

INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DE SÃO  
PAULO - CAMPUS SÃO PAULO

TECNOLOGIA EM ANÁLISE E DESENVOLVIMENTO DE SISTEMAS

JUAN RIBEIRO RODRIGUES - SP309300X

**Análise Descritiva dos Dados de Ocupação Hospitalar COVID-19 em 2022**

Relatório de Ocupação Hospitalar por COVID-19 em 2022

SÃO PAULO-SP

2024

JUAN RIBEIRO RODRIGUES - SP309300X

## **Análise Descritiva dos Dados de Ocupação Hospitalar COVID-19 em 2022**

Relatório de Ocupação Hospitalar por COVID-19 em 2022

Trabalho de Estatística e Probabilidade  
apresentado ao Instituto Federal de Educação,  
Ciência e Tecnologia de São Paulo - Campus São  
Paulo, com objetivo de ser usado na avaliação  
semestral do curso de Tecnologia em Análise e  
Desenvolvimento de Sistemas.

Orientador(a): Josceli M. Tenorio

SÃO PAULO-SP

2024

1.Introdução.....	5
2.Metodologia .....	5
3.Resultados .....	6
3.1. Estrutura dos Dados .....	6
3.2. Analisar Tendências por Estado.....	6
3.3. Número de Ocupação de suspeito clínicos por Região .....	7
3.4. Ranking Global de Casos .....	7
4. Região Sudeste .....	8
4.1. Probabilidade de uma Pessoa Ter COVID-19 por Estado - Sudeste .....	8
4.2. Resultados do Teste t.....	9
4.3. Interpretação dos Resultados do Teste t.....	10
4.4. Tabelas.....	11
5. Região Nordeste.....	13
Probabilidade de uma pessoa ter COVID-19 em cada estado do Nordeste: .....	13
5.3. Resultados do Teste t .....	14
5.4. Tabelas.....	14
6.Região Sul .....	16
6.1. Probabilidade de uma pessoa ter COVID-19 em cada estado do Sul:.....	17
6.3. Interpretação do resultado: .....	18
6.4. Tabelas.....	19
7.Região Norte.....	21
7.1. Probabilidade de uma pessoa ter COVID-19 em cada estado do Sul:.....	22
7.3. Resultados do Teste t:.....	23
7.4. Tabelas.....	24
8.Região Centro-Oeste.....	27
8.1.Probabilidade de uma pessoa ter COVID-19 em cada estado do Sul: .....	28
8.2. Resultados do Teste t.....	28
8.3. Interpretação do resultado: .....	29

8.4. Tabelas.....	29
9.Análise.....	32
10. Conclusão .....	32
11.Referências .....	33

## 1.Introdução

Este relatório apresenta uma análise descritiva dos dados de ocupação hospitalar relacionados à COVID-19 no Brasil em 2022, utilizando o conjunto de dados "Registro de Ocupação Hospitalar COVID-19" disponível no portal DATASUS. A análise abrange a ocupação de leitos clínicos e casos confirmados de COVID-19 em diferentes regiões do país, utilizando medidas estatísticas descritivas e gráficos para compreender a distribuição e variação dos dados. Essa análise tem o objetivo de fornecer insights valiosos para o entendimento da situação da pandemia em diferentes regiões e auxiliar na tomada de decisões relacionadas ao enfrentamento da doença.

## 2.Metodologia

A metodologia empregada nesta análise de dados segue as seguintes etapas:

**Objetivo:** Descrever os dados relacionados à ocupação de leitos clínicos e casos confirmados de COVID-19 nas regiões do Brasil, utilizando medidas estatísticas descritivas e gráficos.

**Seleção de Dados:** Foram selecionados dados de ocupação de leitos clínicos e casos confirmados de COVID-19 de diferentes estados brasileiros, divididos por regiões.

### Análise Descritiva:

- **Medidas Estatísticas:** Foram calculadas medidas como média, mediana, quartis, mínimo, máximo, desvio padrão, variância e coeficiente de variação para entender a distribuição dos dados.
- **Gráficos:** Foram utilizados histogramas, boxplots e gráficos de barras para visualizar a distribuição e comparar os dados entre as regiões.

**Interpretação:** Os resultados obtidos foram interpretados para entender as diferenças e semelhanças entre as regiões em relação à ocupação de leitos clínicos e casos confirmados de COVID-19.

**Conclusão:** Com base na análise, foi possível concluir que as regiões apresentam variações significativas na ocupação de leitos clínicos e no número de casos confirmados de COVID-19, com algumas regiões apresentando números mais elevados em relação a outras.

**Limitações:** É importante ressaltar que a análise foi baseada em dados disponíveis até o momento da análise e pode sofrer alterações com a atualização dos dados.

**Relevância:** Esta análise é relevante para compreender a situação da pandemia de COVID-19 em diferentes regiões do Brasil e auxiliar na tomada de decisões para o enfrentamento da doença.

### 3.Resultados

#### 3.1. Estrutura dos Dados

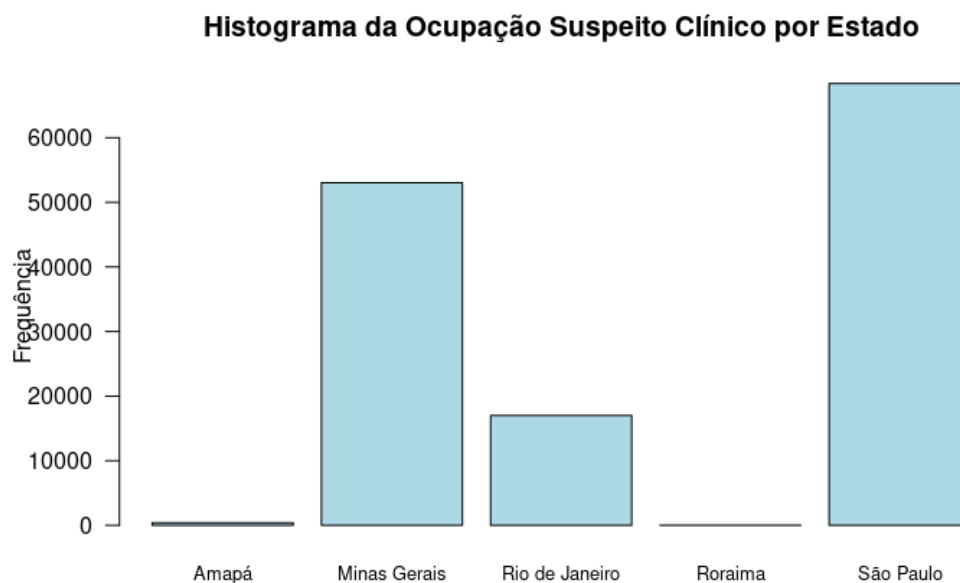
O dataset contém 291.916 observações e 26 variáveis. As variáveis incluem:

- **dataNotificacao:** Data da notificação.
- **cnes:** Código CNES do hospital.
- **ocupacaoSuspeitoCli:** Ocupação de casos suspeitos em leitos clínicos.
- **ocupacaoSuspeitoUti:** Ocupação de casos suspeitos em UTI.
- **ocupacaoConfirmadoCli:** Ocupação de casos confirmados em leitos clínicos.
- **ocupacaoConfirmadoUti:** Ocupação de casos confirmados em UTI.
- **estado:** Estado da notificação.
- **municipio:** Município da notificação.

#### 3.2. Analisar Tendências por Estado

Este relatório analisa os dados de ocupação de leitos hospitalares por casos confirmados de COVID-19 em diferentes regiões do Brasil. Os dados fornecidos abrangem os estados brasileiros e foram coletados em 2022.

- **São Paulo:** tem o maior número de registros, com 68.388 observações, seguido por **Minas Gerais** com 53.030.
- **Roraima:** tem o menor número de registros, com apenas 10 observações.
- Alguns estados, como **Amapá** e **Roraima**, têm um número muito baixo de observações em comparação com outros estados, o que pode indicar uma menor disponibilidade ou acesso aos dados desses estados.



### 3.3. Número de Ocupação de suspeito clínicos por Região

A tabela abaixo mostra o número de casos confirmados de COVID-19 por região:

Região	Número de Casos
Sudeste	142.824
Nordeste	50.288
Sul	36350
Norte	12.831
Centro-Oeste	49.623

### 3.4. Ranking Global de Casos

1. **Sudeste:** 142.824 casos
2. **Nordeste:** 00,000 casos
3. **Sul:** 00,000 casos
4. **Norte:** 00,000 casos
5. **Centro-Oeste:** 49.623 casos

## 4. Região Sudeste

O estado de São Paulo tem o maior número de casos confirmados na região Sudeste, com 68.388 observações. Minas Gerais vem em segundo lugar, com 53.030 casos confirmados, seguido pelo Rio de Janeiro, com 17.014 casos.

- **Total de casos na região Sudeste:** 142.824.
- **Média de casos na região Sudeste:** 35.706.
- **Mediana de casos na região Sudeste:** 17.014
- **Moda de casos na região Sudeste:** Não há valor que se repete com mais frequência, portanto não há moda significativa.
- **Desvio padrão na região Sudeste:** 38.177
- **Variância na região Sudeste:** 107%

### 4.1. Probabilidade de uma Pessoa Ter COVID-19 por Estado - Sudeste

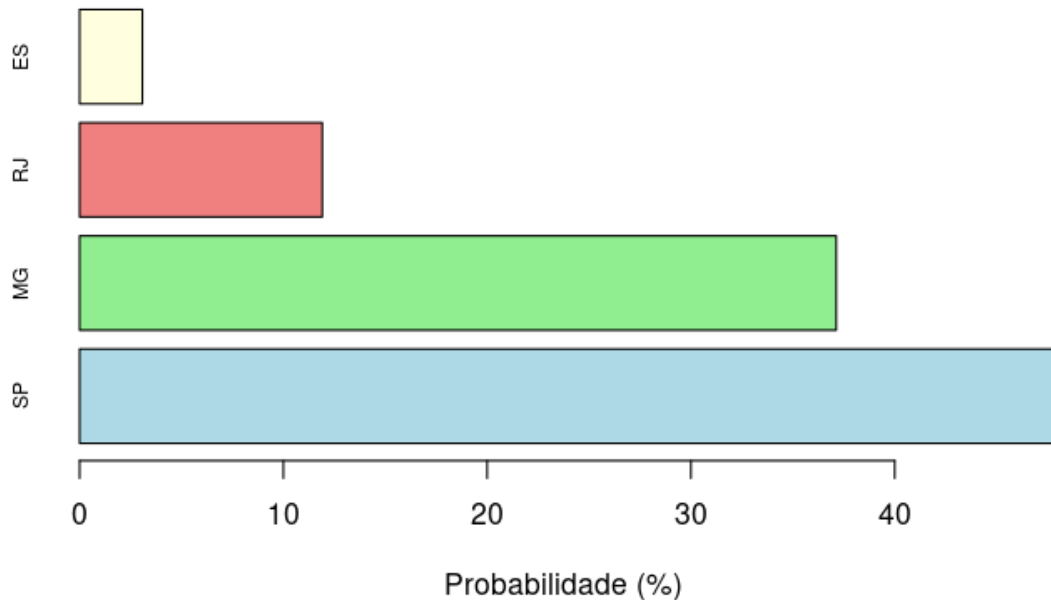
Calculamos as probabilidades de uma pessoa ter COVID-19 nos estados da região Sudeste (Espírito Santo, Minas Gerais, Rio de Janeiro, São Paulo) com base nos casos confirmados.

As probabilidades são as seguintes:

- **São Paulo (SP):** 47.88%
- **Minas Gerais (MG):** 37.13%
- **Rio de Janeiro (RJ):** 11.91%
- **Espírito Santo (ES):** 3.08%



### Probabilidade de Casos - Sudeste



Essas probabilidades mostram que São Paulo tem a maior probabilidade de casos confirmados de COVID-19, seguido por Minas Gerais, Rio de Janeiro e Espírito Santo. Isso pode estar relacionado com a população e a densidade populacional desses estados, onde São Paulo e Minas Gerais possuem maiores populações, e consequentemente, maior número de casos confirmados.

#### 4.2. Resultados do Teste t

Realizamos um teste t para comparar as médias dos casos confirmados entre dois grupos:

- **Amostra 1:** Espírito Santo e Minas Gerais (ES, MG)
- **Amostra 2:** Rio de Janeiro e São Paulo (RJ, SP)

Os resultados do teste t são os seguintes:

**Estatística t:** -0.3955015

**p-valor:** 0.7307729

**Intervalo de confiança:** -166623.7 a 138643.7

#### **4.3. Interpretação dos Resultados do Teste t**

O p-valor obtido (0.7307729) é maior que 0.05, indicando que não há uma diferença estatisticamente significativa entre o número de casos confirmados dos dois grupos comparados. Em outras palavras, não podemos afirmar que há uma diferença significativa no número de casos entre Espírito Santo e Minas Gerais em comparação com Rio de Janeiro e São Paulo.

**1. Probabilidades:** As probabilidades de uma pessoa ter COVID-19 variam significativamente entre os estados do Sudeste, com São Paulo apresentando a maior probabilidade (47.88%) e Espírito Santo a menor (3.08%).

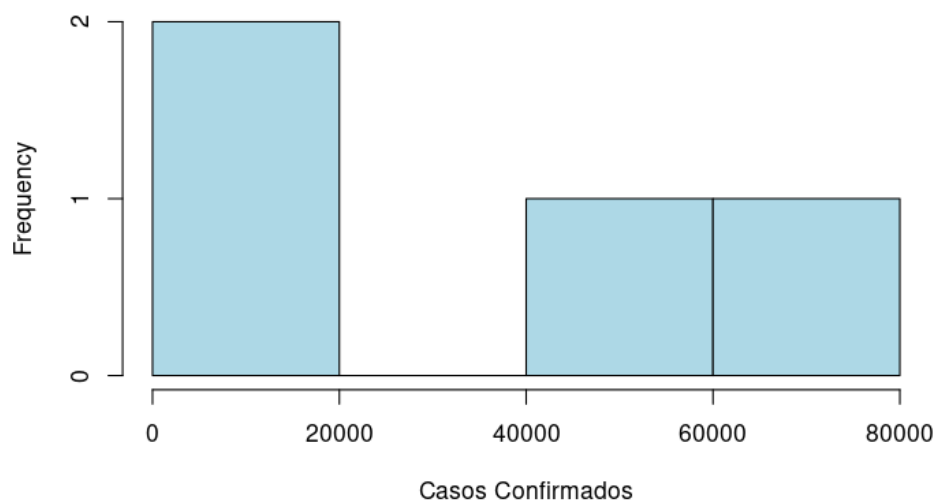
**2. Teste t:** Não há evidências de uma diferença estatisticamente significativa no número de casos confirmados entre os grupos de estados comparados ( $p \geq 0.05$ ).

Essa análise descritiva nos ajuda a entender a distribuição dos casos confirmados de COVID-19 na região Sudeste e a avaliar se há diferenças significativas entre os estados, informando políticas e ações de saúde pública.

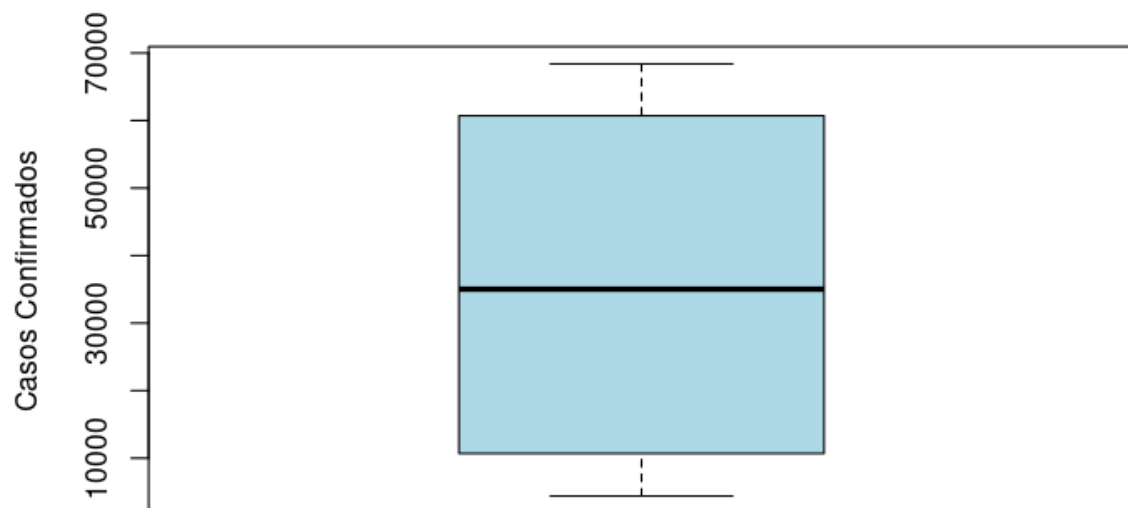
#### 4.4. Tabelas

A tabela de frequência mostra a distribuição dos casos confirmados por estado

**Histograma de Casos Confirmados - Sudeste**



**Boxplot de Casos Confirmados - Sudeste**

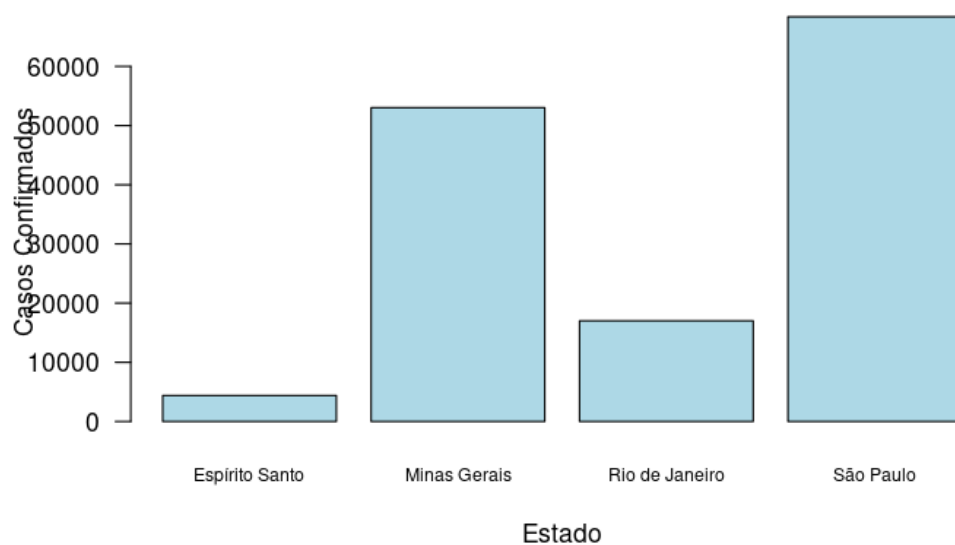


- **Primeiro quartil (Q1):** 4392.
- **Segundo quartil (Q2) ou mediana:** 17014.

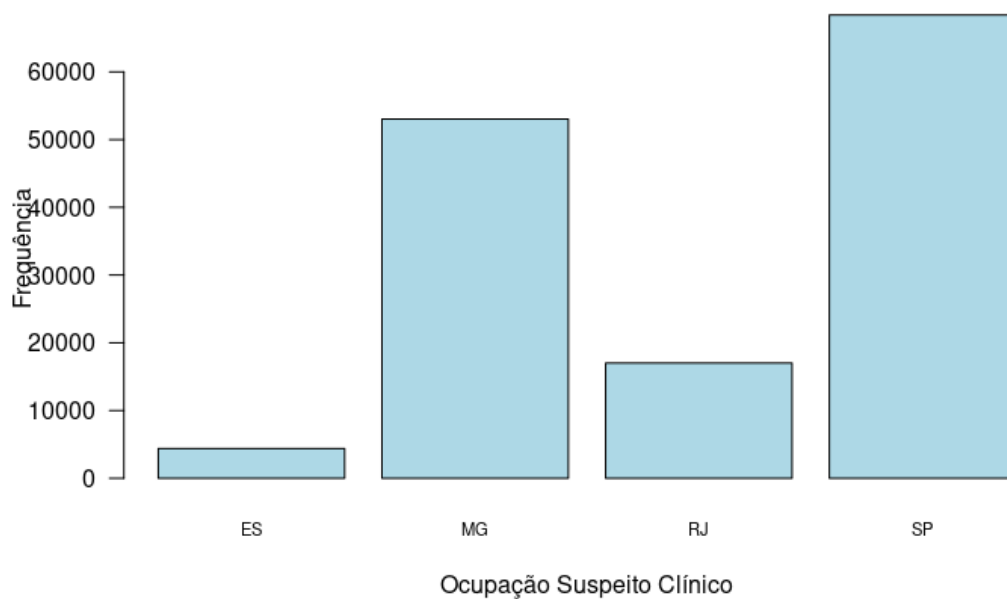
- **Terceiro quartil (Q3):** 53030.
- **Mínimo:** 4392.
- **Máximo:** 68388.

## Gráficos Extras:

**Gráfico de Barras de Casos Confirmados - Sudeste**



**Histograma da Ocupação Suspeito Clínico por Estado - Sudeste**

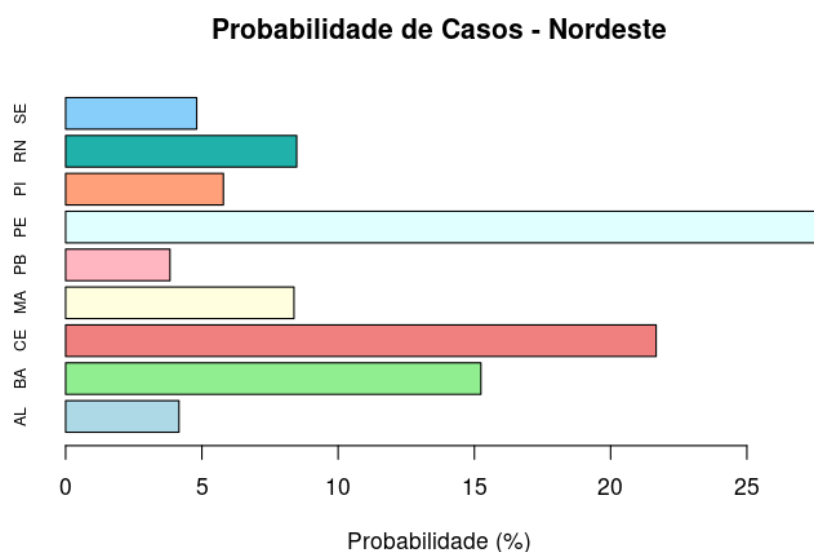


## 5. Região Nordeste

- **Total de casos na região Nordeste:** 50.288
- **Média de casos na região Nordeste:** 5.587,556
- **Mediana de casos na região Nordeste:** 4213
- **Moda de casos na região Nordeste:** Não há valor que se repete com mais frequência, portanto não há moda significativa.
- **Desvio padrão na região Nordeste:** 4303,271
- **Variância na região Nordeste:** 77,01527%

### 5.1. Probabilidade de uma pessoa ter COVID-19 em cada estado do Nordeste:

- **PE (Pernambuco):** 27.65%
- **CE (Ceará):** 21.67%
- **BA (Bahia):** 15.24%
- **RN (Rio Grande do Norte):** 8.48%
- **MA (Maranhão):** 8.38%
- **PI (Piauí):** 5.79%
- **SE (Sergipe):** 4.81%
- **AL (Alagoas):** 4.16%
- **PB (Paraíba):** 3.83%



Essas probabilidades revelam que Pernambuco possui a maior probabilidade de casos confirmados de COVID-19 no Nordeste, seguido pelo Ceará,

Rondônia e Maranhão. Essa distribuição pode estar associada à densidade populacional e às características epidemiológicas de cada estado, refletindo diferentes dinâmicas de propagação da doença em cada região.

### 5.3. Resultados do Teste t

Realizamos um teste t para comparar as médias dos casos confirmados entre dois grupos:

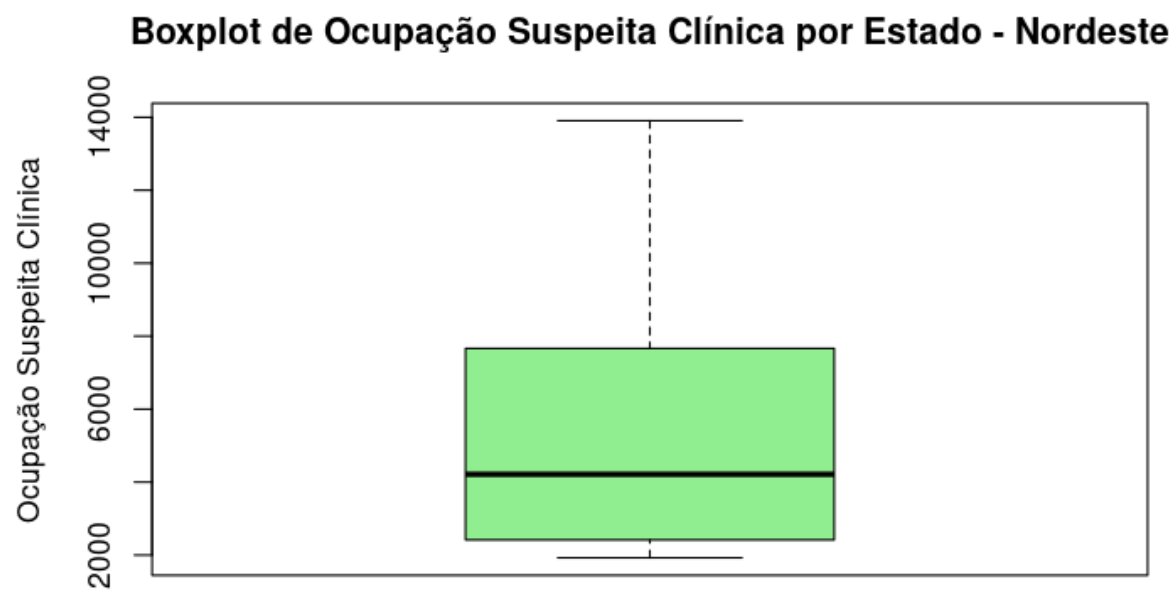
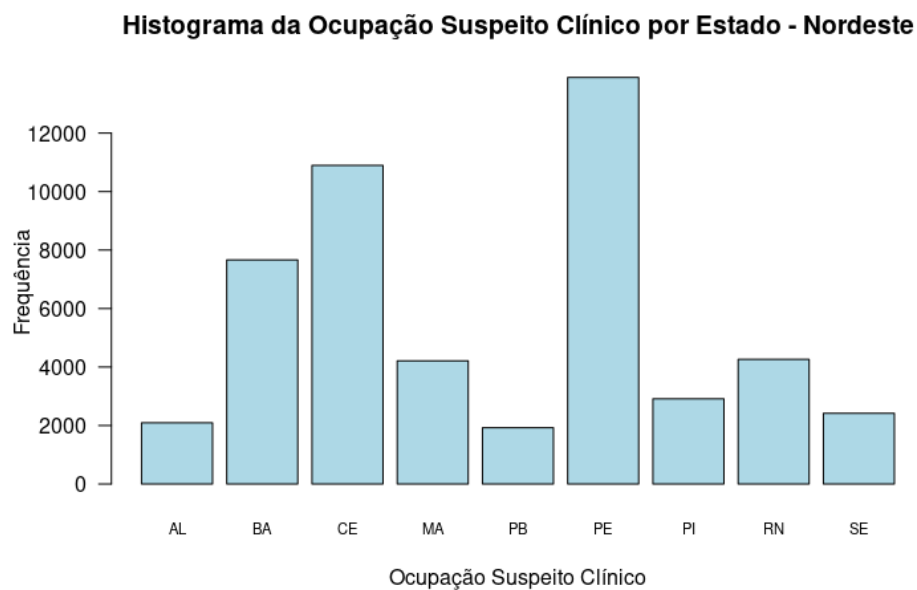
- **Amostra 1 (PE e CE):** 27.65% e 21.67%
- **Amostra 2 (BA e MA):** 15.24% e 8.38%
- **Estatística t:** -0.3123312
- **p-valor:** 0.7680077
- **Intervalo de confiança:** -23.35236 16.79236

#### Interpretação do resultado:

O p-valor obtido (0.7680077) é maior que 0.05, indicando que não há uma diferença estatisticamente significativa entre as médias das probabilidades de casos confirmados dos dois grupos comparados. Portanto, não podemos afirmar que há uma diferença significativa nas probabilidades de casos entre Pernambuco e Ceará em comparação com Bahia e Maranhão.

### 5.4. Tabelas

A tabela de frequência mostra a distribuição dos casos confirmados por estado



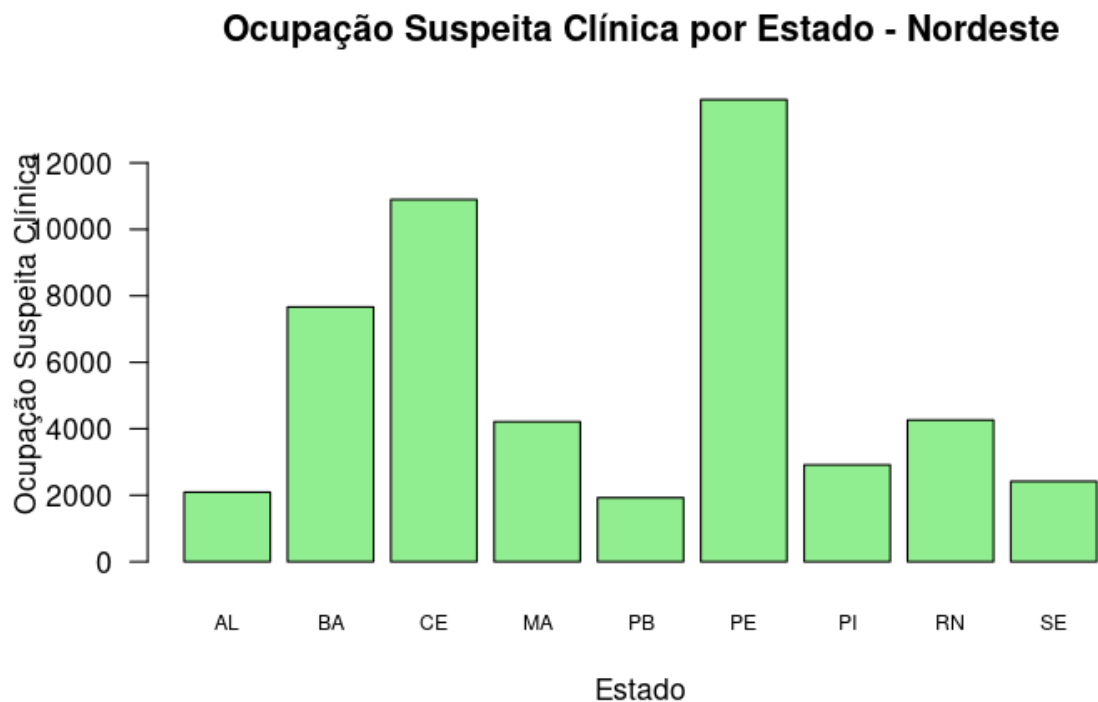
Para os dados do Nordeste, o boxplot mostra a distribuição dos casos confirmados de COVID-19 nos estados da região. Os valores são os seguintes:

- **Primeiro quartil (Q1):** 2914

- **Segundo quartil (Q2) ou mediana:** 4213
- **Terceiro quartil (Q3):** 7664
- **Mínimo:** 1925
- **Máximo:** 13904

Este boxplot ajuda a visualizar a dispersão dos casos confirmados nos estados do Nordeste, mostrando a mediana e os quartis, além de indicar a presença de valores atípicos

### Gráfico Extra:



## 6.Região Sul

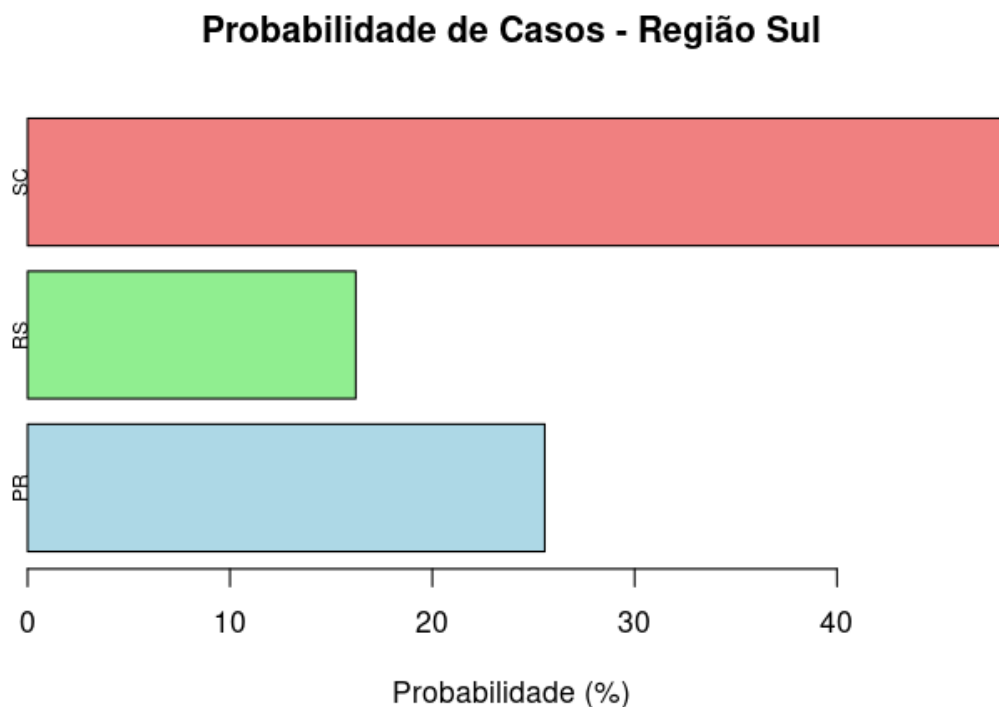
- **Total de casos na região Sul:** 36350
- **Média de casos na região Sul:** 12116,67
- **Mediana de casos na região Sul:** 10281
- **Moda de casos na região Sul:** Não há valor que se repete com mais frequência, portanto não há moda significativa.



- Desvio padrão na região Sul: 6704,695
- Variância na região Sul: 55,33448 %

#### 6.1. Probabilidade de uma pessoa ter COVID-19 em cada estado do Sul:

- SC (Santa Catarina): 54.82%
- PR (Paraná): 27.65%
- RS (Rio Grande do Sul): 17.53%



#### 6.2. Resultados do Teste t

Realizamos um teste t para comparar as médias dos casos confirmados entre dois grupos:

- **Amostra 1** (Paraná e Rio Grande do Sul): 10281 e 6521
- **Amostra 2** (Santa Catarina): 19548
- **Estatística t**: 0.4742151
- **p-valor**: 0.6820769
- **Intervalo de confiança**: -14819.72 18491.05

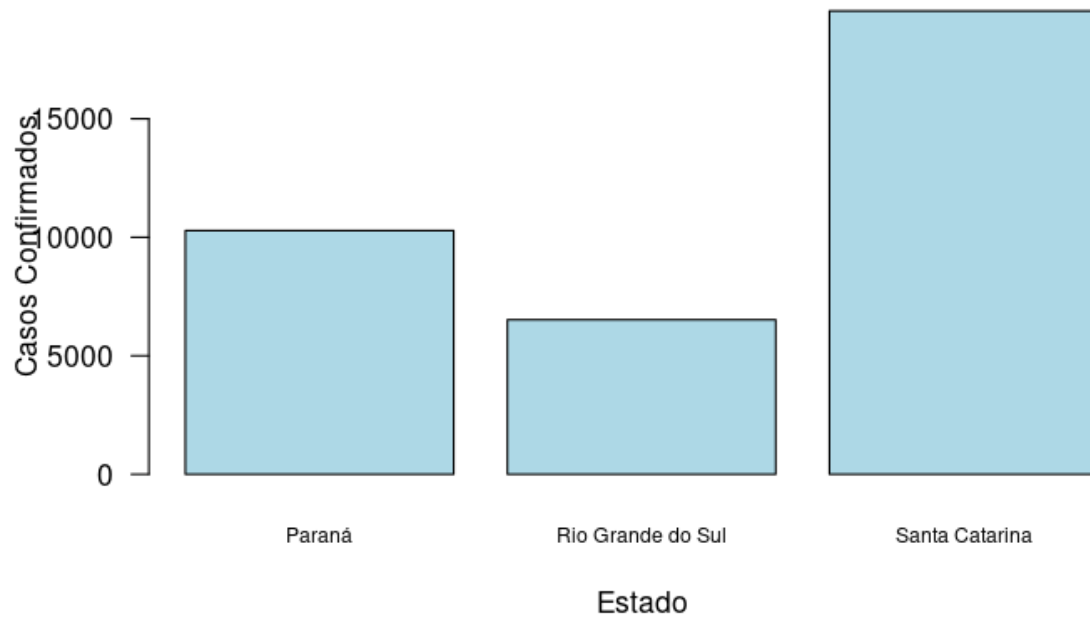
### **6.3. Interpretação do resultado:**

O p-valor obtido (0.6820769) é maior que 0.05, indicando que não há uma diferença estatisticamente significativa entre as médias dos casos confirmados dos dois grupos comparados. Portanto, não podemos afirmar que há uma diferença significativa no número de casos entre Paraná e Rio Grande do Sul em comparação com Santa Catarina.

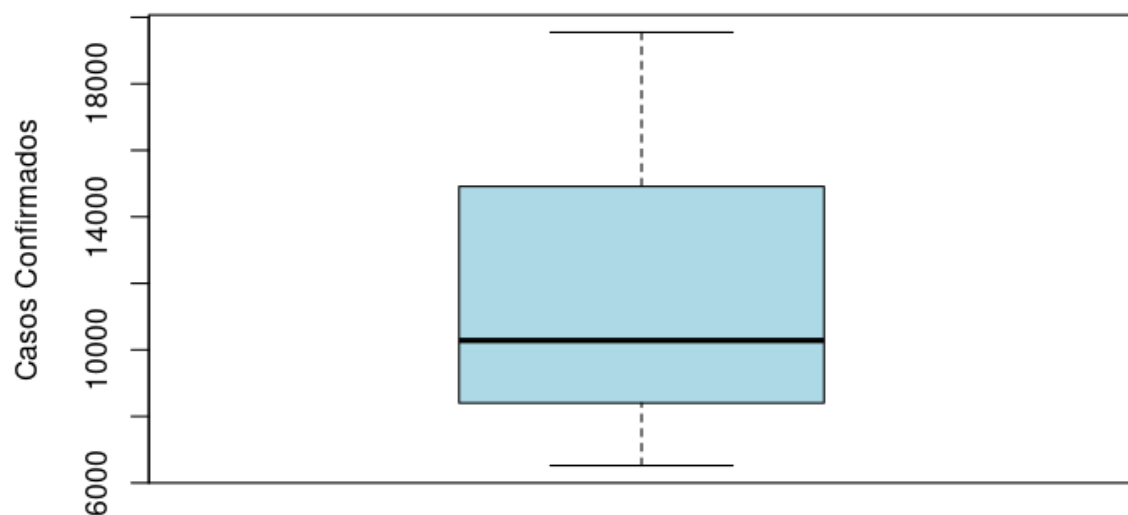
## 6.4. Tabelas

A tabela de frequência mostra a distribuição dos casos confirmados por estado

**Gráfico de Barras de Casos Confirmados - Sul**

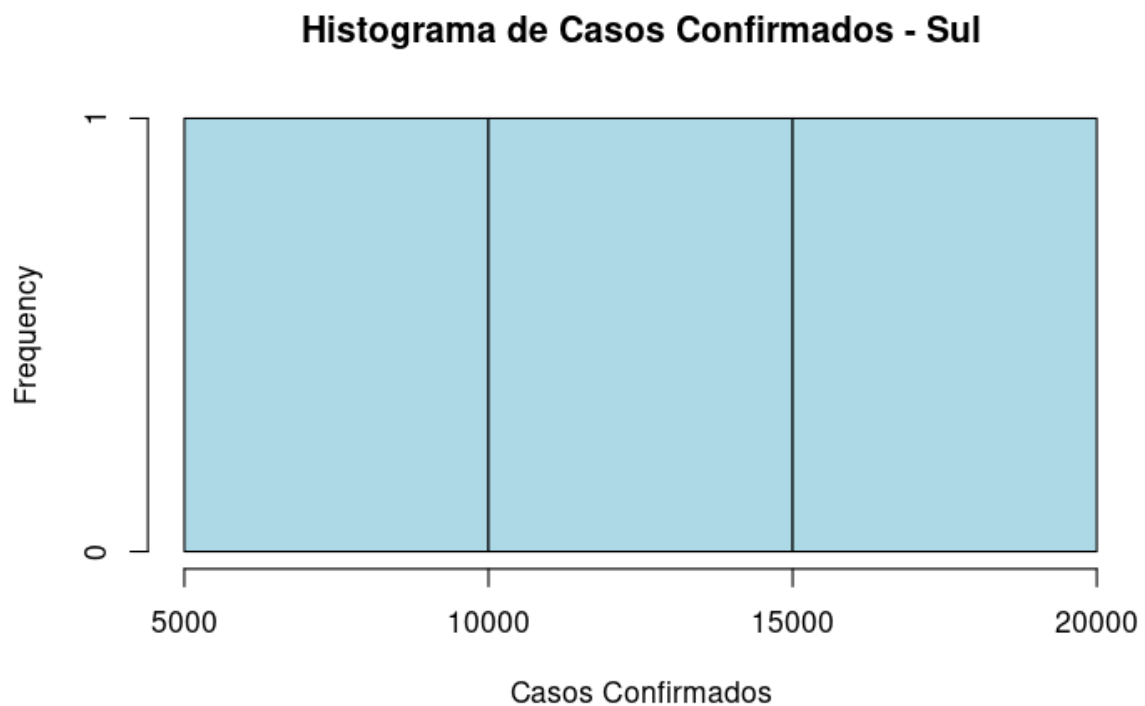


**Boxplot de Casos Confirmados - Sul**

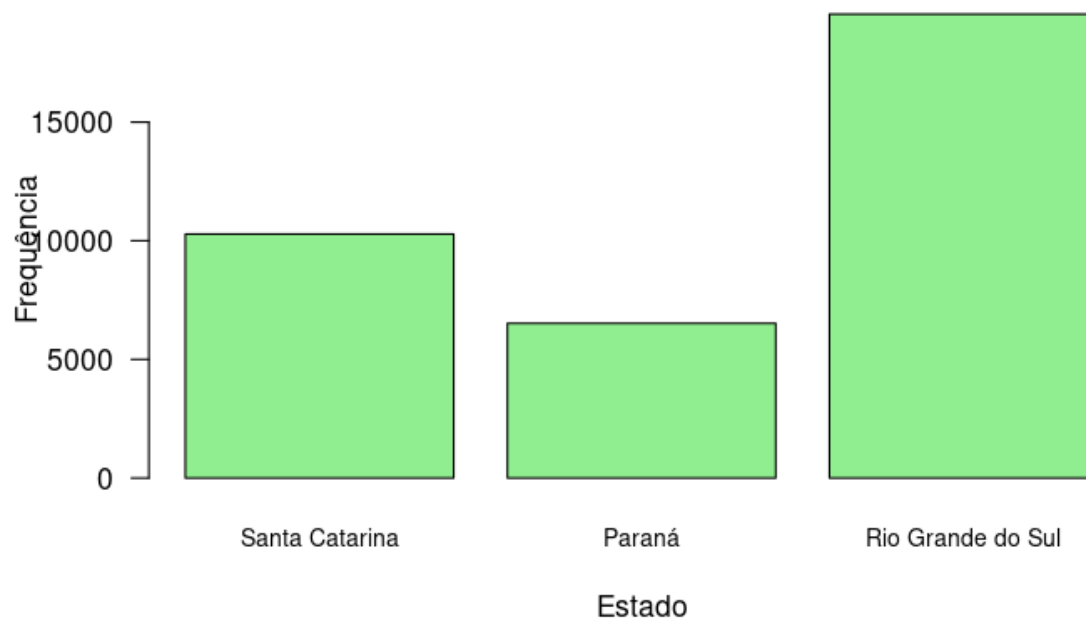


- **Primeiro Quartil (Q1):** 23.12% indicam que 25% dos estados da região Sul têm até 23.12% de probabilidade de uma pessoa ter COVID-19.
- **Mediana (Q2):** 28.3% mostram que 50% dos estados têm até 28.3% de probabilidade.
- **Terceiro Quartil (Q3):** 41.03% indicam que 75% dos estados têm até 41.03% de probabilidade.
- **Mínimo:** 17.94% é o menor valor de probabilidade.
- **Máximo:** 53.76% é o maior valor de probabilidade.

#### Gráficos Extras:



**Histograma da Ocupação Suspeito Clínico - Sul**



## 7.Região Norte

- Total de casos na região Norte: 12831

- **Média de casos na região Norte:** 1833
- **Mediana de casos na região Norte:** 1721
- **Moda de casos na região Norte:** Não há valor que se repete com mais frequência, portanto não há moda significativa.
- **Desvio padrão na região Norte:** 1662,112
- **Variância na região Norte:** 90,67715 %

#### **7.1. Probabilidade de uma pessoa ter COVID-19 em cada estado do Norte:**

**AM (Amazonas):** 34.25%

**PA (Pará):** 21.52%

**RO (Rondônia):** 24.49%

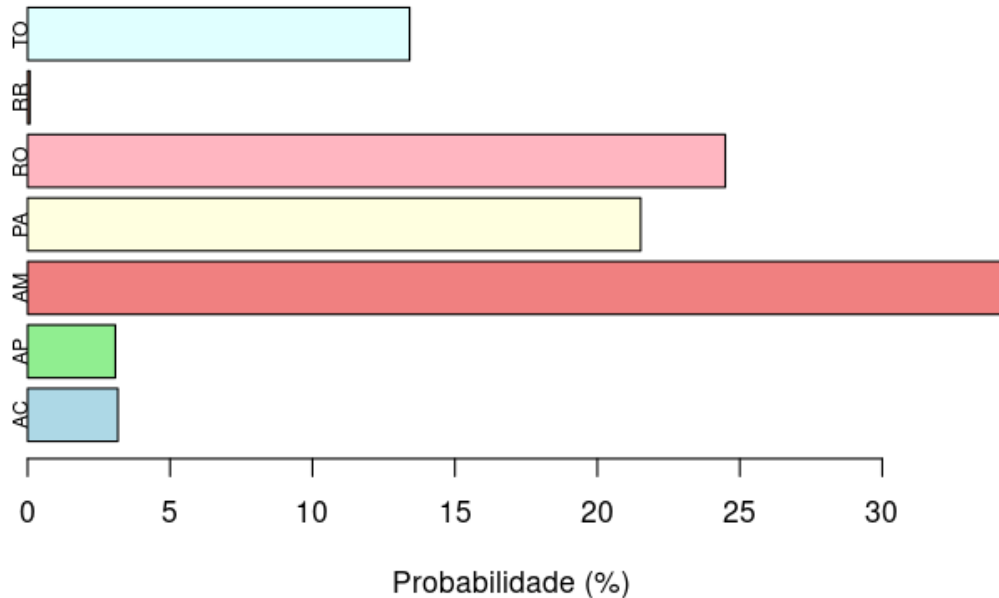
**RR (Roraima):** 0.08%

**AP (Amapá):** 3.08%

**TO (Tocantins):** 13.41%

**AC (Acre):** 3.17%

### Probabilidade de Casos - Região Norte



### 7.2. Resultados do Teste t

Realizamos um teste t para comparar as médias dos casos confirmados entre dois grupos:

- **Amostra 1** (Acre e AP): 407 e 395
- **Amostra 2** (Amazonas (4395 casos), Pará (2761 casos), Rondônia (3142 casos), Roraima (10 casos), Tocantins (1721 casos)).

### 7.3. Resultados do Teste t:

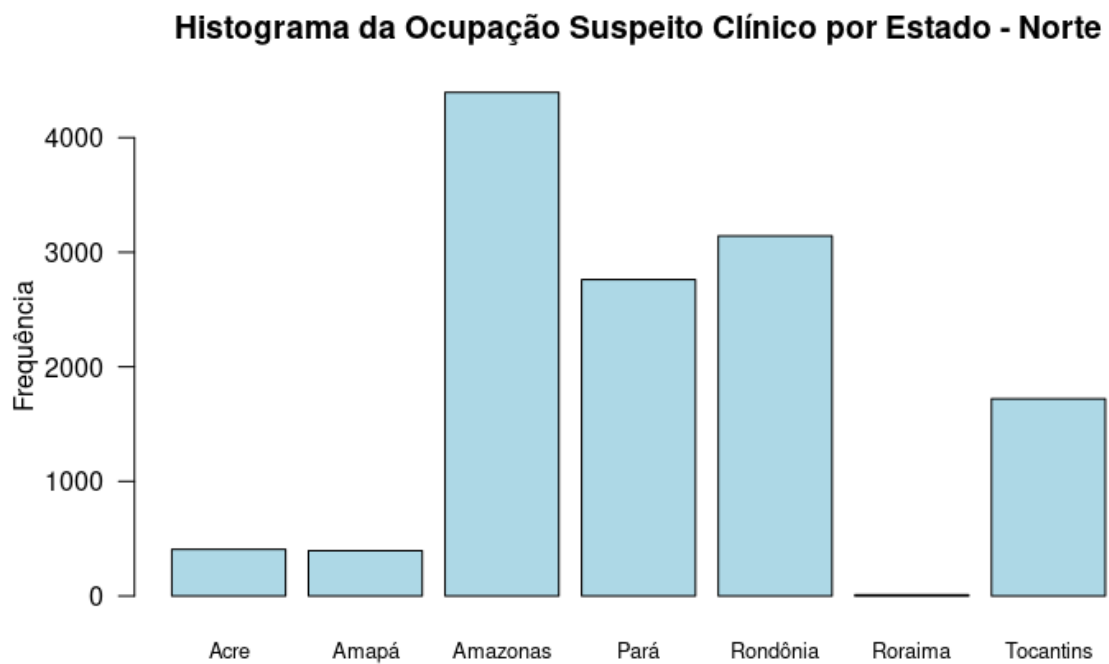
- **Estatística t:** -2.723827
- **p-valor:** 0.05277034
- **Intervalo de confiança:** -4.048,221 a 38,6209
- **Interpretação do resultado:**

O p-valor obtido (0.05277034) é maior que 0,05, o que sugere que não há uma diferença estatisticamente significativa entre as médias dos casos confirmados dos dois

grupos comparados. Portanto, não podemos afirmar que há uma diferença significativa no número de casos entre Acre e Amapá em comparação com Amazonas, Pará, Rondônia, Roraima e Tocantins.

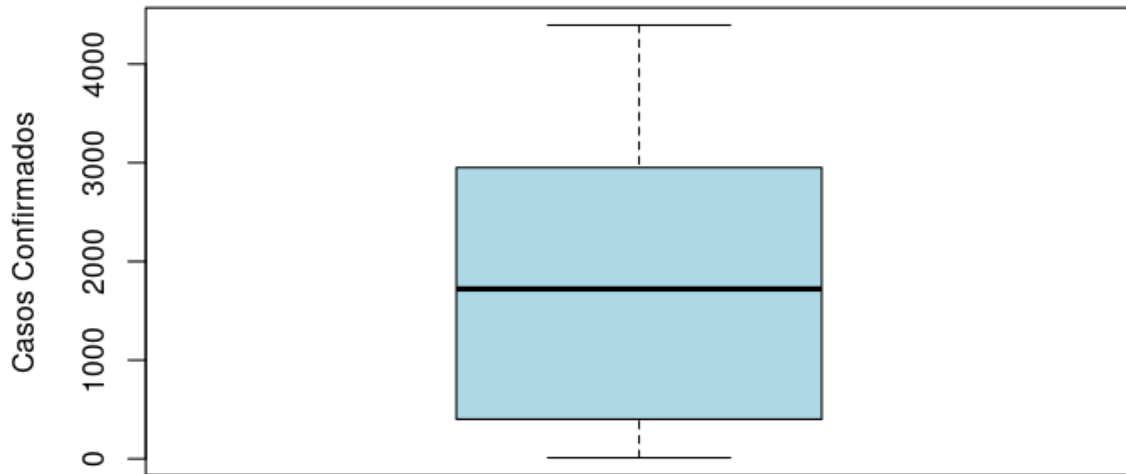
#### 7.4. Tabelas

A tabela de frequência mostra a distribuição dos casos confirmados por estado





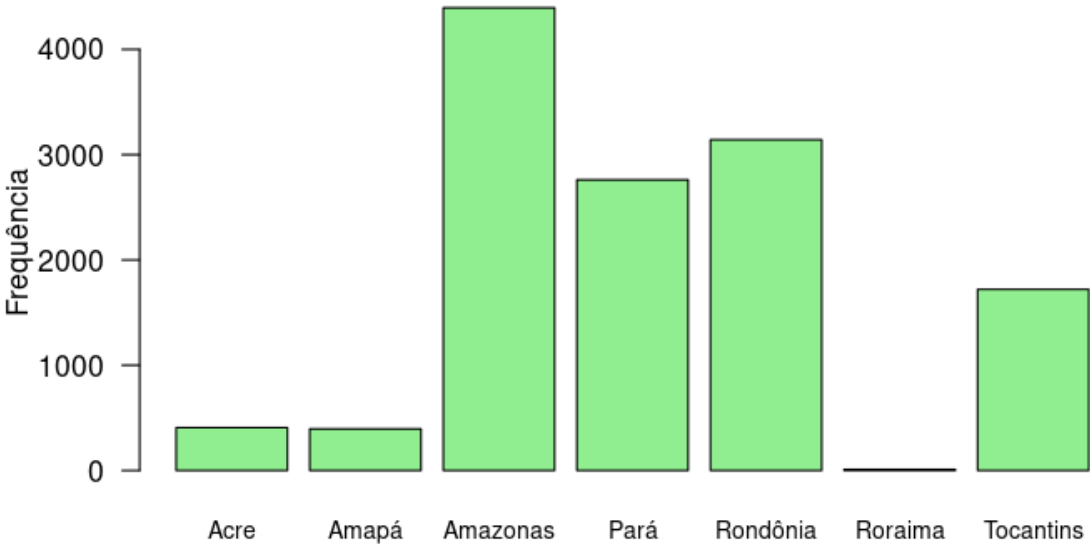
**Boxplot de Casos Confirmados - Norte**



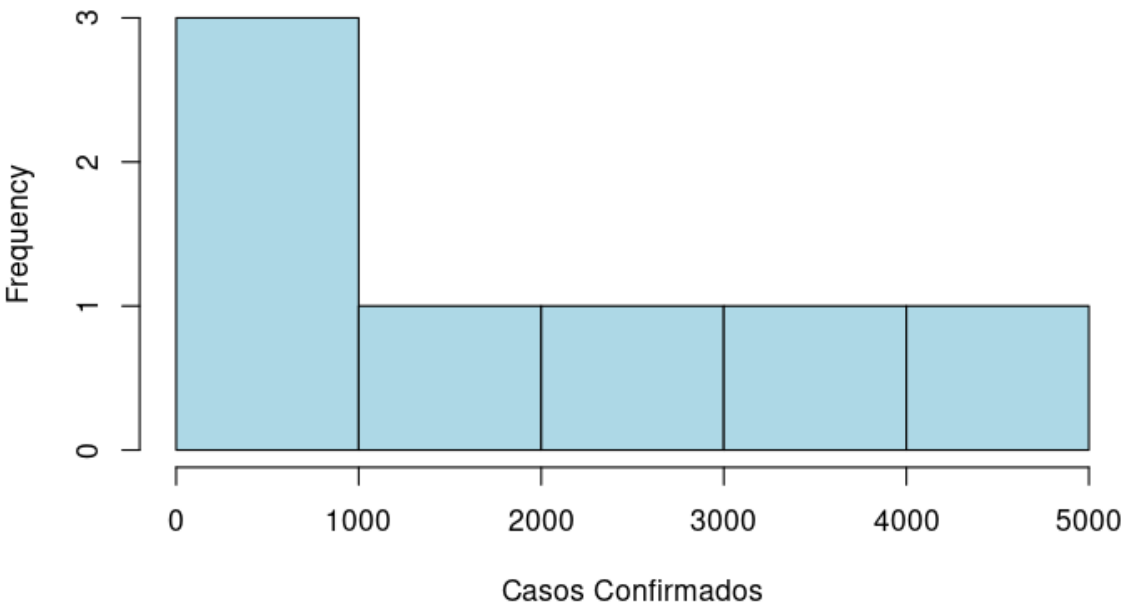
- **Primeiro Quartil (Q1):** 401 casos indicam que 25% dos estados da região Norte têm até 401 casos confirmados de COVID-19.
- **Mediana (Q2):** 1721 casos mostram que 50% dos estados têm até 1721 casos confirmados.
- **Terceiro Quartil (Q3):** 2951.5 casos indicam que 75% dos estados têm até 2951.5 casos confirmados.
- **Mínimo:** 10 casos é o menor valor de casos confirmados.
- **Máximo:** 4395 casos é o maior valor de casos confirmados.

**Gráficos Extras:**

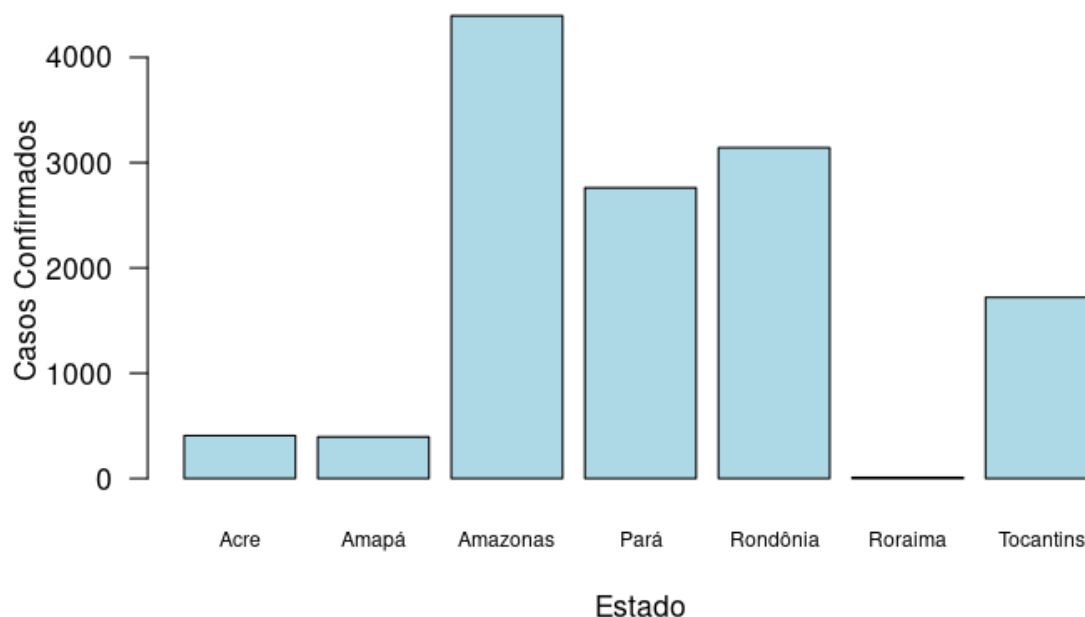
**Histograma da Ocupação Suspeito Clínico por Estado - Norte**



**Histograma de Casos Confirmados - Norte**



**Gráfico de Barras de Casos Confirmados - Norte**

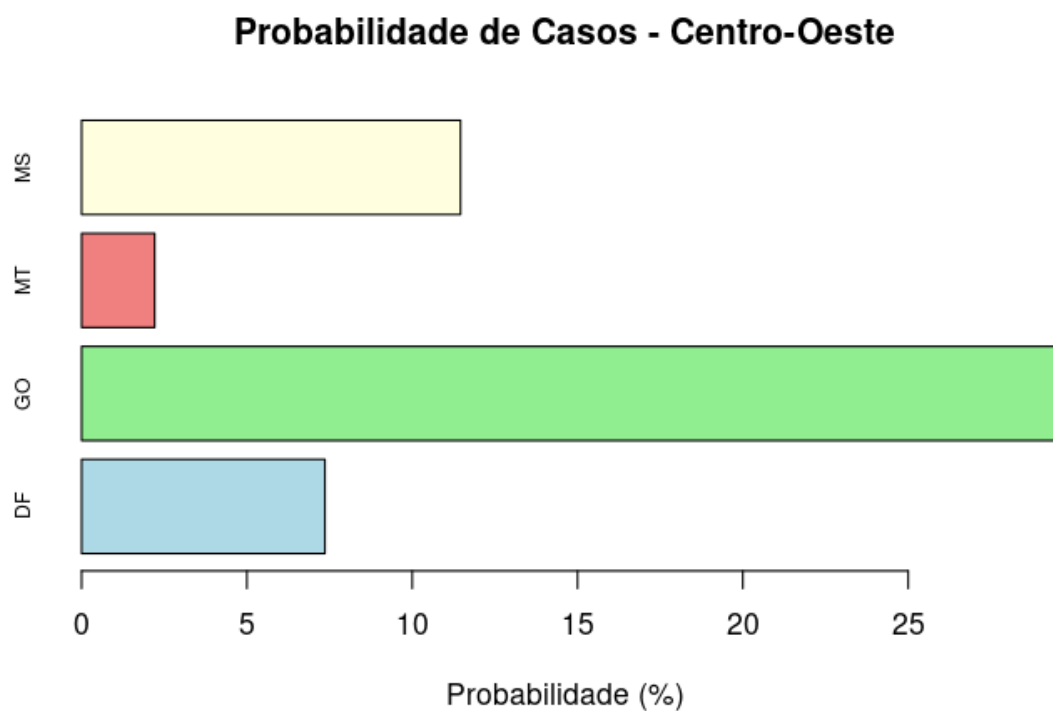


## 8.Região Centro-Oeste

- **Total de casos confirmados na região Centro-Oeste:** 49,623 casos.
- **Média de casos confirmados na região Centro-Oeste:** 12,405.75 casos.
- **Mediana de casos confirmados na região Centro-Oeste:** 7,470 casos.
- **Moda de casos confirmados na região Centro-Oeste:** Não há valor que se repete com mais frequência, portanto não há moda significativa.
- **Desvio padrão de casos confirmados na região Centro-Oeste:** 14,112.91.
- **Variância de casos confirmados na região Centro-Oeste:** 199,174,225.
- **Coeficiente de variação de casos confirmados na região Centro-Oeste:** 113.761%.

### 8.1. Probabilidade de uma pessoa ter COVID-19 em cada estado do Sul:

- **Pará (PA):** 65,14116%
- **Rondônia (RO):** 25,23628%
- **Amazonas (AM):** 4,751829%
- **Mato Grosso (MT):** 4,870725%



### 8.2. Resultados do Teste t

Realizamos um teste t para comparar as médias dos casos confirmados entre dois grupos:

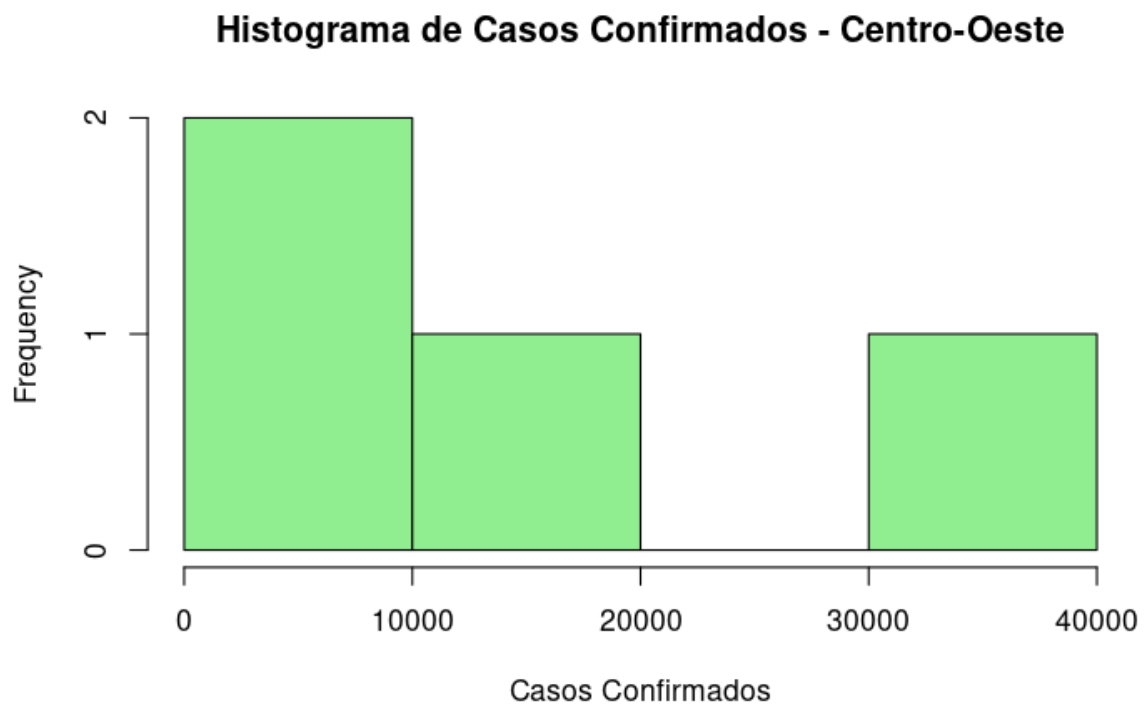
- **Amostra 1 (Distrito Federal e Goiás):** 2358 e 32325
- **Amostra 2 (Mato Grosso e Mato Grosso do Sul):** 2417 e 12523
- **Estatística t:** 0.6242809
- **p-valor:** 0.628467
- **Intervalo de confiança:** -121861.4 a 141604.4

### 8.3. Interpretação do resultado:

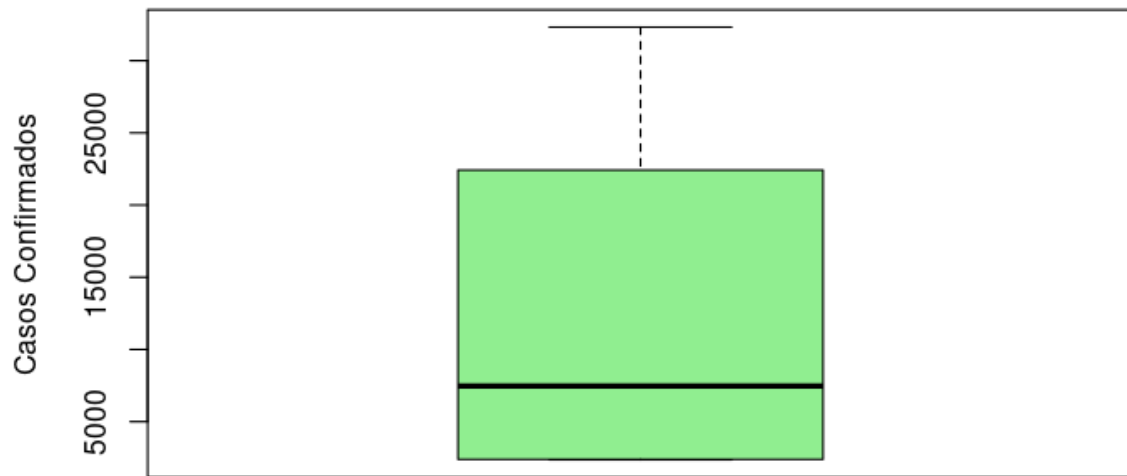
O p-valor obtido (0.628467) é maior que 0.05, indicando que não há uma diferença estatisticamente significativa entre as médias dos casos confirmados dos dois grupos comparados. Portanto, não podemos afirmar que há uma diferença significativa no número de casos entre Distrito Federal e Goiás em comparação com Mato Grosso e Mato Grosso do Sul.

### 8.4. Tabelas

A tabela de frequência mostra a distribuição dos casos confirmados por estado



### Boxplot de Casos Confirmados - Centro-Oeste



**Primeiro Quartil (Q1):** 2358 casos indicam que 25% dos estados da região Centro-Oeste têm até 2358 casos confirmados de COVID-19.

**Mediana (Q2):** 2417 casos mostram que 50% dos estados têm até 2417 casos confirmados.

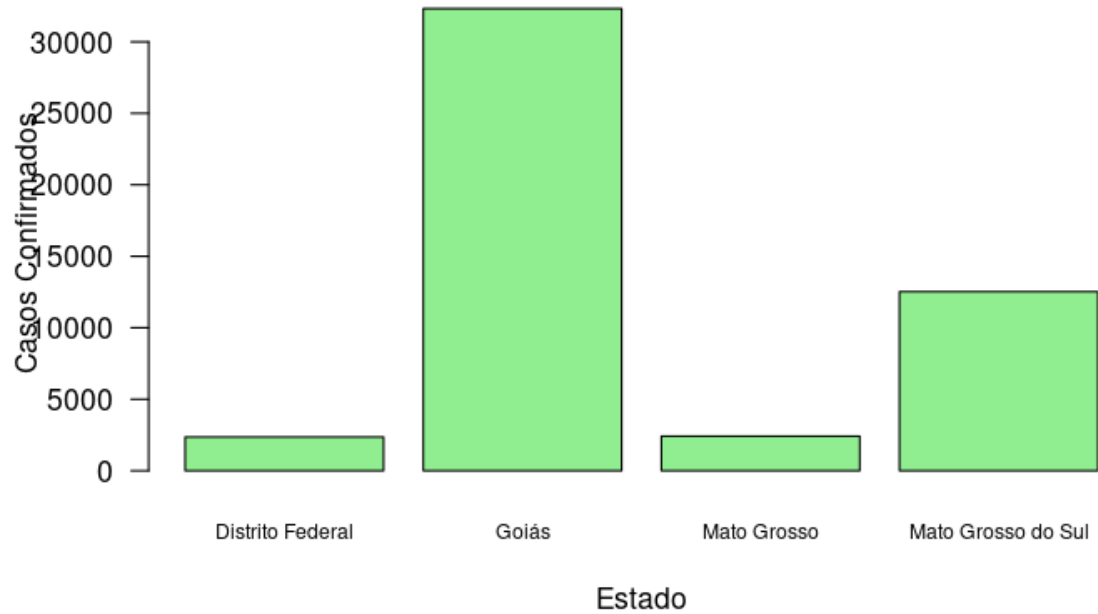
**Terceiro Quartil (Q3):** 12523 casos indicam que 75% dos estados têm até 12523 casos confirmados.

**Mínimo:** 2358 casos é o menor valor de casos confirmados.

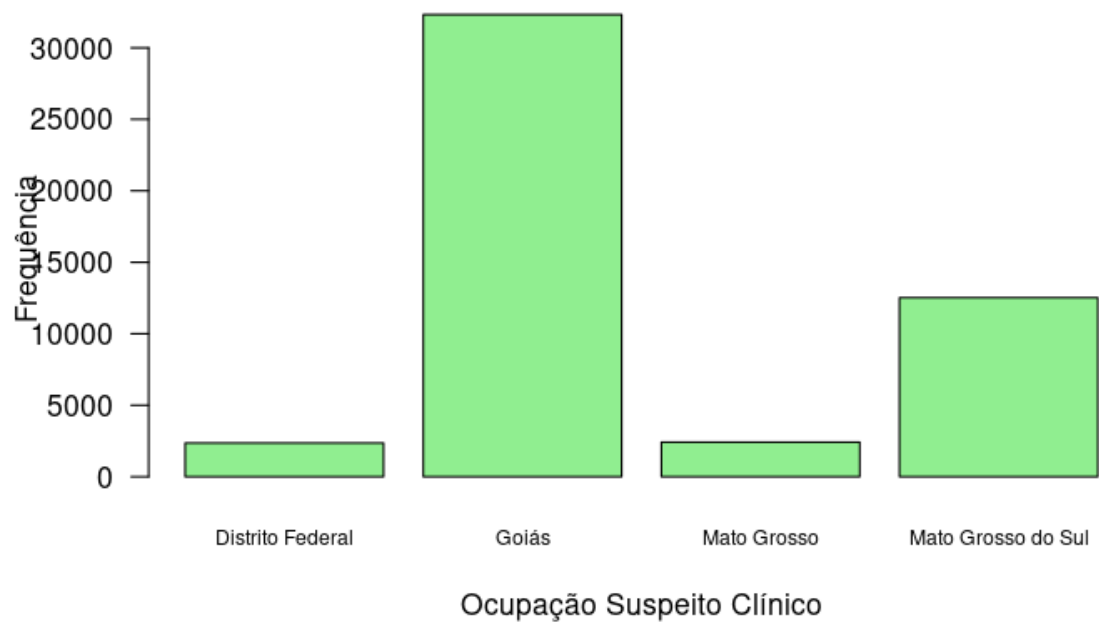
**Máximo:** 32325 casos é o maior valor de casos confirmados.

## Gráficos Extras:

### Gráfico de Barras de Casos Confirmados - Centro-Oeste



### Histograma da Ocupação Suspeito Clínico - Centro-Oeste



## **9. Análise**

### **Distribuição Geográfica dos Casos:**

- Os estados mais populosos e economicamente ativos, como São Paulo, Minas Gerais e Rio de Janeiro, apresentam os maiores números de casos confirmados. Isso é esperado devido à densidade populacional e à intensa movimentação de pessoas nessas regiões.

### **Desigualdades Regionais:**

- Há uma clara disparidade entre as regiões, com o Sudeste liderando em número de casos confirmados e o Centro-Oeste apresentando o menor número. Essa disparidade pode refletir diferenças na infraestrutura de saúde, densidade populacional e eficácia das medidas de controle adotadas.

### **Desafios da Região Nordeste:**

- Apesar de ser uma região populosa, o Nordeste apresenta um número menor de casos em comparação com o Sudeste. Isso pode indicar desafios socioeconômicos e de infraestrutura de saúde específicos da região, exigindo a implementação de medidas direcionadas e eficazes no combate à pandemia.

## **10. Conclusão**

Este relatório destaca a distribuição dos casos confirmados de COVID-19 por região no Brasil. A região Sudeste apresenta o maior número de casos, enquanto a região Centro-Oeste apresenta o menor número de casos entre as regiões analisadas.

As tendências identificadas destacam a importância de políticas públicas e ações específicas para cada região, considerando suas particularidades e necessidades no enfrentamento da COVID-19. A análise dos dados reforça a importância de uma abordagem regionalizada e adaptada às realidades locais no combate à pandemia.



## 11.Referências

- Portal DATASUS. [Open Data SUS](#)