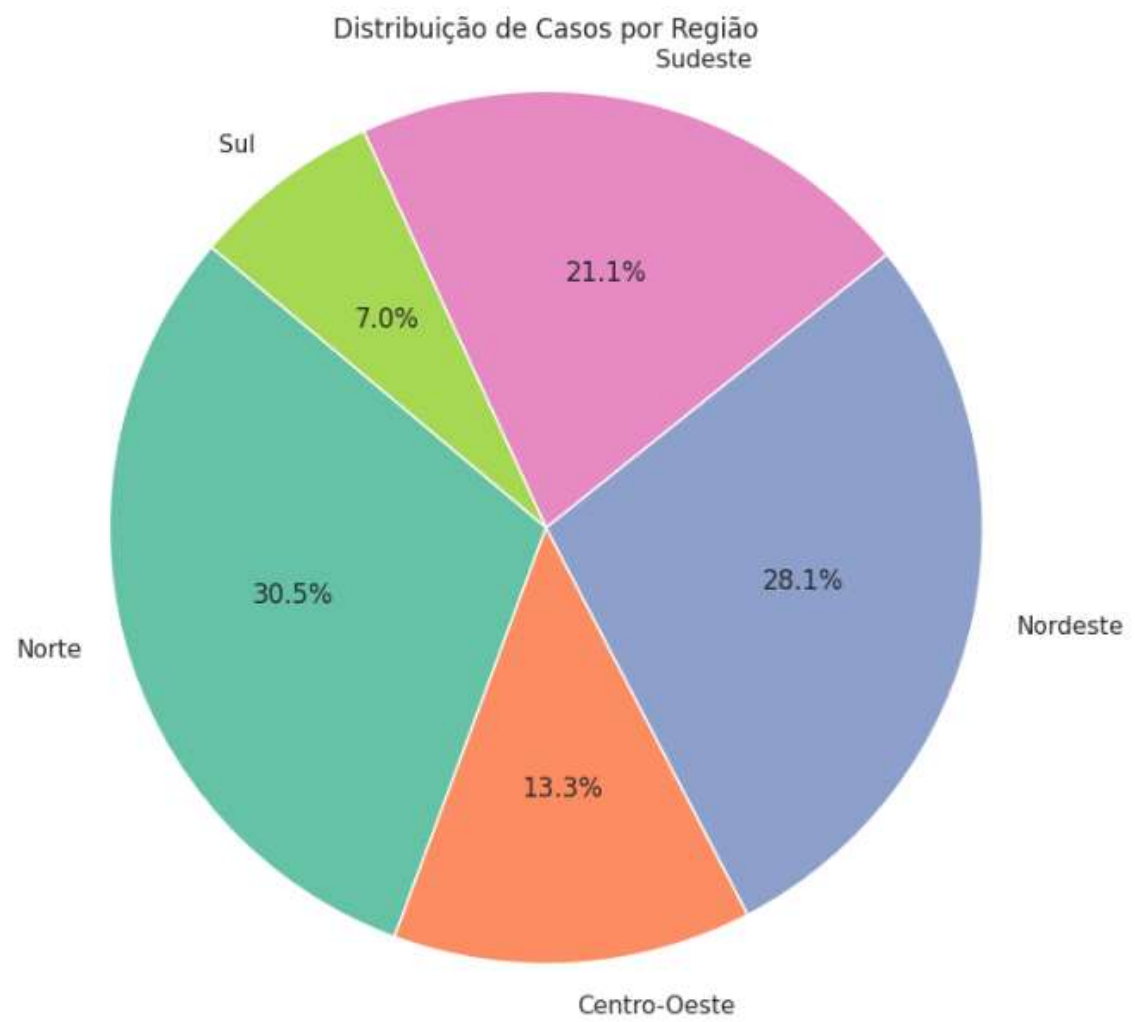
Dado o volume de 23,79% infectados, é importante analisarmos aonde esta localizado o maior volume de pessoas, assim podemos adiantar o processo de preparo e disponibilidade de equipes médicas.



“Vale lembrar que a Grande São Luís (Região que abrange a capital maranhense e os municípios de Raposa, Paço do Lumiar e São José de Ribamar) foi a primeira do Brasil a decretar o 'lockdown' (confinamento obrigatório).”

Estado	Qtd	Percentual
Maranhão	1740	8,0%
São Paulo	1518	7,0%
Rio de Janeiro	1421	6,0%
Ceará	1210	5,0%
Pará	1111	5,0%
Amazonas	1100	5,0%
Goiás	1094	5,0%
Santa Catarina	980	4,0%
Espírito Santo	975	4,0%
Roraima	961	4,0%
Alagoas	842	4,0%
Minas Gerais	808	4,0%
Mato Grosso	854	4,0%
Bahia	724	3,0%
Distrito Federal	745	3,0%
Piauí	693	3,0%
Sergipe	658	3,0%
Paraíba	633	3,0%
Pernambuco	596	3,0%
Paraná	592	3,0%
Acre	564	3,0%
Rio Grande do Sul	551	2,0%
Rio Grande do Norte	544	2,0%
Amapá	480	2,0%
Tocantins	403	2,0%
Rondônia	350	2,0%
Mato Grosso do Sul	338	2,0%

Ao reavaliarmos a disposição dos dados e reformarmos eles para uma nova visualização e análise, podemos identificar que as regiões norte e nordeste são as mais impactadas com a contaminação da doença. Considerando as pesquisas realizadas pelo IDH (índice de desenvolvimento humano) identificamos uma lacuna de desenvolvimento nos estados destas regiões. Os estados onde as taxas mais aumentaram são os mais pobres.



Além da região aonde estão localizados o foco da contaminação, o tipo de residência também é um fator importante à considerarmos, ao analisarmos os dados é possível identificar que a maior parte dos casos está localizado na (UF Unidade da Federação, excluindo a região metropolitana e a RIDE)

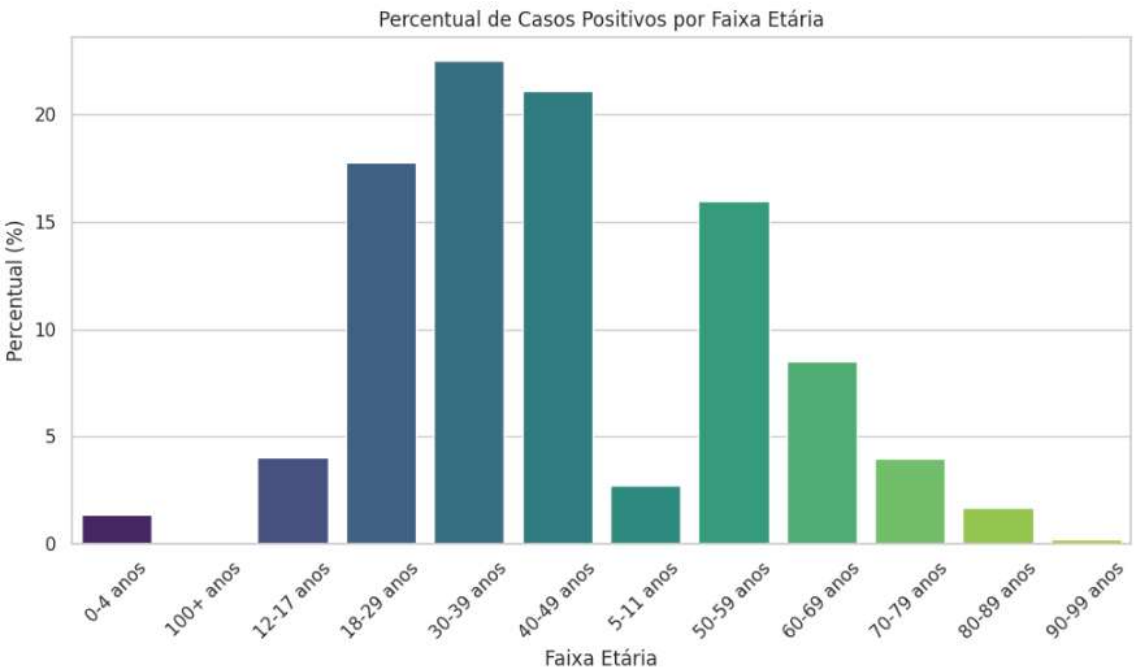
Tipo	Total	Percentual
UF	11791	52%
Capital	7545	33%
RM	2844	13%
RIDE	305	2%

Outras importantes características para analisarmos, são os fatores relacionados a população em si.

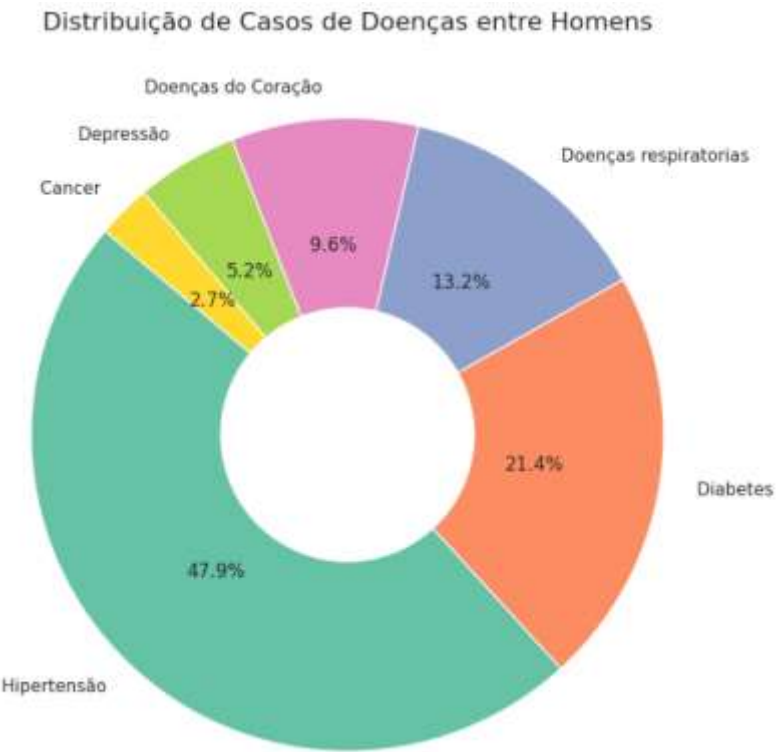
O maior volume de pessoas infectadas são mulheres, representando 55% da nossa base.

Sexo	Qtd	Percentual
Mulher	12453	55%
Homem	10032	45%

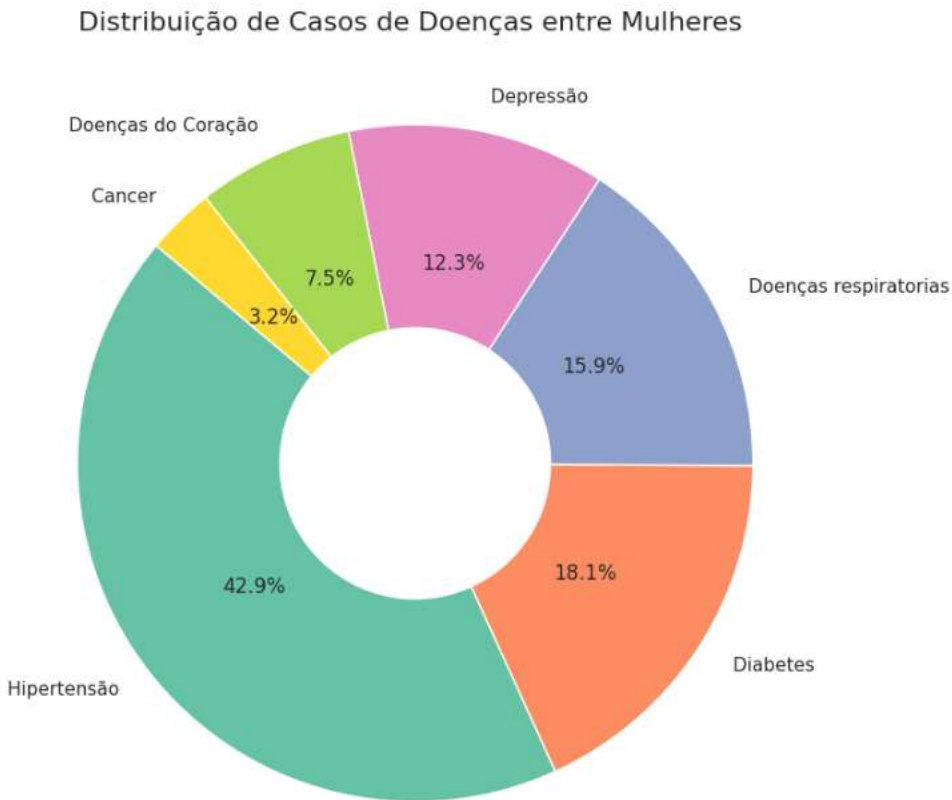
E o maior volume de pessoas contagiadas, possuem idade entre 30 a 39 anos, seguidas por 40 a 49 anos e assim em diante.



Dentre 45% dos homens que fizeram o teste e constataram positivo, 3601 indivíduos possuíam alguma doença pré-existente.



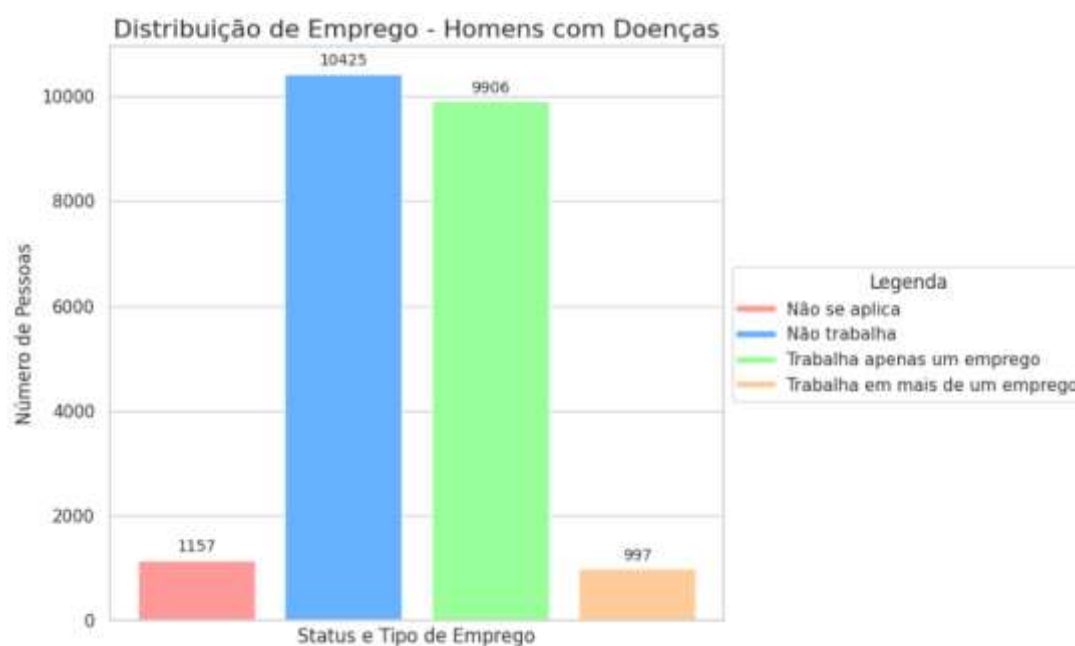
Para as mulheres o número de pessoas chega a 5780



Os fatores pré-existent aumentam o risco do desenvolvimento de uma forma grave da doença. Para estas pessoas é importante que os demais pontos continuem sendo tratados de forma integral com acompanhamento intensivo.

Também é importante que façamos análise e adicionemos considerações aos pontos relacionados a jornada profissional. A pandemia gerou grande impacto sobre a população, além da saúde prejudicada de um grande volume de indivíduos, a situação financeira e profissional também foi bastante impactada.

- Durante o período analisado, 46,36% das pessoas não estavam trabalhando.
- 4,43 das pessoas atuam em mais de um emprego para complementar renda.
- 44,06% atua em uma jornada única de trabalho.



O desemprego, ou falta de renda é um dos principais fatores que geram impacto ao cuidado com a saúde, visto que o acesso ao plano médico ainda é exorbitante quando comparamos a média de renda financeira da população ao custo que é gerado para atender esta demanda.

Somente 31% das pessoas que foram diagnosticadas possuíam convênio médico vinculado ao seu nome.

Possui Convênio	Qtd	Percentual
Não respondeu	9	3%
Não	15066	66%
Sim	7410	31%

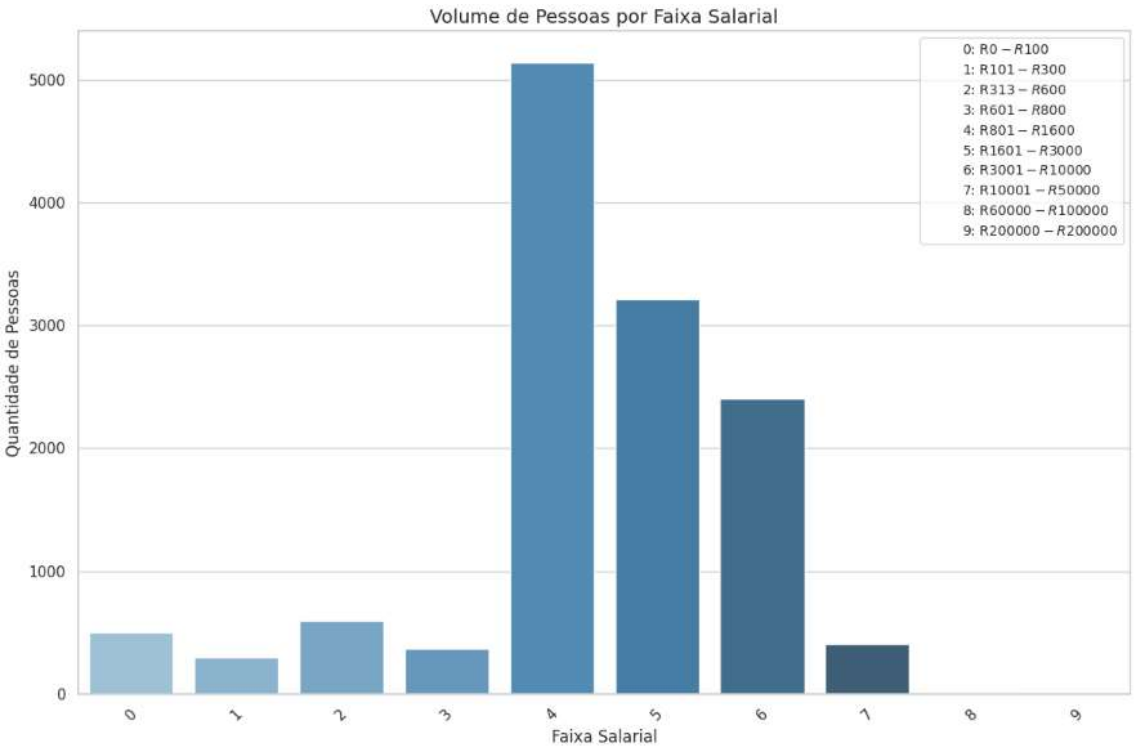
Também consideramos analisar os casos aonde houve pedido de empréstimo dentre as famílias, o resultado foi:

Status	Qtd	Percentual
Não solicitou	9525	99,08%
Sim, e pelo menos um morador conseguiu	766	0,79%
Sim, mas nenhum morador conseguiu	134	0,13%

Mas segundo as pesquisas as famílias ainda que conseguissem estes valores, priorizariam outros segmentos como, alimentação, e moradia ao invés de saúde.

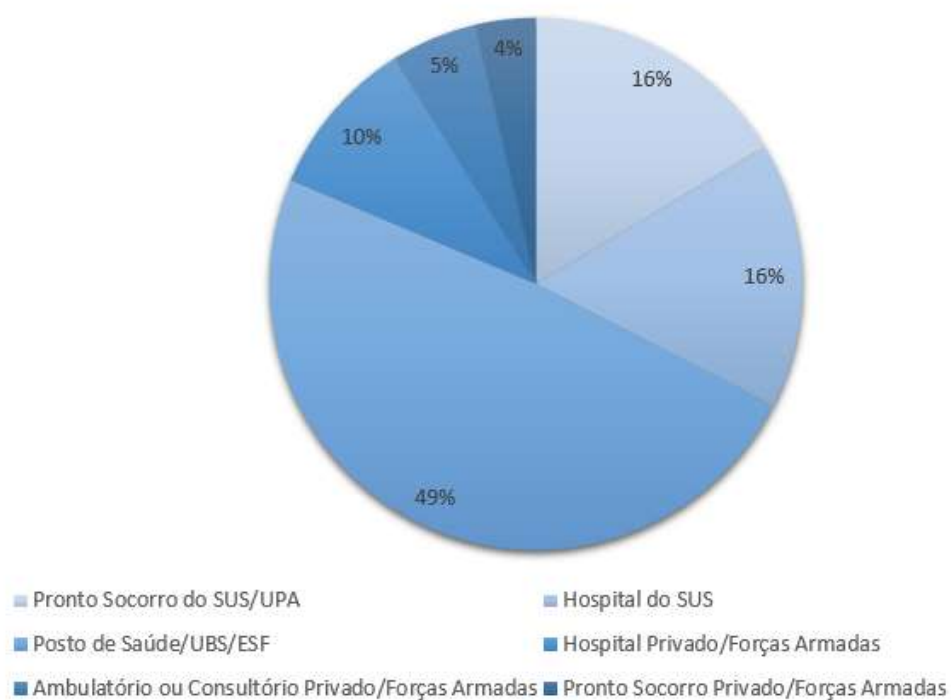
A renda média da população está em concentrada no valor de R\$2.668,17, mas considerando o custo aproximado de um convênio para uma pessoa com idade entre em 30 e 39 anos, o valor pode representar em média 33% da renda.

Podemos analisar no gráfico a distribuição de salários de acordo com o volume de pessoas infectadas



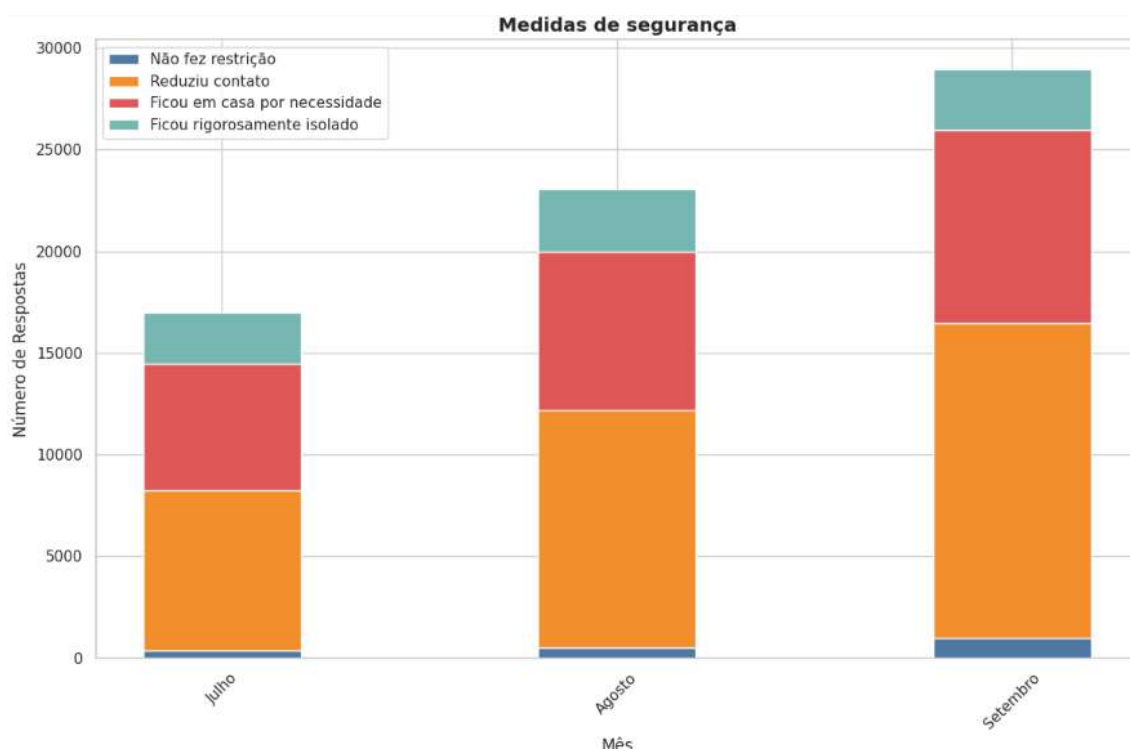
Por fim, ao analisarmos o comportamento da população, é possível identificarmos que somente 14,88% das pessoas que foram diagnosticadas buscaram ajuda, as demais não buscaram ou não quiseram responder.

Percentual de pessoas que buscaram ajuda



Além disso 346 pessoas foram internadas e 130 necessitaram de respirador artificial.

Sob o percentual de pessoas que realizaram o teste e obtiveram resultado negativo, é possível observar que a redução de contato e o isolamento foram uma das medidas mais adotadas.



Plano de ação

Com base nas análises realizadas sobre o comportamento, saúde, jornada de trabalho e moradia, é evidente que precisamos de novos protocolos para atender a população caso ocorra uma nova pandemia.

Os entrevistados possuem uma renda consideravelmente pequena, para suportar um volume expressivo de gastos com saúde, ainda mais quando levamos em conta o tamanho médio das famílias que residem no país e o custo de vida.

Por este motivo, é interessante que programas sejam construídos preferencialmente em parceria com o governo para que pessoas com doenças pré-existentes, possam ser tratadas sem intervenção, medidas de conscientização sejam reforçadas e para que o primeiro acesso ao teste seja incentivado em caso de sintomas intensos.

Para os colaboradores que possuem um ou mais empregos, é ideal que haja acompanhamento por profissionais da saúde dentro de um intervalo recorrente e periódico nas empresas a fim de garantir a continuidade desta jornada profissional, reduzindo o impacto financeiro da sociedade.

Também é válido avaliar a disposição da estrutura dos hospitais segmentando de forma rígida os casos leves, intermediários e graves, evitando a recorrente contaminação no ambiente hospitalar.

Por fim, o incentivo ao cuidado com a saúde é um dos pontos que podemos trabalhar de forma recorrente e sem a necessidade de um surto tão grande para que seja colocado em prática. Então programas que possam contextualizar as pessoas sobre a importância do cuidado consigo mesmo, com o lar e os familiares, as possíveis formas de administrar sintomas leves e o reforço sobre a necessidade de acompanhamento periódico com o médico, podem ser adotados e bem utilizados neste cenário.

Referencias

https://www.ibge.gov.br/estatisticas/investigacoes-experimentais/estatisticas-experimentais/27946-divulgacao-semanal-pnadcovid1?t=o-que-e&utm_source=covid19&utm_medium=hotsite&utm_campaign=covid_19

<https://covid19.ibge.gov.br/pnad-covid/>

<https://covid19.ibge.gov.br/pnad-covid/saude.php>

<https://basedosdados.org/dataset/c747a59f-b695-4d19-82e4-fef703e74c17?table=5894e1ac-dc08-465d-91a3-703683da85ba>

<https://basedosdados.org/dataset/c747a59f-b695-4d19-82e4-fef703e74c17?table=5dcaf8f0-6509-4dea-958b-4d23bc2a8695>

<https://cloud.google.com/bigquery/docs/reference/legacy-sql?hl=pt-br>

<https://www.alura.com.br/conteudo/google-bigquery-funcoes-logicas-numericas?srsId=AfmBOoqGK-ZXpzG1aqlRsp-3BznLRmzGnKZGQctQiPC08DNWGYX-u-G>

<https://g1.globo.com/ma/maranhao/noticia/2020/05/26/maranhao-e-o-segundo-pior-estado-do-brasil-em-distanciamento-social-diz-pesquisa-nacional.ghtml>

<https://cognatis.com.br/norte-e-nordeste-concentram-os-municipios-com-menor-idh-do-pais/>

<https://www.revistasindico.com.br/economia/2021/o-impacto-da-pandemia-na-saude-financeira-das-pessoas/482916>

<https://portalconectado.com.br/pandemia-impactos-na-saude-financeira-das-familias-no-brasil/>

Query

Etapas 01

```
SELECT
    column_name,
    data_type,
    is_nullable
FROM
    `tcmodulo03.PNAD_COVID.INFORMATION_SCHEMA.COLUMNS`
WHERE
    table_name = 'agosto';
```

Tipos de dados existentes

- V1031- Float
- V1032 - Float
- Rm_Ride – String
- Demais colunas – Int64

Validação do mês que estou analisando, garantindo que minha tabela esta integra com o que criei.

```
SELECT *
FROM `tcmodulo03.PNAD_COVID.agosto`
where v1013 = 8
```

Validação da idade das pessoas que responderam esta pesquisa. Nesta analise identifiquei que muitas pessoas possuíam idade menor que 16 anos, o que poderia impactar a analise visto a possível inconsistência das respostas. Porém como estamos trabalhando com dados estatísticos experimentais segundo a própria classificação do PNAD, optei por seguir com todo o volume de dados da base.

```
SELECT *
FROM `tcmodulo03.PNAD_COVID.agosto`
where (A002 < 0) or (A002 > 130)

CREATE OR REPLACE VIEW `tcmodulo03.PNAD_COVID.view_detalhe_menores_16` AS
SELECT
    A002 AS idade,
    COUNT(*) AS total_por_idade,
    COUNT(*) * 100.0 / (SELECT COUNT(*) FROM `tcmodulo03.PNAD_COVID.view_dados_PNAD`
WHERE A002 <= 16) AS percentual_por_idade
FROM
    `tcmodulo03.PNAD_COVID.view_dados_PNAD`
WHERE
    A002 <= 16
GROUP BY
    A002
ORDER BY
```

idade;

A conferência do valor financeiro de renda, para identificar possíveis outliers

```
SELECT *
FROM `tcm modulo03.PNAD_COVID.agosto`
where (C01012 >100000)
```

Isso foi feito para as três tabelas que estou analisando, além de me garantir a integridade de algumas informações e apoiar esta visualização de outliers, eu também fui me ambientando com as colunas e ganhando conhecimento sobre como responderia as perguntas.

Etapa 02

Quantidade de pessoas entrevistadas

```
-- Contar o total de entrevistados em cada tabela e somar
SELECT
(SELECT COUNT(*) FROM `tcm modulo03.PNAD_COVID.julho`) AS total_julho,
(SELECT COUNT(*) FROM `tcm modulo03.PNAD_COVID.agosto`) AS total_agosto,
(SELECT COUNT(*) FROM `tcm modulo03.PNAD_COVID.setembro`) AS total_setembro,
-- Somar os totais para obter o número geral de entrevistados
(
(SELECT COUNT(*) FROM `tcm modulo03.PNAD_COVID.julho`) +
(SELECT COUNT(*) FROM `tcm modulo03.PNAD_COVID.agosto`) +
(SELECT COUNT(*) FROM `tcm modulo03.PNAD_COVID.setembro`)
) AS total_entrevistados
```

Resultado 1.157.984 pessoas

Aqui eu criei uma view para não ter que rodar tabela por tabela, assim eu unifico tudo em uma consulta.

```
##Criar a junção das três
CREATE OR REPLACE VIEW `tcm modulo03.PNAD_COVID.view_dados_PNAD` AS
SELECT *, "julho" AS mes FROM `tcm modulo03.PNAD_COVID.julho`
UNION ALL
SELECT *, "agosto" AS mes FROM `tcm modulo03.PNAD_COVID.agosto`
UNION ALL
SELECT *, "setembro" AS mes FROM `tcm modulo03.PNAD_COVID.setembro`;
```

Quem respondeu a pesquisa

```
SELECT
view_dados.*,
associativa.Descricao AS tipo_respondente
FROM
`tcm modulo03.PNAD_COVID.view_dados_PNAD` AS view_dados
LEFT JOIN
`tcm modulo03.PNAD_COVID.condicaoDomicilio` AS associativa
ON
view_dados.A001A = CAST(associativa.Cod AS INT64);
```

Percentual de pessoas com sintomas

Query – Aqui criei uma view nova só para separar quem teve algum sintoma

```
CREATE OR REPLACE VIEW `tcmodulo03.PNAD_COVID.view_casos_sintomas` AS
SELECT
*
FROM
`tcmodulo03.PNAD_COVID.view_dados_PNAD`
WHERE
B0011 = 1 OR B0012 = 1 OR B0013 = 1 OR B0014 = 1 OR
B0015 = 1 OR B0016 = 1 OR B0017 = 1 OR B0018 = 1 OR
B0019 = 1 OR B00110 = 1 OR B00111 = 1 OR B00112 = 1 OR
B00113 = 1;
```

Distribuição dos sintomas

Ai criei uma validação dos tipos de sintomas, contando o total por tipo e qual o percentual em cima da base de todos os com sintomas.

```
CREATE OR REPLACE VIEW `tcmodulo03.PNAD_COVID.view_analise_sintomas` AS
WITH casos_unificados AS (
  SELECT
    CASE
      WHEN CAST(B0011 AS INT64) = 1 THEN 1
      WHEN CAST(B0012 AS INT64) = 1 THEN 2
      WHEN CAST(B0013 AS INT64) = 1 THEN 3
      WHEN CAST(B0014 AS INT64) = 1 THEN 4
      WHEN CAST(B0015 AS INT64) = 1 THEN 5
      WHEN CAST(B0016 AS INT64) = 1 THEN 6
      WHEN CAST(B0017 AS INT64) = 1 THEN 7
      WHEN CAST(B0018 AS INT64) = 1 THEN 8
      WHEN CAST(B0019 AS INT64) = 1 THEN 9
      WHEN CAST(B00110 AS INT64) = 1 THEN 10
      WHEN CAST(B00111 AS INT64) = 1 THEN 11
      WHEN CAST(B00112 AS INT64) = 1 THEN 12
      WHEN CAST(B00113 AS INT64) = 1 THEN 13
      ELSE NULL
    END AS cod_sintoma
  FROM
    `tcmodulo03.PNAD_COVID.view_casos_sintomas`
),
casos_filtrados AS (
  SELECT
    cod_sintoma
  FROM
    casos_unificados
  WHERE
    cod_sintoma IS NOT NULL
)
SELECT
  s.descricao AS tipo_sintoma, -- Descrição do sintoma
  COUNT(c.cod_sintoma) AS total_por_tipo, -- Total de casos por tipo
  COUNT(c.cod_sintoma) * 100.0 / (SELECT COUNT(*) -- Percentual em relação ao
  total
  FROM casos_filtrados) AS percentual
FROM
```

```

    casos_filtrados c
JOIN
    `tcmodulo03.PNAD_COVID.sintomas` s
    ON c.cod_sintoma = s.cod -- Vinculando pelo código do
    sintoma
GROUP BY
    s.descricao
ORDER BY
    total_por_tipo DESC;

```

Quantas fizeram teste

```

CREATE OR REPLACE VIEW `tcmodulo03.PNAD_COVID.view_realizouTeste` AS
SELECT
    CASE
        WHEN B008 = 1 THEN 'Sim'
        WHEN B008 = 2 THEN 'Não'
        WHEN B008 = 9 THEN 'Ignorado'
        ELSE 'Não informado'
    END AS realizou_teste,
    COUNT(*) AS total_respostas, -- Total de pessoas por
    resposta
    COUNT(*) * 100.0 / (SELECT COUNT(*) FROM `tcmodulo03.PNAD_COVID.view_dados_PNAD`)
AS percentual -- Percentual em relação ao total
FROM
    `tcmodulo03.PNAD_COVID.view_dados_PNAD`
WHERE
    B008 IN (1, 2, 9) -- Considera somente
    respostas válidas
GROUP BY
    realizou_teste
ORDER BY
    total_respostas DESC;

```

Percentual de resultados positivos e negativos

Resultado positivo

```

CREATE OR REPLACE VIEW `tcmodulo03.PNAD_COVID.view_resultados_positivos` AS
SELECT *
FROM `tcmodulo03.PNAD_COVID.view_dados_PNAD`
WHERE
    B009B = 1 OR
    B009D = 1 OR
    B009F = 1;

```

Resultado negativo

```

CREATE OR REPLACE VIEW `tcmodulo03.PNAD_COVID.view_resultados_negativos` AS
SELECT *
FROM `tcmodulo03.PNAD_COVID.view_dados_PNAD`
WHERE
    B009B = 2 OR
    B009D = 2 OR
    B009F = 2;

```

Dados para criação do gráfico dos tipos de resultado

```
CREATE OR REPLACE VIEW `tcmodulo03.PNAD_COVID.view_resultadoTeste` AS
WITH resultados AS (
  SELECT
    CASE
      WHEN B009B IN (1, 2, 3, 4, 9) THEN B009B
      WHEN B009D IN (1, 2, 3, 4, 9) THEN B009D
      WHEN B009F IN (1, 2, 3, 4, 9) THEN B009F
      ELSE NULL
    END AS resultado_teste
  FROM `tcmodulo03.PNAD_COVID.view_dados_PNAD`
  WHERE
    B009B IN (1, 2, 3, 4, 9) OR
    B009D IN (1, 2, 3, 4, 9) OR
    B009F IN (1, 2, 3, 4, 9)
)
SELECT
  resultado_teste,
  CASE
    WHEN resultado_teste = 1 THEN 'Positivo'
    WHEN resultado_teste = 2 THEN 'Negativo'
    WHEN resultado_teste = 3 THEN 'Inconclusivo'
    WHEN resultado_teste = 4 THEN 'Ainda não recebeu o resultado'
    WHEN resultado_teste = 9 THEN 'Ignorado'
    ELSE 'Não aplicável'
  END AS descricao,
  COUNT(*) AS total_respostas,
  COUNT(*) * 100.0 / (SELECT COUNT(*) FROM `tcmodulo03.PNAD_COVID.view_dados_PNAD`)
  AS percentual
FROM resultados
GROUP BY
  resultado_teste, descricao
ORDER BY
  total_respostas DESC;
```

Positivos - Região em que moram

View para separação por estado

```
CREATE OR REPLACE VIEW `tcmodulo03.PNAD_COVID.view_casos_positivos_por_estado` AS
WITH estados AS (
  SELECT 11 AS UF, 'Rondônia' AS estado
  UNION ALL
  SELECT 12, 'Acre'
  UNION ALL
  SELECT 13, 'Amazonas'
  UNION ALL
  SELECT 14, 'Roraima'
  UNION ALL
  SELECT 15, 'Pará'
  UNION ALL
  SELECT 16, 'Amapá'
  UNION ALL
  SELECT 17, 'Tocantins'
  UNION ALL
  SELECT 21, 'Maranhão'
```

```

UNION ALL
SELECT 22, 'Piauí'
UNION ALL
SELECT 23, 'Ceará'
UNION ALL
SELECT 24, 'Rio Grande do Norte'
UNION ALL
SELECT 25, 'Paraíba'
UNION ALL
SELECT 26, 'Pernambuco'
UNION ALL
SELECT 27, 'Alagoas'
UNION ALL
SELECT 28, 'Sergipe'
UNION ALL
SELECT 29, 'Bahia'
UNION ALL
SELECT 31, 'Minas Gerais'
UNION ALL
SELECT 32, 'Espírito Santo'
UNION ALL
SELECT 33, 'Rio de Janeiro'
UNION ALL
SELECT 35, 'São Paulo'
UNION ALL
SELECT 41, 'Paraná'
UNION ALL
SELECT 42, 'Santa Catarina'
UNION ALL
SELECT 43, 'Rio Grande do Sul'
UNION ALL
SELECT 50, 'Mato Grosso do Sul'
UNION ALL
SELECT 51, 'Mato Grosso'
UNION ALL
SELECT 52, 'Goiás'
UNION ALL
SELECT 53, 'Distrito Federal'
)
SELECT
    e.estado,
    COUNT(*) AS total_casos,
    COUNT(*) * 100.0 / SUM(COUNT(*)) OVER () AS percentual
FROM
    `tcmodulo03.PNAD_COVID.view_resultados_positivos` p
JOIN
    estados e
ON
    p.UF = e.UF
GROUP BY
    e.estado
ORDER BY
    total_casos DESC;

```

Aqui adicionei a região

```
CREATE OR REPLACE VIEW `tcmodulo03.PNAD_COVID.casos_por_regiao` AS
```

```

SELECT
    'Norte' AS regioao,
    8442 AS total_casos,
    39 AS percentual
UNION ALL
SELECT
    'Centro-Oeste' AS regioao,
    4026 AS total_casos,
    17 AS percentual
UNION ALL
SELECT
    'Nordeste' AS regioao,
    8527 AS total_casos,
    36 AS percentual
UNION ALL
SELECT
    'Sudeste' AS regioao,
    6412 AS total_casos,
    27 AS percentual
UNION ALL
SELECT
    'Sul' AS regioao,
    2123 AS total_casos,
    9 AS percentual;

```

View para residência

```

CREATE OR REPLACE VIEW `tcmodulo03.PNAD_COVID.positivos_residencia` AS
SELECT
    CASE
        WHEN V1023 = 1 THEN 'Capital'
        WHEN V1023 = 2 THEN 'Resto da RM'
        WHEN V1023 = 3 THEN 'Resto da RIDE'
        WHEN V1023 = 4 THEN 'Resto da UF'
        ELSE 'Não especificado'
    END AS tipo_residencia,
    COUNT(*) AS total_casos,
    COUNT(*) * 100.0 / SUM(COUNT(*)) OVER () AS percentual
FROM
    `tcmodulo03.PNAD_COVID.view_resultados_positivos`
GROUP BY
    tipo_residencia
ORDER BY
    total_casos DESC;

```

Segmentação de pessoas com sintomas por sexo

```

CREATE OR REPLACE VIEW `tcmodulo03.PNAD_COVID.positivos_porsexo` AS
SELECT
    CASE
        WHEN A003 = 1 THEN 'Homem'
        WHEN A003 = 2 THEN 'Mulher'
        ELSE 'Não especificado'
    END AS sexo,
    COUNT(*) AS total_casos,
    COUNT(*) * 100.0 / SUM(COUNT(*)) OVER () AS percentual
FROM
    `tcmodulo03.PNAD_COVID.view_resultados_positivos`
GROUP BY
    sexo
ORDER BY
    total_casos DESC;

```

```

END AS sexo,
COUNT(*) AS total_casos,
COUNT(*) * 100.0 / SUM(COUNT(*)) OVER () AS percentual
FROM
`tcmodulo03.PNAD_COVID.view_resultados_positivos`
GROUP BY
sexo
ORDER BY
total_casos DESC;

```

Faixa etária das pessoas infectadas

```

CREATE OR REPLACE VIEW `tcmodulo03.PNAD_COVID.view_faixas_etarias_positivos` AS
SELECT
CASE
WHEN A002 BETWEEN 0 AND 4 THEN '0-4 anos'
WHEN A002 BETWEEN 5 AND 11 THEN '5-11 anos'
WHEN A002 BETWEEN 12 AND 17 THEN '12-17 anos'
WHEN A002 BETWEEN 18 AND 29 THEN '18-29 anos'
WHEN A002 BETWEEN 30 AND 39 THEN '30-39 anos'
WHEN A002 BETWEEN 40 AND 49 THEN '40-49 anos'
WHEN A002 BETWEEN 50 AND 59 THEN '50-59 anos'
WHEN A002 BETWEEN 60 AND 69 THEN '60-69 anos'
WHEN A002 BETWEEN 70 AND 79 THEN '70-79 anos'
WHEN A002 BETWEEN 80 AND 89 THEN '80-89 anos'
WHEN A002 BETWEEN 90 AND 99 THEN '90-99 anos'
ELSE '100+ anos'
END AS faixa_etaria,
COUNT(*) AS total_casos,
COUNT(*) * 100.0 / SUM(COUNT(*)) OVER () AS percentual
FROM
`tcmodulo03.PNAD_COVID.view_resultados_positivos`
GROUP BY
faixa_etaria
ORDER BY
faixa_etaria ASC;

```

Possuem outras doenças

Homem

```

CREATE OR REPLACE VIEW `tcmodulo03.PNAD_COVID.homens_com_doencas` AS
SELECT
'Homem' AS sexo,
CASE
WHEN td.string_field_1 LIKE '%Doenças docora%' THEN 'Doenças do Coração' --
Alteração para abranger possíveis variações de codificação
ELSE td.string_field_1
END AS tipo_doenca,
COUNT(*) AS total_casos,
COUNT(*) * 100.0 / SUM(COUNT(*)) OVER () AS percentual
FROM
`tcmodulo03.PNAD_COVID.view_resultados_positivos` vrp
CROSS JOIN UNNEST([
STRUCT(1 AS codigo, vrp.B0101 AS resposta),

```



```

        STRUCT(2 AS codigo, vrp.B0102 AS resposta),
        STRUCT(3 AS codigo, vrp.B0103 AS resposta),
        STRUCT(4 AS codigo, vrp.B0104 AS resposta),
        STRUCT(5 AS codigo, vrp.B0105 AS resposta),
        STRUCT(6 AS codigo, vrp.B0106 AS resposta)
    ]) associacao
JOIN
    `tcm modulo03.PNAD_COVID.TiposDoencas` td
ON
    associacao.codigo = td.int64_field_0
WHERE
    vrp.A003 = 1 -- Filtrar apenas homens
    AND associacao.resposta = 1 -- Respondeu "Sim" para a doença
GROUP BY
    tipo_doenca
ORDER BY
    total_casos DESC;

```

Mulher

```

CREATE OR REPLACE VIEW `tcm modulo03.PNAD_COVID.mulheres_com_doencas` AS
SELECT
    'Mulher' AS sexo,
    CASE
        WHEN td.string_field_1 LIKE '%Doenças docora%' THEN 'Doenças do Coração' --
        Alteração para abranger possíveis variações de codificação
        ELSE td.string_field_1
    END AS tipo_doenca,
    COUNT(*) AS total_casos,
    COUNT(*) * 100.0 / SUM(COUNT(*)) OVER () AS percentual
FROM
    `tcm modulo03.PNAD_COVID.view_resultados_positivos` vrp
CROSS JOIN UNNEST([
    STRUCT(1 AS codigo, vrp.B0101 AS resposta),
    STRUCT(2 AS codigo, vrp.B0102 AS resposta),
    STRUCT(3 AS codigo, vrp.B0103 AS resposta),
    STRUCT(4 AS codigo, vrp.B0104 AS resposta),
    STRUCT(5 AS codigo, vrp.B0105 AS resposta),
    STRUCT(6 AS codigo, vrp.B0106 AS resposta)
]) associacao
JOIN
    `tcm modulo03.PNAD_COVID.TiposDoencas` td
ON
    associacao.codigo = td.int64_field_0
WHERE
    vrp.A003 = 2 -- Filtrar apenas mulheres
    AND associacao.resposta = 1 -- Respondeu "Sim" para a doença
GROUP BY
    tipo_doenca
ORDER BY
    total_casos DESC;

```

Possuem trabalho

```

CREATE OR REPLACE VIEW `tcm modulo03.PNAD_COVID.view_emprego_positivos` AS
SELECT
    CASE

```

```

    WHEN C001 = 1 THEN 'Trabalha (trabalhou ou fez bico na última semana)'
    WHEN C001 = 2 THEN 'Não trabalha'
    WHEN C001 IS NULL THEN 'Não se aplica'
    ELSE 'Não aplicável'
END AS status_emprego,

CASE
    WHEN C001 = 1 AND C006 = 1 THEN 'Trabalha em mais de um emprego'
    WHEN C001 = 1 AND C006 = 2 THEN 'Trabalha em apenas um emprego'
    WHEN C001 = 2 THEN 'Não trabalha'
    WHEN C001 IS NULL THEN 'Não se aplica'
    ELSE 'Não aplicável'
END AS tipo_emprego,

COUNT(*) AS total_pessoas,
COUNT(*) * 100.0 / SUM(COUNT(*)) OVER () AS percentual
FROM
    `tcmodulo03.PNAD_COVID.view_resultados_positivos`
GROUP BY
    status_emprego, tipo_emprego
ORDER BY
    status_emprego, tipo_emprego;

```

Possui plano de saúde

```

CREATE OR REPLACE VIEW `tcmodulo03.PNAD_COVID.view_planosaude_positivos` AS
SELECT
    CASE
        WHEN B007 = 1 THEN 'Sim'
        WHEN B007 = 2 THEN 'Não'
        WHEN B007 = 9 THEN 'Ignorado'
        ELSE 'Não aplicável'
    END AS possui_plano_saude,
    COUNT(*) AS total_pessoas,
    COUNT(*) * 100.0 / SUM(COUNT(*)) OVER () AS percentual
FROM
    `tcmodulo03.PNAD_COVID.view_resultados_positivos`
GROUP BY
    possui_plano_saude
ORDER BY
    possui_plano_saude;

```

Essas pessoas desempregadas tentaram algum tipo de empréstimo, elas conseguiram?

```

CREATE OR REPLACE VIEW `tcmodulo03.PNAD_COVID.view_emprestimo_desempregados` AS
SELECT
    CASE
        WHEN E001 = 1 THEN 'Sim, e pelo menos um morador conseguiu'
        WHEN E001 = 2 THEN 'Sim, mas nenhum morador conseguiu'
        WHEN E001 = 3 THEN 'Não solicitou'
        ELSE 'Não aplicável'
    END AS situacao_emprestimo,
    COUNT(*) AS total_pessoas,
    COUNT(*) * 100.0 / SUM(COUNT(*)) OVER () AS percentual
FROM

```

```

        `tcmodulo03.PNAD_COVID.view_resultados_positivos`
WHERE
    C001 = 2 -- Filtrar apenas os que não trabalham
GROUP BY
    situacao_emprestimo
ORDER BY
    situacao_emprestimo;

```

Qual a renda em média

Query para média geral

```

SELECT
    AVG(view_resultados_positivos.C011A12) AS MediaSalario
FROM
    `tcmodulo03.PNAD_COVID.view_resultados_positivos` AS view_resultados_positivos
WHERE
    view_resultados_positivos.C011A12 IS NOT NULL;

```

Query para segmentação por faixa

```

SELECT
    view_resultados_positivos.C011A11 AS FaixaSalarial,
    MAX(view_resultados_positivos.C011A12) AS MaxSalario,
    MIN(view_resultados_positivos.C011A12) AS MinSalario,
    AVG(view_resultados_positivos.C011A12) AS MediaSalarial,
    COUNT(*) AS QuantidadePessoas
FROM
    `tcmodulo03.PNAD_COVID.view_resultados_positivos` AS view_resultados_positivos
JOIN
    `tcmodulo03.PNAD_COVID.faixaSalarial` AS faixaSalarial
ON
    view_resultados_positivos.C011A11 = faixaSalarial.FaixaSalarial
WHERE
    view_resultados_positivos.C011A12 IS NOT NULL
GROUP BY
    view_resultados_positivos.C011A11
ORDER BY
    QuantidadePessoas DESC;

```

Qual foi o local que buscaram ajuda

-- Criando uma view para consolidar os dados de locais de ajuda

```

CREATE OR REPLACE VIEW `tcmodulo03.PNAD_COVID.view_busca_ajuda` AS
SELECT
    CASE
        WHEN B0041 = 1 THEN 'Posto de Saúde/UBS/ESF'
        WHEN B0042 = 1 THEN 'Pronto Socorro do SUS/UPA'
        WHEN B0043 = 1 THEN 'Hospital do SUS'
        WHEN B0044 = 1 THEN 'Ambulatório ou Consultório Privado/Forças Armadas'
        WHEN B0045 = 1 THEN 'Pronto Socorro Privado/Forças Armadas'
        WHEN B0046 = 1 THEN 'Hospital Privado/Forças Armadas'
        ELSE 'Outros/Não Aplicável'
    END AS local_ajuda,

```

```

COUNT(CASE WHEN B0041 = 1 THEN 1 END) AS total_posto_saude,
COUNT(CASE WHEN B0042 = 1 THEN 1 END) AS total_sus_upa,
COUNT(CASE WHEN B0043 = 1 THEN 1 END) AS total_hospital_sus,
COUNT(CASE WHEN B0044 = 1 THEN 1 END) AS total_ambulatorio_privado,
COUNT(CASE WHEN B0045 = 1 THEN 1 END) AS total_pronto_socorro_privado,
COUNT(CASE WHEN B0046 = 1 THEN 1 END) AS total_hospital_privado,
COUNT(*) AS total_geral
FROM
    `tcmodulo03.PNAD_COVID.view_resultados_positivos`
GROUP BY
    local_ajuda;

```

Quantos internados? E quantos Precisaram de respirador

```

-- Criando uma view para calcular os dados de internação e uso de
ventilador/entubação
CREATE OR REPLACE VIEW `tcmodulo03.PNAD_COVID.view_internacao_ventilacao` AS
SELECT
    -- Dados sobre internação
    CASE
        WHEN B005 = 1 THEN 'Sim, ficou internado'
        WHEN B005 = 2 THEN 'Não ficou internado'
        WHEN B005 = 3 THEN 'Não foi atendido'
        WHEN B005 = 9 THEN 'Ignorado'
        ELSE 'Não aplicável'
    END AS status_internacao,
    COUNT(CASE WHEN B005 = 1 THEN 1 END) AS total_internados,
    COUNT(CASE WHEN B005 = 2 THEN 1 END) AS total_nao_internados,
    COUNT(CASE WHEN B005 = 3 THEN 1 END) AS total_nao_atendidos_internacao,
    COUNT(CASE WHEN B005 = 9 THEN 1 END) AS total_ignorados_internacao,
    -- Dados sobre entubação ou ventilador
    CASE
        WHEN B006 = 1 THEN 'Sim, utilizou entubação/ventilador'
        WHEN B006 = 2 THEN 'Não utilizou entubação/ventilador'
        WHEN B006 = 3 THEN 'Não foi atendido'
        WHEN B006 = 9 THEN 'Ignorado'
        ELSE 'Não aplicável'
    END AS status_ventilacao,
    COUNT(CASE WHEN B006 = 1 THEN 1 END) AS total_usaram_ventilador,
    COUNT(CASE WHEN B006 = 2 THEN 1 END) AS total_nao_usaram_ventilador,
    COUNT(CASE WHEN B006 = 3 THEN 1 END) AS total_nao_atendidos_ventilacao,
    COUNT(CASE WHEN B006 = 9 THEN 1 END) AS total_ignorados_ventilacao,
    COUNT(*) AS total_geral
FROM
    `tcmodulo03.PNAD_COVID.view_resultados_positivos`
GROUP BY
    status_internacao, status_ventilacao;

```

Negativos

Qual foi a atitude perante a crise

```

SELECT
    DB.ANO,
    DB.MES,

```

```

SUM(CASE WHEN CAST(DB.B011 AS INT64) = 1 THEN 1 ELSE 0 END) AS Nao_fez_restricao,
SUM(CASE WHEN CAST(DB.B011 AS INT64) = 2 THEN 1 ELSE 0 END) AS Reduziu_contato,
SUM(CASE WHEN CAST(DB.B011 AS INT64) = 3 THEN 1 ELSE 0 END) AS
Ficou_em_casa_necessidade,
SUM(CASE WHEN CAST(DB.B011 AS INT64) = 4 THEN 1 ELSE 0 END) AS
Ficou_rigorosamente_isolado,
COUNT(*) AS Total_Respostas
FROM
`tcmodulo03.PNAD_COVID.view_resultados_negativos` DB
WHERE
CAST(DB.B011 AS INT64) IN (1, 2, 3, 4)
GROUP BY
DB.ANO,
DB.MES
ORDER BY
DB.ANO,
DB.MES;

```