

PNAD Covid-19

Projeto analítico

Victoria Costi Menezes | RM361292

Desenvolvido para a faculdade FIAP

PNAD Covid-19



Objetivo





Analisaremos os dados coletados pela pesquisa PNAD Covid-19 para entender o impacto da pandemia, e identificar fraquezas a serem remediadas pelo hospital fictício a qual esse material é entregue.

O projeto traz como foco aspectos socioeconômicos e regionais, visando identificar padrões como a Covid afetou e foi abordada nessas divisões.

Utilizamos as tecnologias AWS, e Apache Spark via Python. As visualizações foram feitas com o PowerBI.

github.com/victoria-menezes/fiap-data-analytics-TC3

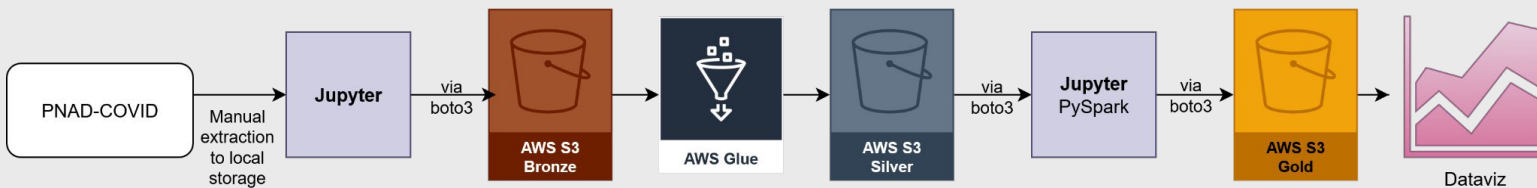
Etapas

- 1  Extração dos dados de **3** meses
- 2  Tratamento dos dados e definição da schema
- 3  Tornando os dados legíveis utilizando os dicionários providenciados, e escolhendo as perguntas
- 4  Análise visual

PNAD Covid-19



Arquitetura



O notebook está disponível no [GitHub](#).

Foram utilizados dados de **Setembro, Outubro e Novembro**, totalizando 1.15 milhões de registros. Esses registros foram unidos em uma só base pelo AWS Glue; perguntas diferentes entre os questionários de Setembro-Outubro e Novembro foram descartadas.

PNAD Covid-19



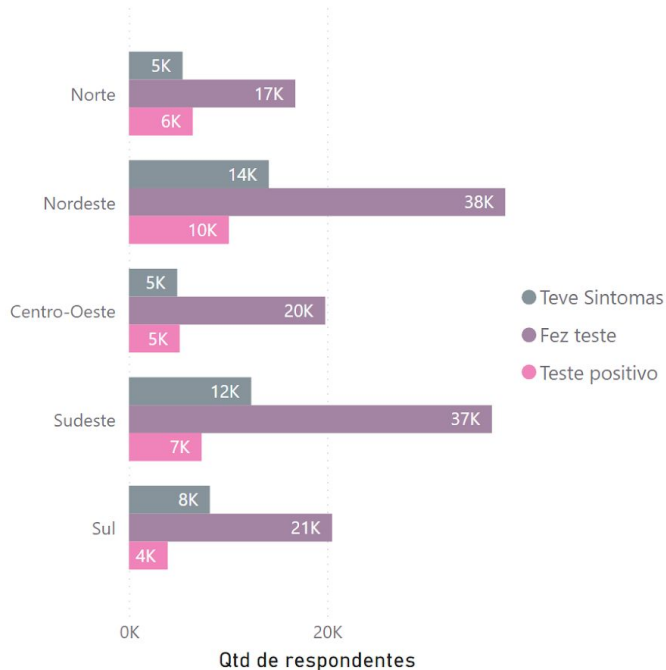
Arquitetura

Seguem as perguntas escolhidas

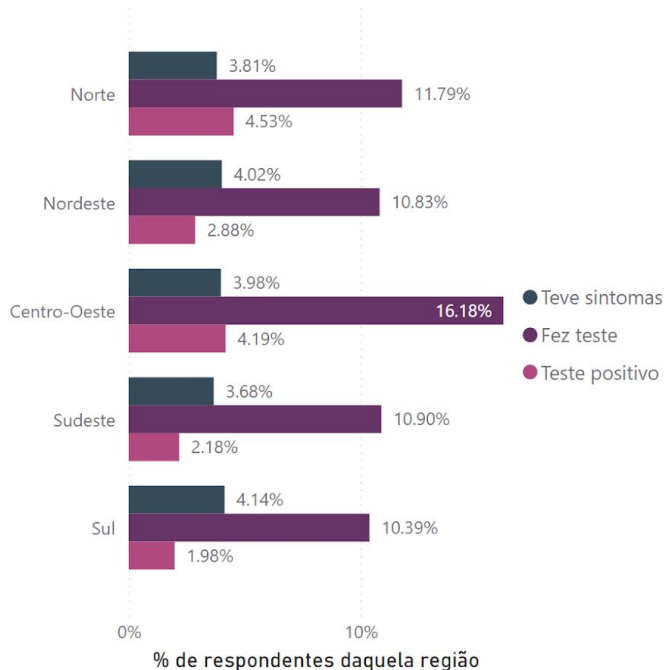
- UF Unidade da Federação
- V1022 Situação de Domicílio
- A002 Qual sua idade?
- A004 Cor ou raça
- A005 Qual sua escolaridade?
- B001 Na semana passada, o(a) Sr(a) teve (sintomas)?
- B002 Por causa (de sintomas), foi a algum estabelecimento de saúde?
- B004 Em que local(is) buscou o atendimento?
- B007 O(A) Sr(a) tem algum plano de saúde médico, seja particular, de empresa ou de órgão público?
- B009B (SWAB) Qual o resultado?
- B009D (Furo no dedo) Qual o resultado?
- B009F (Veia no braço) Qual o resultado?
- B010 Algum médico já lhe deu o diagnóstico de alguma dessas doenças?
- B011 Na semana passada, devido à pandemia do Coronavírus, em que medida o(a) Sr(a) restringiu o contato com as pessoas?
- C002 Na semana passada, estava temporariamente afastado de algum trabalho?
- C004 O(A) Sr(a) continuou a ser remunerado (mesmo que parcialmente) por esse trabalho?
- C007 No trabalho (único ou principal) que tinha nessa semana, o(a) Sr(a) era:
- C010 Número de faixa do rendimento / retirada em dinheiro (NORMALMENTE)
- C011 Número de faixa do rendimento / retirada em dinheiro (NO MES ATUAL)
- C013 Na semana passada, o(a) Sr(a) estava em trabalho remoto (home office ou teletrabalho)? (C013)

A apresentação visual a seguir pode também ser visualizada pelo [PowerBI](#)

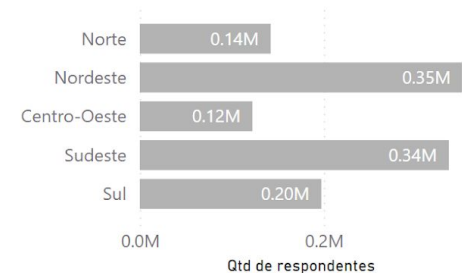
1. Sintomas, testes feitos e positivos por região



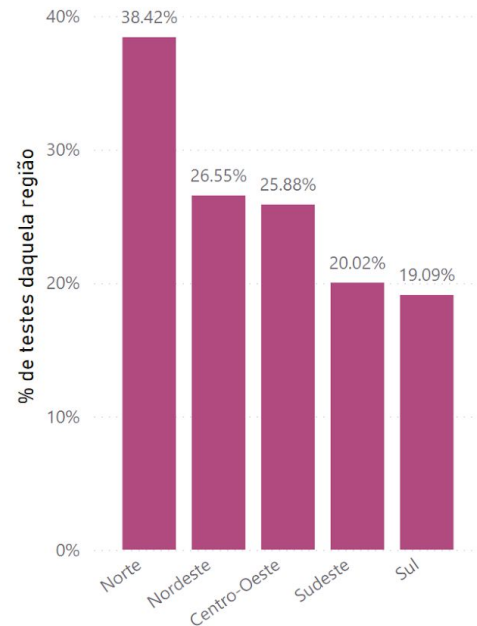
2. Sintomas, testes feitos e positivos por região



3. Quantidade de respondentes por região



4. Testes feitos que retornaram positivos

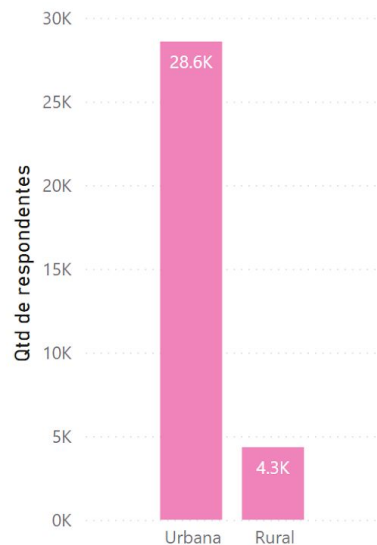


A **Covid-19** afetou todas as regiões, em números brutos sendo mais prevalente nas regiões com a maior quantidade de respondentes (**figuras 1 e 3**). Para melhor analisar como a doença se comportou em diferentes regiões e grupos, faremos nossas análises **em proporção**.

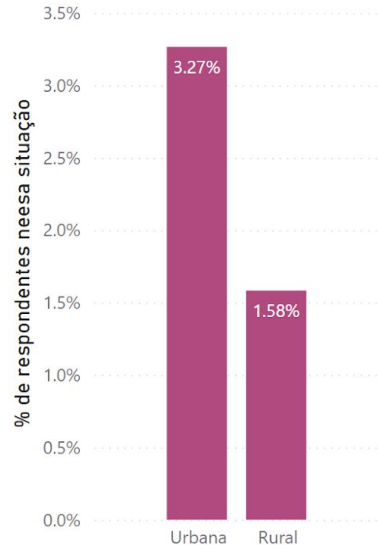
Com essa perspectiva, vemos pela **figura 2** que a **região Norte** foi a que teve a **maior proporção de casos de Covid confirmados** nesse estudo, tendo 4.5% de testes positivos com aproximadamente a mesma proporção de **testes feitos** a outras regiões (~10% dos respondentes da região fizeram o teste).

Observando a **figura 4**, que denota **quantos dos testes feitos foram positivos**, vemos um padrão similar: **a região Norte foi a mais proporcionalmente afetada**. A região centro-oeste, com a maior proporção de testes feitos (16%), é a terceira mais afetada.

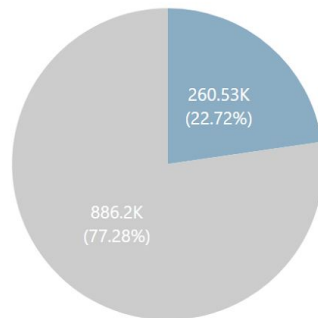
1. Testes positivos por situação domiciliar



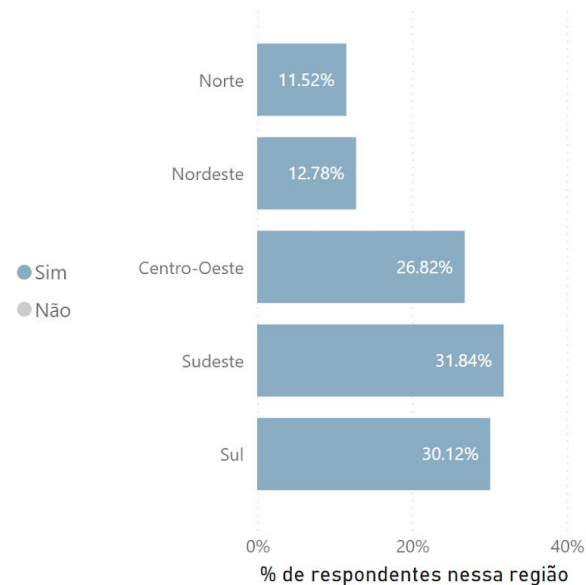
2. Testes positivos por situação domiciliar



3. Tem plano de saúde?



4. Plano de saúde por região



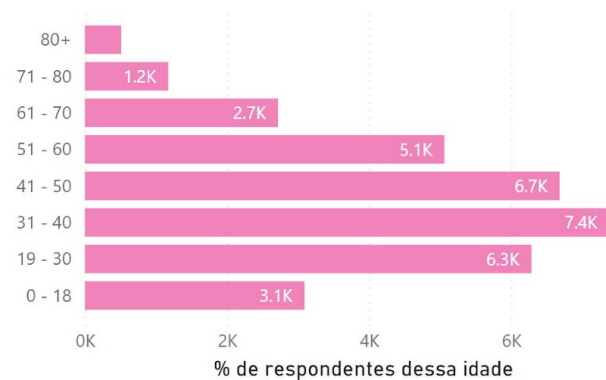
Finalmente, as **áreas urbanas** são as mais afetadas pelo Covid-19.

Isso se aplica em **quantidade bruta (figura 1)** e em **proporção (figura 2)**.

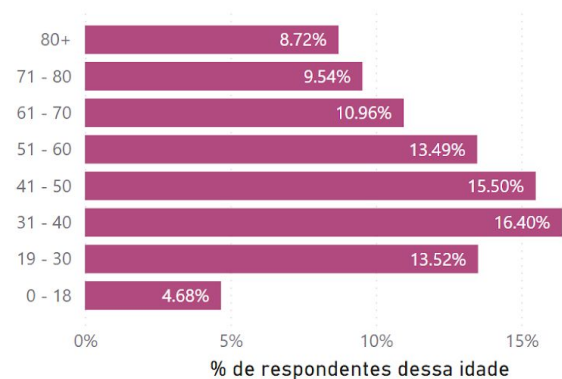
77.3% dos respondentes dizem não ter plano de saúde (**figura 3**). Isso significa que uma majoritária parcela da população **depende do SUS**.

Além disso, as **regiões com maior proporção de testes positivos** são, também as **regiões que menos tem plano de saúde (figura 4)**.

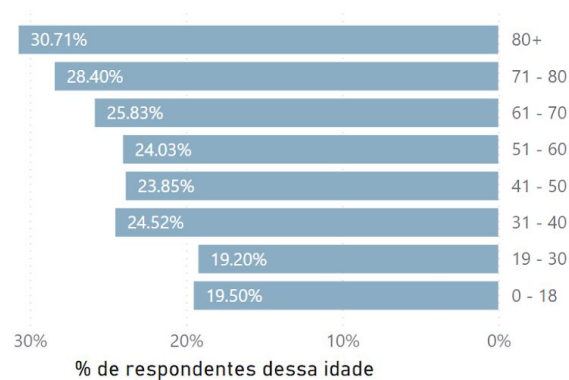
1. Testes positivos por idade



2. Testes positivos por idade



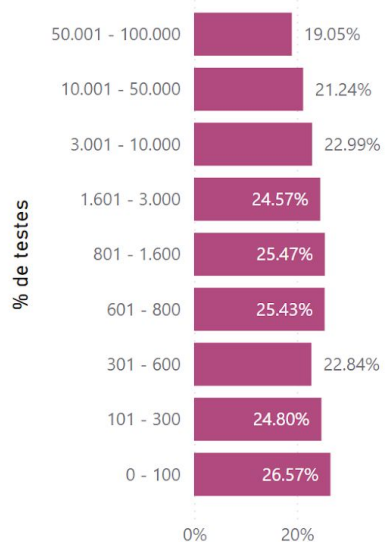
3. Plano de saúde por idade



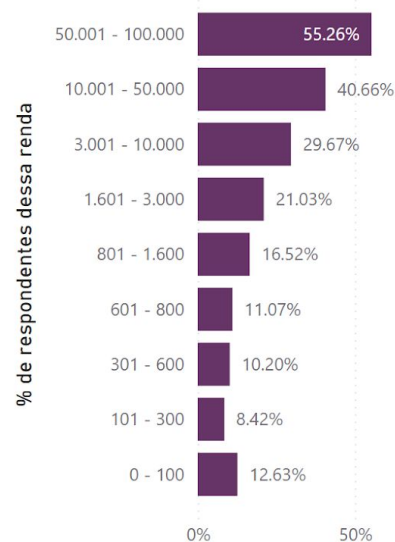
Em **quantidade bruta (figura 1)** e **proporção (figura 2)**, pessoas na **faixa etária trabalhadora retornam mais testes positivos**.

Além disso, **respondentes mais jovens tem menos chance de ter plano de saúde**.

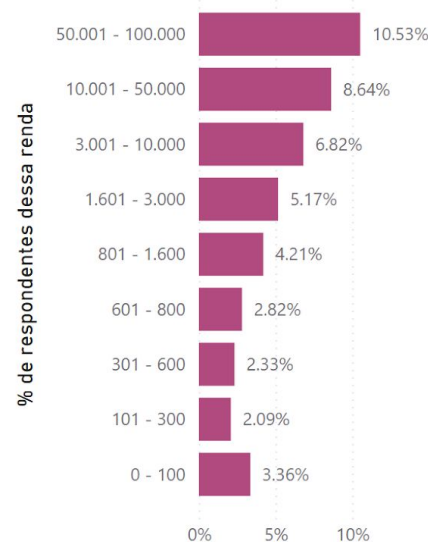
1. Testes feitos positivos por renda



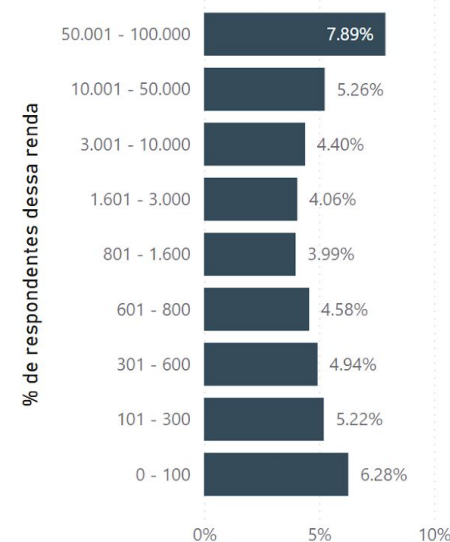
2. Testes feitos por renda



3. Testes positivos por renda



4. Sentimento de sintomas por renda

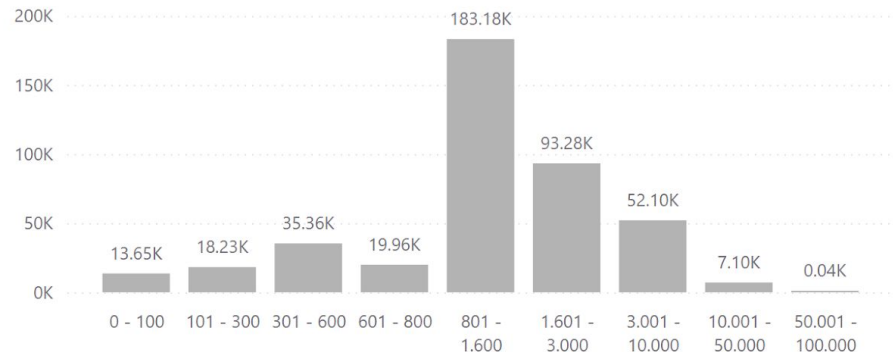


Analisando agora a distribuição **por renda**, se observa pela **figura 1** que estatisticamente, testes feitos retornam taxas similares (20 - 25%) de resultados positivos.

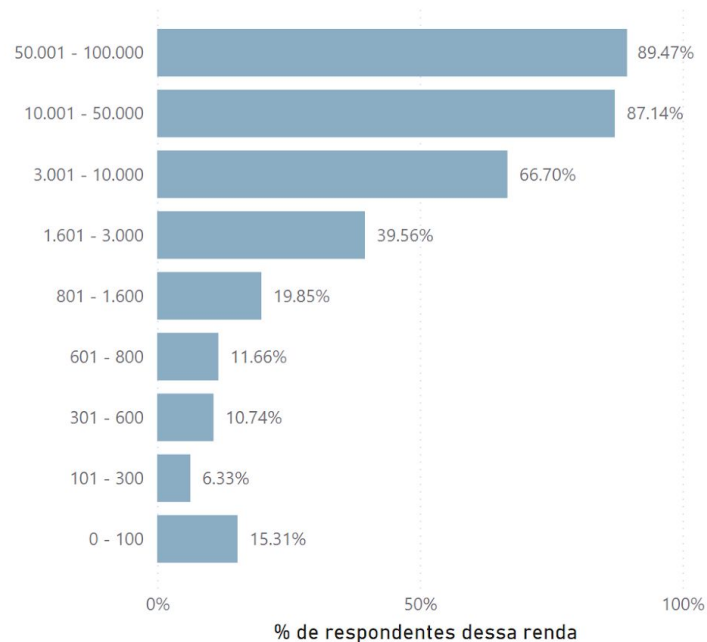
Porém, também se observa pelas **figuras 2 e 4** que respondentes com rendas menores fazem o teste em menor proporção, **mesmo sentindo sintomas na mesma, ou maior, proporção que as faixas de renda maiores.**

Há portanto uma **possibilidade de que há uma maior incidência de Covid-19 nessas rendas** que não está aqui registrada.

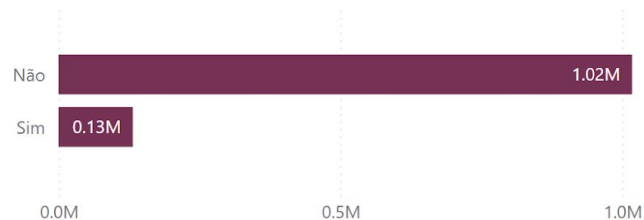
Quantidade de respondentes por renda



1. Plano de saúde por renda



2. Quando tiveram sintomas, os respondentes fizeram um teste COVID?



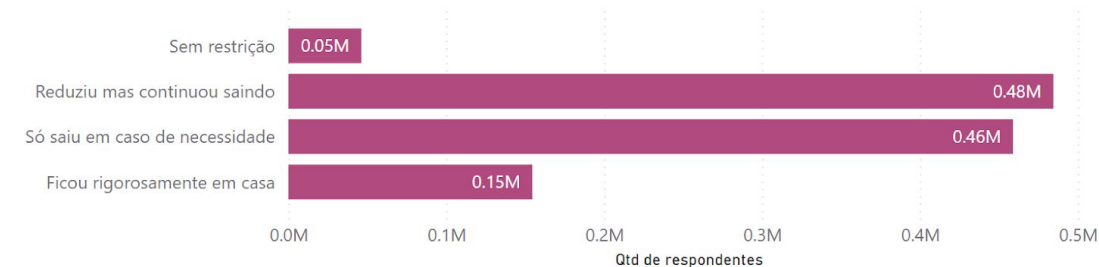
Via **figura 1**, planos de saúde são mais prevalentes em faixas de renda mais elevadas. Isso acarreta em uma **maior dependência no SUS nas faixas mais baixas**.

Via **figura 2**, no geral, **os respondentes não tiveram um forte hábito de fazer o teste COVID**, mesmo com sintomas. Esse fato novamente nos alerta ao risco de estatísticas não completamente fiéis à realidade.

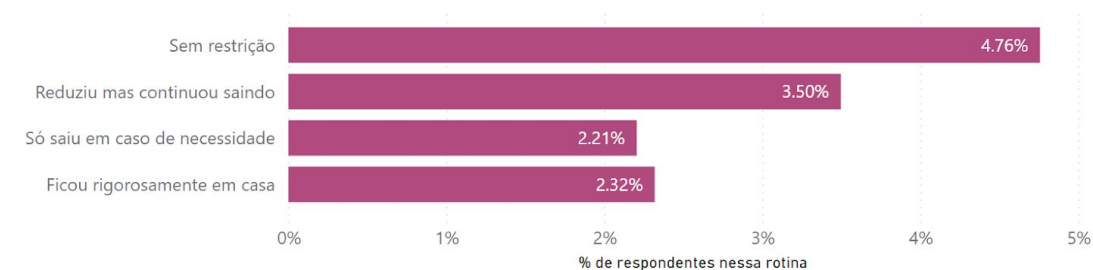
Também não foram todos que aderiram às recomendações de distanciamento, via **figura 3**.

Isso teve o resultado esperado, via **figura 4**: **pessoas que não restringiram seu contato com outras retornaram maior taxa de testes positivos**.

3. Aderência à cada rotina

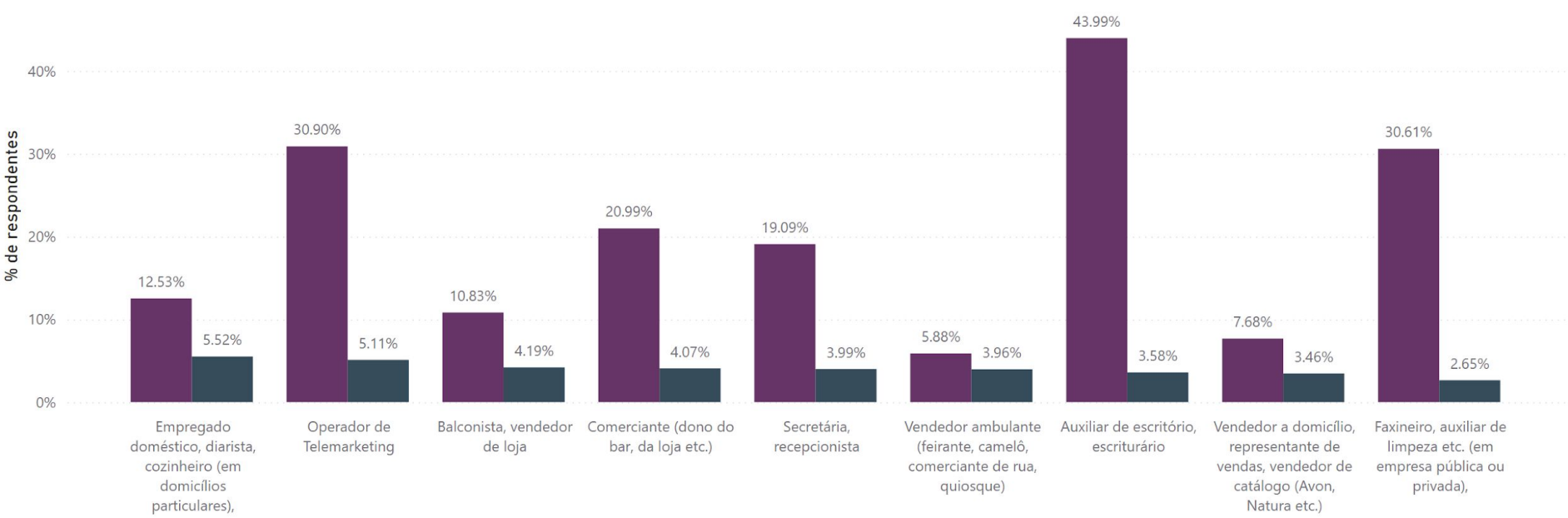


4. Testes positivos por rotina



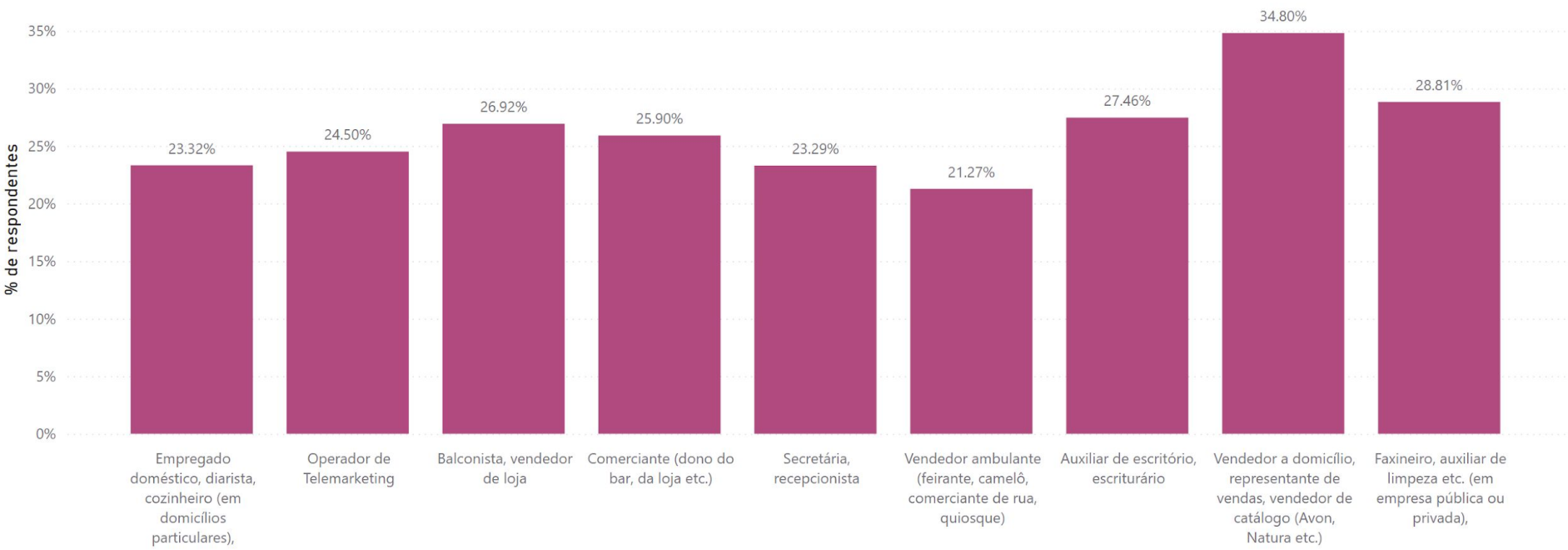
1. Testes feitos e sintomas identificados por função no trabalho

● Fez teste ● Teve sintomas



Certas profissões estavam mais expostas ao COVID. Via **figura 1**, As profissões que mais auto-identificaram sintomas **não foram** as que mais aplicaram os testes de Covid, novamente criando uma discrepância entre sintomas e casos confirmados. **Auxiliares de escritório** foram os que mais aplicaram o teste mesmo tendo uma relativa baixa incidência de sintomas. Isso pode apontar para uma iniciativa das empresas nesse setor, ou a uma iniciativa específica desses trabalhadores.

1. Quantas pessoas sentiram sintomas por função no trabalho



Entre casos confirmados por teste, via **figura 1** os vendedores ambulantes foram os menos afetados pelo Covid, possivelmente devido ao ambiente aberto onde trabalham. Em contraste, **vendedores a domicílio**, que encontram com diversas pessoas em ambientes fechados, foram os mais afetados.

PNAD Covid-19



Conclusões

- Enquanto as regiões mais numericamente afetadas foram as mais populosas, **a região Norte em particular foi afetada desproporcionalmente pela pandemia;**
- Áreas **urbanas** foram mais afetadas pela Covid, **em números brutos E proporcionalmente;**
- Mais de 75% da população **não conta como plano de saúde;**
- Aqueles que não tem plano de saúde **tendem a ser das menores faixas de renda;**
- Como tendência geral, as demografias mais proporcionalmente afetadas pela Covid são também as que não tendem ter planos de saúde;
- Para fins estatísticos, **a baixa execução dos diversos testes COVID, mesmo com a presença de sintomas, diminui o nível de confiança de certos dados;**
- Certas profissões estão a mais risco da Covid que outras;

PNAD Covid-19



Recomendações

- Responder de acordo com as demografias da região na qual o hospital se situa; no caso de um novo surto, é provável que ao menos inicialmente, **as proporções de contaminação de mantenham;**
- **Planejamentos de rota pela área urbana e possível construção de postos pela cidade** caso haja a necessidade, podendo assim atender a população em suas residências quando possível e **evitando o superlotamento do hospital principal;**
- Considerar **parceria entre o SUS e os grandes hospitais privados**, seja em compartilhamento de infraestrutura ou redirecionamento de pacientes; a maior parte da população não conta com seguro de saúde e portanto a distribuição de pacientes é desproporcional entre o setor privado e público. Isso pode acarretar em mais casos quando pessoas desistem de ir ao hospital ou recebem cuidado inadequado; uma parceria seria portanto benéfica a todos;
- Desenvolvimento de uma campanha conscientizante **para empresas e indivíduos**. Indivíduos precisam saber o guia básico de prevenção (distanciamento, máscaras, etc) e empresas precisam **permitir que esse guia seja seguido;**
- Possível parceria **com empresas privadas para de alguma maneira incentivar o item acima;**
- Fornecimento acessível de **testes de Covid**, importante para fins estatísticos;

PNAD Covid-19



Links

- github.com/victoria-menezes/fiap-data-analytics-TC3
- <https://app.powerbi.com/view?r=eyJrIjoieWRmMmY5NzQtZWl0Zi00MzUyLTgzY2QtMDM1NTE5YmVkNGE5IiwidCI6ImExZGJiZmUyLTg5YjgtNDU0OS1iZTEwLWNiYzM2NGU1OTU1MSIsImMiOiR9>