

IFSP - Instituto Federal de Educação, Ciência e Tecnologia
Câmpus São Paulo

KARIMAN GOMES FERREIRA DO NASCIMENTO	SP3068811
MARCELO FERREIRA CRUZ	SP3068862
MARIANA CAVALLE	SP307336X
VICTOR YOSHINOBU LISBOA SAGAWA	SP3073513

Análise Exploratoria de Dados

São Paulo - SP - Brasil

2023

IFSP - Instituto Federal de Educação, Ciência e Tecnologia
Câmpus São Paulo

KARIMAN GOMES FERREIRA DO NASCIMENTO	SP3068811
MARCELO FERREIRA CRUZ	SP3068862
MARIANA CAVALLE	SP307336X
VICTOR YOSHINOBU LISBOA SAGAWA	SP3073513

Análise Exploratoria de Dados

Relatório da Análise exploratória de dados
da Relatório sobre a Síndrome Respiratória
Aguda Grave (SRAG)

Professor: JOSCELI MARIA TENORIO

IFSP - Instituto Federal de Educação, Ciência e Tecnologia
Câmpus São Paulo

Tecnologia em Análise e Desenvolvimento de Sistemas

ESP1A5 - Estatística e Probabilidade

São Paulo - SP - Brasil

2023

Sumário

1	INTRODUÇÃO	3
1.1	Descrição do Problema	3
1.2	Proposta de Solução	3
1.3	Descrição da Base de Dados	5
2	ANÁLISES	6
2.1	Análise de Variáveis	6
2.2	Análise Descritiva	8
2.3	Análise de Frequência	10
2.4	Análise Probabilística	17
2.5	Análise de Vacinação	24
2.6	Análise de Sintomas	34
2.7	Testes de Hipótese	41
3	CONCLUSÃO	52

1 Introdução

A Síndrome Respiratória Aguda Grave (SRAG) é uma condição clínica caracterizada por uma infecção respiratória aguda severa que afeta os pulmões e pode levar a complicações graves, incluindo insuficiência respiratória aguda. A SRAG é causada por diversos agentes patogênicos, incluindo vírus, bactérias e outros microrganismos.

A análise de dados relacionados à SRAG desempenha um papel crucial na compreensão e no controle dessa síndrome. Ao examinar as variáveis pertinentes, podemos identificar padrões, tendências e fatores de risco associados a essa condição. As variáveis selecionadas para análise são importantes, pois fornecem informações essenciais sobre os pacientes, suas características epidemiológicas, além de dados clínicos relevantes.

1.1 Descrição do Problema

Os profissionais de saúde enfrentam o desafio de lidar com essa doença complexa e, para isso, é fundamental contar com análises de dados detalhadas. A compreensão dos dados relacionados à SRAG desempenha um papel crucial na identificação de padrões, fatores de risco e respostas terapêuticas adequadas. Por meio de análises precisas, os profissionais de saúde podem obter insights valiosos sobre a doença, seu comportamento e a eficácia das medidas adotadas. Essas informações são essenciais para o desenvolvimento de estratégias mais efetivas de prevenção, diagnóstico e tratamento da SRAG.

1.2 Proposta de Solução

Com o objetivo de enfrentar a complexidade da Síndrome Respiratória Grave (SRG) e auxiliar os profissionais de saúde, propõe-se a aplicação de análises estatísticas adequadas a cada contexto. Essas análises têm o potencial de extrair informações valiosas dos dados disponíveis, visando aprimorar a prevenção, diagnóstico e tratamento da SRG.

- **Análise de Variáveis:**

A análise de variáveis é fundamental para entender as características dos pacientes afetados pela SRG. Essa análise permite explorar informações como idade, sexo, raça/cor, e outras variáveis relevantes. Ao investigar essas variáveis, é possível identificar padrões e diferenças no perfil dos pacientes, fornecendo insights valiosos sobre fatores de risco e grupos mais suscetíveis à SRG. Essa análise desempenha um papel crucial na tomada de decisões informadas e na implementação de estratégias de prevenção e tratamento mais direcionadas.

- **Análise Descritiva:**

A análise descritiva envolve a identificação e o estudo das características básicas dos pacientes, como idade, sexo, sintomas e outros atributos relevantes. Por meio dessa análise, é possível obter informações como a distribuição etária dos pacientes, o perfil de sintomas mais comuns e possíveis associações entre variáveis. Esses insights fornecem uma visão abrangente da doença, auxiliando na identificação de grupos de risco, no diagnóstico precoce e no desenvolvimento de estratégias eficazes de prevenção e tratamento.

- **Análise de Frequência:**

A análise de frequência permite identificar a prevalência de diferentes eventos ou características relacionadas à SRAG. Ela envolve a contagem e a classificação de ocorrências, como sintomas relatados pelos pacientes e a cobertura vacinal contra a gripe. Essa análise fornece insights sobre os sintomas mais comuns, a eficácia da vacinação e a distribuição de eventos ao longo do tempo. Essas informações são fundamentais para o diagnóstico, a avaliação da eficácia das medidas preventivas e a identificação de possíveis padrões ou tendências, contribuindo para o aprimoramento das estratégias de controle. Com a identificação dos sintomas mais comuns relatados pelos pacientes, como febre, tosse e dor de garganta podemos auxiliar o profissional na identificação de sintomas-chave para o diagnóstico precoce da SRG.

- **Análise Probabilística:**

A análise probabilística permite avaliar a probabilidade de ocorrência de determinados eventos ou resultados, como a relação entre a idade dos pacientes e a duração dos sintomas. Ao aplicar técnicas de análise de correlação, é possível identificar padrões e tendências nos dados, fornecendo insights sobre a evolução da doença e fatores que podem influenciar a gravidade e a duração dos sintomas. Essa análise probabilística auxilia na compreensão mais profunda da SRAG, contribuindo para a tomada de decisões clínicas informadas e o desenvolvimento de estratégias de prevenção e controle mais eficazes.

- **Análise de Vacinação:**

Por meio da análise de vacinação, é possível avaliar a cobertura vacinal entre os pacientes e investigar possíveis diferenças entre grupos demográficos, como sexo e faixa etária. Essa análise fornece informações valiosas sobre a proteção conferida pela vacinação e sua influência na prevenção da SRAG. Além disso, ao comparar a incidência da doença entre vacinados e não vacinados, é possível determinar a efetividade da vacina na redução dos casos. A análise de vacinação contribui para embasar estratégias de imunização e aprimorar as medidas de controle da SRAG.

- **Análise de Sintomas:**

A análise de sintomas identifica os sintomas mais comuns relatados pelos pacientes, como febre, tosse e dor de garganta, proporcionando insights importantes para o diagnóstico precoce da doença. Além disso, a análise de sintomas pode examinar a relação entre a idade dos pacientes e o número de sintomas apresentados, bem como a duração dos sintomas, contribuindo para uma melhor compreensão da evolução da doença e identificando fatores que influenciam a gravidade e a duração dos sintomas. Essa análise é essencial para embasar decisões clínicas informadas e melhorar o tratamento e o cuidado dos pacientes com SRAG.

- **Testes de Hipótese:**

Os testes de hipótese permitem comparar médias, proporções ou correlações entre grupos específicos, como pacientes vacinados e não vacinados contra a gripe, pacientes com e sem febre, entre outros. Ao aplicar testes de hipótese, é possível avaliar se existem diferenças estatisticamente significativas entre esses grupos, fornecendo insights valiosos sobre a relação entre variáveis e a presença ou gravidade da SRAG. Essa análise estatística ajuda os profissionais de saúde a compreender melhor os fatores de risco, a tomar decisões clínicas fundamentadas e a implementar estratégias de prevenção e controle adequadas para a doença.

1.3 Descrição da Base de Dados

A base de dados utilizada neste projeto é o **"SRAG 2013 a 2018 - Banco de Dados de Síndrome Respiratória Aguda Grave"**. Para a realização das análises, foi utilizada exclusivamente a base de dados referente ao ano de **2013**, cujo arquivo CSV correspondente é denominado **"INFLUD13"**.

A escolha de iniciar a análise pelos dados de 2013 é permitir estabelecer uma linha de base e compreender o panorama inicial da síndrome respiratória aguda grave. Ao examinar os dados deste período inicial, é possível identificar possíveis tendências, padrões e comportamentos que podem servir como referência para as análises subsequentes dos anos seguintes. Ao analisar esses dados vamos proporcionar uma visão mais abrangente sobre a ocorrência da SRAG nesse ano específico de 2013, permitindo futuras comparações com os demais anos do banco de dados. É possível estabelecer uma base sólida para a investigação da SRAG, explorando as características iniciais e fundamentais da doença e fornecendo um contexto essencial para as análises comparativas entre os anos subsequentes.

2 Análises

As análises de dados desempenham um papel fundamental no campo da saúde, oferecendo insights valiosos aos profissionais. Ao examinar padrões e tendências, é possível identificar fatores que influenciam a saúde da população. Com base nessas descobertas, os profissionais de saúde podem desenvolver estratégias mais eficazes de prevenção, intervenção e alocação de recursos, a fim de melhorar os resultados de saúde e garantir cuidados mais personalizados e direcionados.

2.1 Análise de Variáveis

- **Data dos primeiros sintomas / Diagnóstico (DT_SIN_PRI):**
 - Tipo: Categórica (Qualitativa) Nominal
 - Categoria: dd/mm/aaaa
 - Descrição: Data dos primeiros sintomas do caso de SRAG. Data do diagnóstico do caso de SRAG crônico e DRTs. Data do acidente nos acidentes da saúde do trabalhador.
- **Idade (NU_IDADE_N):**
 - Tipo: Numérica (Quantitativa) Discreta
 - Categoria: Número de anos
 - Descrição: Idade do paciente. A composição da variável segue o critério: 1º dígito: 1. Hora, 2. Dia, 3. Mês, 4. Ano. Por exemplo, 3009 significa nove meses e 4018 significa dezoito anos.
- **Gestante (CS_GESTANT):**
 - Tipo: Categórica (Qualitativa) Ordinal
 - Categoria: Código de gestação
 - Descrição: Idade gestacional da paciente. Os códigos representam: 1. 1º Trimestre, 2. 2º Trimestre, 3. 3º Trimestre, 4. Idade gestacional ignorada, 5. Não, 6. Não se aplica.
- **Sexo (CS_SEXO):**
 - Tipo: Categórica (Qualitativa) Nominal
 - Categoria: M- Masculino F- Feminino I- Ignorado

- Descrição: Sexo do paciente.
- **Raça/Cor (CS_RACA):**
 - Tipo: Categórica (Qualitativa) Nominal
 - Categoria: Código de raça/cor
 - Descrição: Classificação da raça/cor do paciente. Os códigos representam: 1. Branca, 2. Preta, 3. Amarela, 4. Parda, 5. Indígena, 9. Ignorado.
- **Recebeu vacina contra gripe (VACINA):**
 - Tipo: Categórica (Qualitativa) Nominal
 - Categoria: Indica se o paciente recebeu vacina contra gripe.
 - Descrição: Indica se o paciente recebeu a vacina contra a gripe. 1: Sim 2: Não 9: Ignorado.
- **Sinais e sintomas - febre (FEBRE):**
 - Tipo: Categórica (Qualitativa) Nominal
 - Categoria: Indica se o paciente apresentou febre.
 - Descrição: Indica se o paciente apresentou febre. 1: Sim 2: Não 9: Ignorado.
- **Sinais e sintomas - tosse (TOSSE):**
 - Tipo: Categórica (Qualitativa) Nominal
 - Categoria: Indica se o paciente apresentou tosse.
 - Descrição: Indica se o paciente apresentou tosse. 1: Sim 2: Não 9: Ignorado.
- **Sinais e sintomas - dor de garganta (GARGANTA):**
 - Tipo: Categórica (Qualitativa) Nominal
 - Categoria: Indica se o paciente apresentou dor de garganta.
 - Descrição: Indica se o paciente apresentou dor de garganta. 1: Sim 2: Não 9: Ignorado

2.2 Análise Descritiva

A análise descritiva em estatística é uma forma de resumir, organizar e descrever os principais aspectos de um conjunto de dados. Ela visa fornecer uma visão geral dos dados, identificar padrões, tendências e características relevantes por meio de medidas e métodos estatísticos. Essas informações podem auxiliar no diagnóstico precoce na percepção de padrões para planejamento e prevenção de ocorrências de saúde, além de mitigar gastos e danos causados pelas SRAG. Para essa análise descritiva, foi utilizada a função "**mean**" para calcular as médias. Além disso, a biblioteca **ggplot2** foi utilizada para criar os gráficos de barras e de caixa. Os códigos e funções utilizadas podem ser encontradas no arquivo **Analise Descritiva.R**.

- **Idade (NU_IDADE_N):**
 - Tipo: Numérica (Quantitativa) Discreta
 - Categoria: Número de anos
 - Descrição: Para a realização de uma análise descritiva nessa variável, iremos considerar apenas valores maiores que 4000, pois assim, de acordo com a composição da variável, poderemos analisar obter dados quantitativos sobre a idade em anos.

Para isso, gera-se uma amostra filtrada nas condições descritas anteriormente. Assim, ao subtrair 4000, pode-se obter a idade dos pacientes e suas implicações:

	DT_NOTIFIC	ID_MUNICIP	SEM_NOT	NU_ANO	SG_UF_NOT	DT_SIN_PRI	DT_NASC	NU_IDADE_N	CS_SEXO	CS
1	04/02/2013	110020	201306	2013	11	01/02/2013	14/02/1986	4026	M	
2	08/04/2013	110004	201315	2013	11	05/03/2013	16/02/1983	4030	F	
3	30/03/2013	110004	201313	2013	11	30/03/2013	14/11/1955	4057	F	
4	26/03/2013	110020	201313	2013	11	21/03/2013	28/10/2011	4001	M	
5	11/03/2013	110020	201311	2013	11	07/03/2013	14/11/2011	4001	M	
6	08/03/2013	110020	201310	2013	11	02/03/2013	02/03/2011	4002	M	
7	02/04/2013	110020	201314	2013	11	29/03/2013	18/02/1990	4023	M	
8	22/04/2013	110020	201317	2013	11	18/04/2013	13/12/1985	4027	F	
9	24/04/2013	110020	201317	2013	11	20/04/2013	26/09/1965	4047	M	
10	09/05/2013	110002	201319	2013	11	01/05/2013	02/10/1983	4029	F	
11	22/05/2013	110020	201321	2013	11	07/05/2013	21/04/1948	4065	M	
12	04/06/2013	110004	201323	2013	11	29/05/2013	03/02/1977	4036	F	
13	07/06/2013	110004	201323	2013	11	05/06/2013	24/06/2008	4004	F	

Figura 1 – Amostra de Idade / Idade - Resultado

Arquivo: Analise Descritiva.R

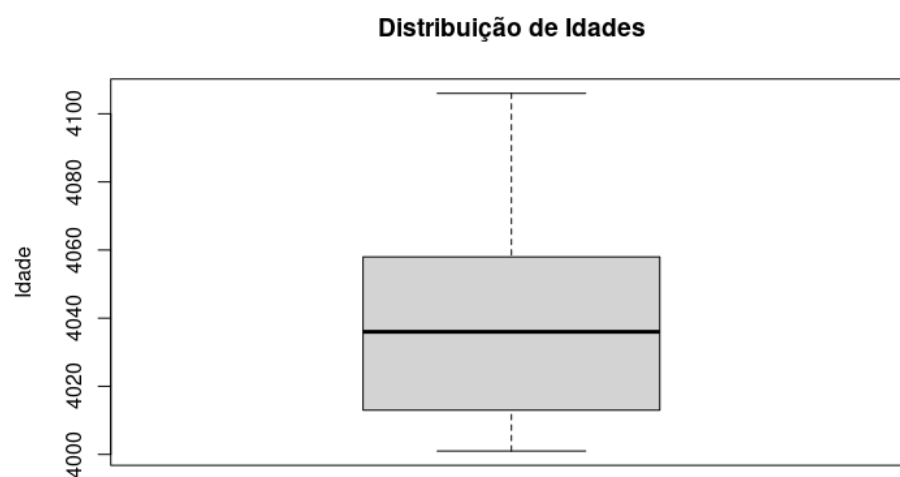


Figura 2 – Idade / Idade - Gráfico de Caixa

Arquivo: Analise Descritiva.R

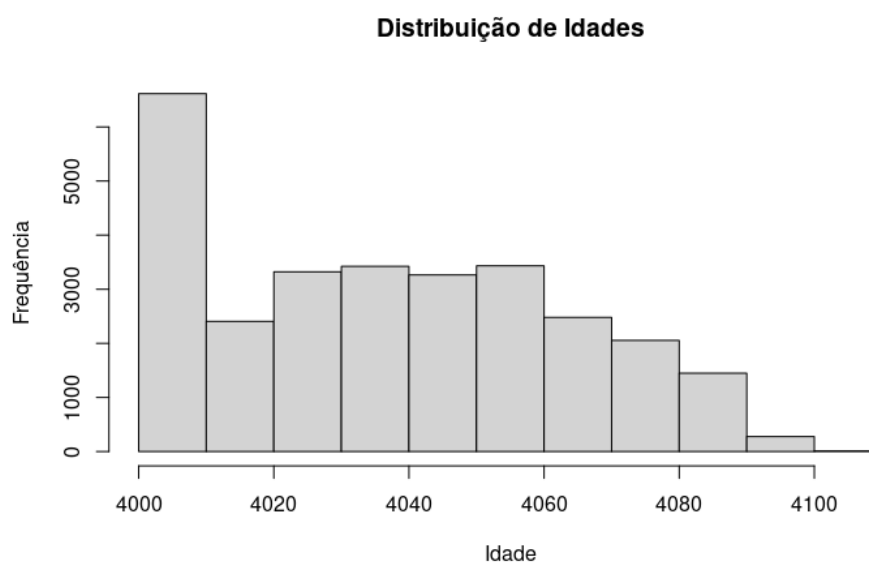


Figura 3 – Idade / Idade - Gráfico de Barras

Arquivo: Analise Descritiva.R

As demais variáveis não são quantitativas e são melhores representadas na Análise de Frequência.

2.3 Análise de Frequência

A análise de frequência pode ajudar os profissionais de saúde a entender melhor a distribuição e a prevalência de características, sintomas e fatores demográficos relacionados à doença. Essas informações podem auxiliar no diagnóstico precoce, tratamento adequado, planejamento de políticas de saúde e estratégias de prevenção direcionadas. Para essa análise de frequência, foi utilizada a função **"table"** para calcular as frequências de cada uma das variáveis. Além disso, a biblioteca **ggplot2** foi utilizada para criar os gráficos de barras. Os códigos e funções utilizadas podem ser encontradas no arquivo **Analise de Frequencia.R**.

- **Data dos primeiros sintomas / Diagnóstico:** Houve um aumento significativo de casos no mês de maio. Durante o verão, de junho a agosto, as contagens permaneceram relativamente estáveis, mas com algumas quedas em agosto. Ao longo dos meses de setembro a dezembro, as contagens diminuiriam gradualmente. Esses padrões sugerem uma possível correlação entre a sazonalidade e a incidência de sintomas ou diagnósticos.



```
R 4.2.2 · ~/
```

```
> # Data dos primeiros sintomas / Diagnóstico
```

```
> table(dados$DT_SIN_PRI)
```

01/01/2013	01/02/2013	01/03/2013	01/04/2013	01/05/2013	01/06/2013
12	26	43	77	204	335
01/07/2013	01/08/2013	01/09/2013	01/10/2013	01/11/2013	01/12/2013
358	224	155	84	61	39
02/01/2013	02/02/2013	02/03/2013	02/04/2013	02/05/2013	02/06/2013
18	14	26	43	170	249
02/07/2013	02/08/2013	02/09/2013	02/10/2013	02/11/2013	02/12/2013
195	178	105	78	49	32
03/01/2013	03/02/2013	03/03/2013	03/04/2013	03/05/2013	03/06/2013
16	14	29	44	160	306

Figura 4 – Data dos Primeiros Sintomas - Frequência (Amostra)

Arquivo: Analise de Frequencia.R

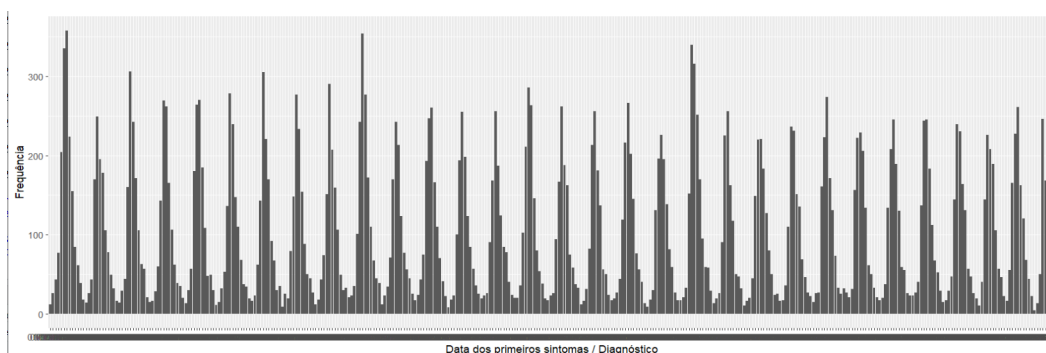


Figura 5 – Data dos Primeiros Sintomas - Gráfico de Barras

Arquivo: Analise de Frequencia.R

- **Gestante:** Podemos observar que há uma distribuição relativamente equilibrada entre os trimestres de gestação, sendo o segundo trimestre o mais frequente. No entanto, é importante ressaltar que um número significativo de casos apresenta idade gestacional desconhecida. Essa informação é essencial para uma abordagem clínica adequada e compreensão dos desfechos clínicos associados à SRAG em gestantes. A análise desses padrões pode auxiliar na identificação de possíveis grupos de risco e na adaptação dos cuidados médicos durante o tratamento de pacientes com SRAG gestacional.

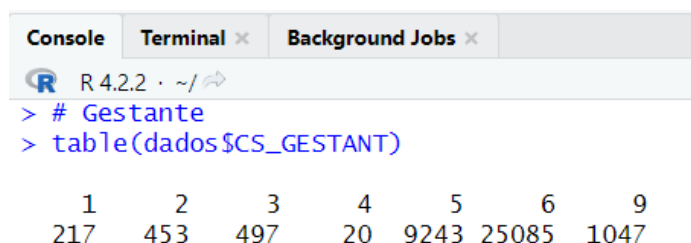


Figura 6 – Gestante - Frequência

Arquivo: Analise de Frequencia.R

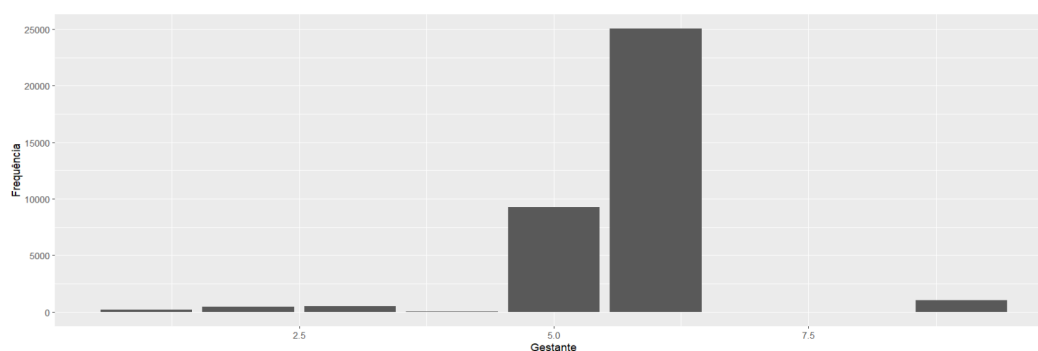


Figura 7 – Gestante - Gráfico de Barras

Arquivo: Analise de Frequencia.R

- **Sexo:** Ao analisar os dados relacionados ao sexo dos pacientes com SRAG, podemos observar a distribuição dos casos nas diferentes categorias. A maioria dos casos é representada pelo sexo feminino, com um total de 18.353 pacientes. Por outro lado, há apenas 8 casos em que o sexo é desconhecido e 18.202 casos do sexo masculino. Essa informação é relevante para compreender a distribuição da doença entre os sexos e pode fornecer insights sobre possíveis diferenças na suscetibilidade, manifestação ou busca por atendimento médico entre homens e mulheres com SRAG.

```
Console Terminal x Background Jobs x
R 4.2.2 · ~/
> # Sexo
> table(dados$CS_SEX0)

      F      I      M
18353      8 18202
```

Figura 8 – Sexo - Frequência

Arquivo: Analise de Frequencia.R

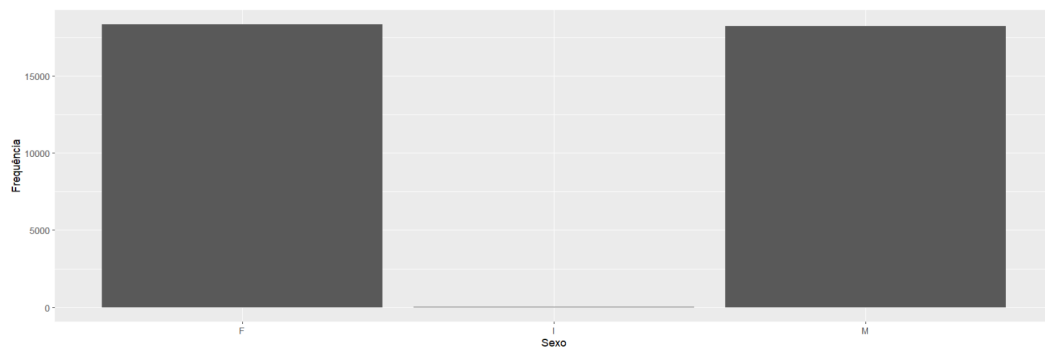


Figura 9 – Sexo - Gráfico de Barras

Arquivo: Analise de Frequencia.R

- **Raça/Cor:** Ao analisar os dados, observamos que a maior quantidade de casos está classificada como "Branca", seguida pela raça/cor "Parda". Em menor quantidade, temos os casos classificados como "Preta", "Amarela" e "Indígena". Além disso, também foi identificado um número significativo de casos em que a raça/cor é "Ignorada". Essa análise permite compreender a distribuição dos casos de SRAG em relação às diferentes raças/cor, fornecendo informações relevantes para os profissionais da saúde sobre possíveis disparidades e necessidades específicas de cada grupo racial/cor.

```

Console Terminal x Background Jobs x
R 4.2.2 · ~/
> # Raça/Cor
> table(dados$CS_RACA)

      1      2      3      4      5      9
22028 1368    270  6778    82  4297

```

Figura 10 – Raça/Cor - Frequência

Arquivo: Analise de Frequencia.R

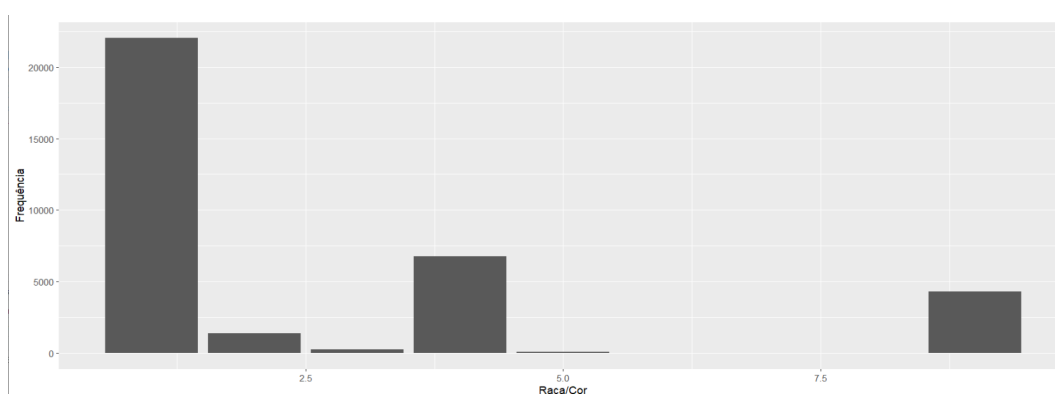


Figura 11 – Raça/Cor - Gráfico de Barras

Arquivo: Analise de Frequencia.R

- **Recebeu vacina contra gripe:** Ao analisar os dados, constatamos que a maioria dos pacientes não recebeu a vacina contra a gripe, representando a categoria "2" na variável. Por outro lado, um número significativo de pacientes indicou ter recebido a vacina contra a gripe, representado pela categoria "1". Além disso, alguns casos possuem essa informação como "Ignorada" (categoria "9"). Essa análise sugere a importância de investigar e compreender a cobertura vacinal contra a gripe entre os pacientes com SRAG. Profissionais de saúde podem utilizar esses insights para avaliar a eficácia da vacinação e a possível relação entre a imunização contra a gripe e a ocorrência da SRAG, identificando grupos de maior risco e direcionando estratégias de prevenção adequadas.

```

Console Terminal x Background Jobs x
R 4.2.2 · ~/
> # Recebeu vacina contra gripe
> table(dados$VACINA)

      1      2      9
8500 20397  6787

```

Figura 12 – Recebeu Vacina contra Gripe - Frequência

Arquivo: Analise de Frequencia.R

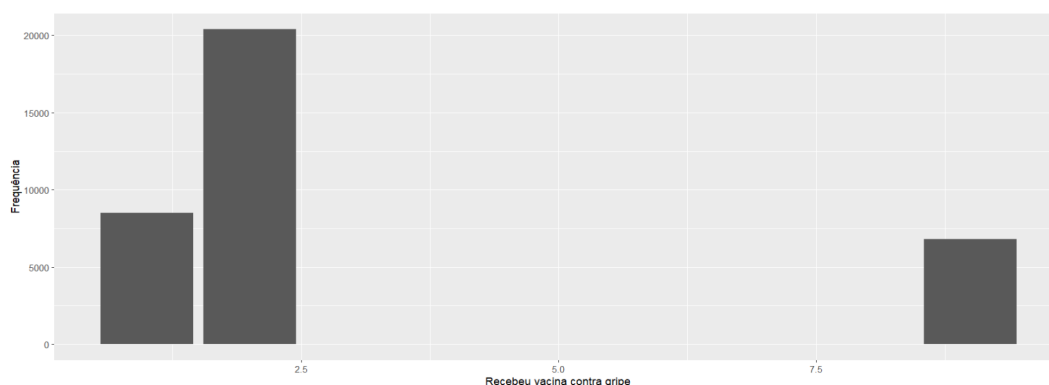


Figura 13 – Recebeu Vacina contra Gripe - Gráfico de Barras

Arquivo: Analise de Frequencia.R

- **Sinais e sintomas - febre:** Podemos observar que a maioria dos pacientes apresentou febre, representada pela categoria "1" na variável. Por outro lado, uma parcela menor dos pacientes não apresentou febre, correspondendo à categoria "2". Também encontramos alguns casos em que essa informação está registrada como "Ignorada" (categoria "9"). Essa análise sugere que a presença de febre é um sintoma comum entre os pacientes com SRAG. Profissionais de saúde podem utilizar esses insights para identificar a febre como um sinal relevante na suspeita e diagnóstico da SRAG, além de considerar a febre como um critério importante na avaliação clínica e triagem de pacientes com sintomas respiratórios. Isso pode auxiliar na tomada de decisões relacionadas ao encaminhamento, tratamento e monitoramento adequados dos casos de SRAG.

```
Console Terminal x Background Jobs x
R 4.2.2 · ~/
> # Sinais e sintomas - febre
> table(dados$FEBRE)

  1    2    9
29959 5883 391
```

Figura 14 – Sintoma de Febre - Frequência

Arquivo: Analise de Frequencia.R

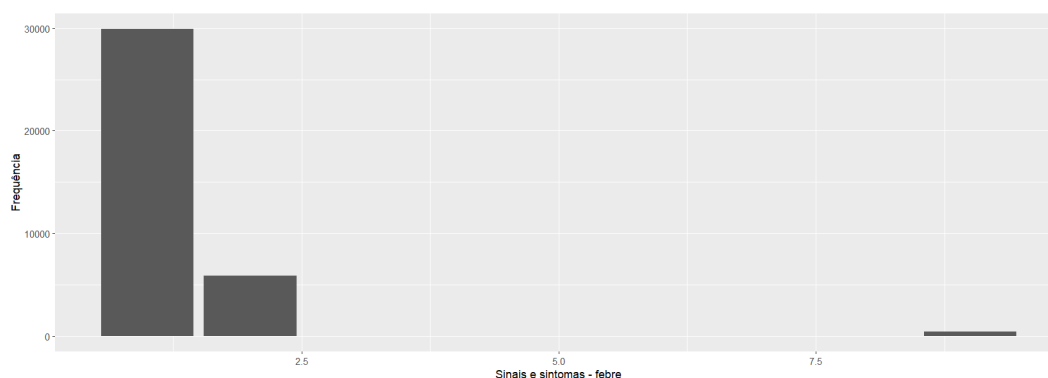


Figura 15 – Sintoma de Febre - Gráfico de Barras

Arquivo: Analise de Frequencia.R

- **Sinais e sintomas - tosse:** Podemos observar que a grande maioria dos pacientes com SRAG apresentou tosse, representada pela categoria "1" na variável. Por outro lado, uma parcela muito menor de pacientes não apresentou tosse, correspondendo à categoria "2". Também é importante mencionar que em alguns casos a informação sobre a tosse está registrada como "Ignorada" (categoria "9"). Essa análise indica que a tosse é um sintoma bastante comum entre os pacientes com SRAG. A presença da tosse pode ser um indicativo importante para a suspeita e diagnóstico da SRAG. Além disso, a tosse pode servir como um critério para a triagem e identificação de pacientes com sintomas respiratórios relevantes.

Console	Terminal ×	Background Jobs ×
R 4.2.2 · ~/		
> # Sinais e sintomas - tosse		
> table(dados\$TOSSE)		
1	2	9
32867	3095	344

Figura 16 – Sintoma de Tosse - Frequência

Arquivo: Analise de Frequencia.R

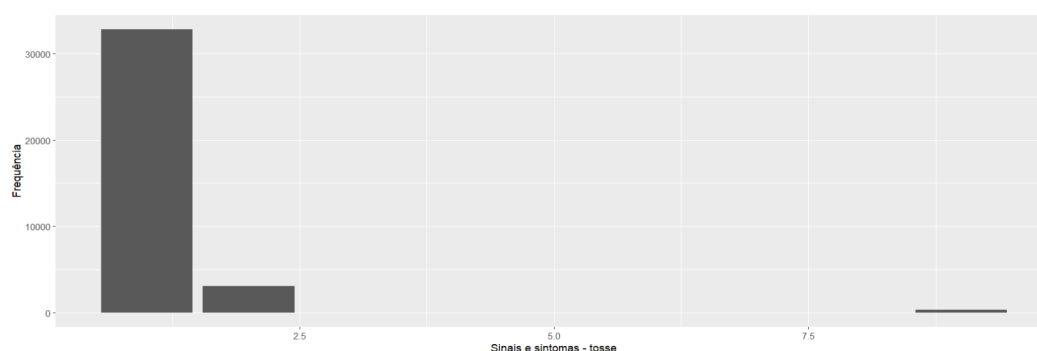


Figura 17 – Sintoma de Tosse - Gráfico de Barras

Arquivo: Analise de Frequencia.R

- **Sinais e sintomas - dor de garganta:** Ao analisar a dor de garganta, podemos observar que uma pequena parcela dos pacientes apresentou dor de garganta, representada pela categoria "1" na variável. Por outro lado, a maioria dos pacientes não apresentou esse sintoma, correspondendo à categoria "2". Essa análise indica que a dor de garganta não é um sintoma tão predominante quanto a tosse ou a febre, nos casos de SRAG. No entanto, sua presença pode ser relevante em alguns pacientes e deve ser considerada durante a avaliação clínica.


```
Console Terminal x Background Jobs x  
R 4.2.2 · ~/   
> # Sinais e sintomas - dor de garganta  
> table(dados$GARGANTA)  
  
1      2      9  
7949 23397 4003
```

Figura 18 – Sintoma de Dor de garganta - Frequência
Arquivo: Analise de Frequencia.R

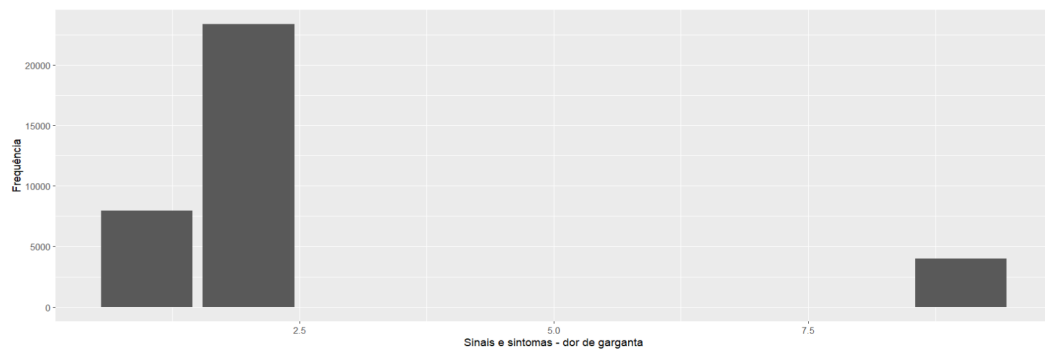


Figura 19 – Sintoma de Dor de garganta - Gráfico de Barras
Arquivo: Analise de Frequencia.R

2.4 Análise Probabilística

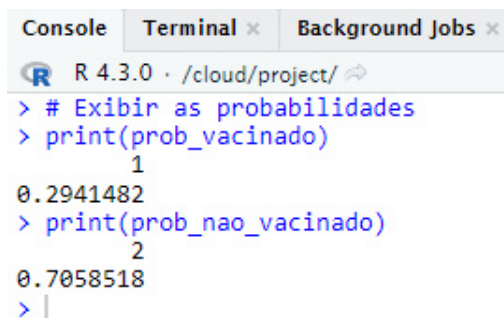
A análise probabilística é uma ferramenta fundamental para os profissionais da saúde, pois permite a compreensão e interpretação de dados com base em probabilidades. Ao analisar informações quantitativas, como taxas de incidência, prevalência de doenças ou eficácia de tratamentos, os profissionais de saúde podem identificar padrões, tendências e estimar riscos associados a diferentes condições de saúde. Durante a análise dos dados, foram utilizadas as funções **prop.table()**, empregada para calcular as proporções das frequências das variáveis, permitindo uma visão clara das distribuições dos dados. A função **subset()** foi utilizada para filtrar os dados com base em critérios específicos. A função **nrow()** foi aplicada para separar as linhas de interesse. O cálculo da soma foi realizado usando a função **sum()**. Para a visualização dos dados, as bibliotecas **ggplot2** e **geom_bar()** foram utilizadas para criar gráficos de barras. Os códigos e funções utilizadas podem ser encontradas no arquivo **Análise Probabilística.R**.

- **Estimativa de probabilidade de Vacinação**

Variáveis:

- Recebeu vacina contra gripe

Análise: A análise das probabilidades mostra que, a probabilidade de uma pessoa ter se vacinado contra gripe é de aproximadamente 0.294 (29.4%), enquanto a probabilidade de uma pessoa não ter se vacinado é de aproximadamente 0.706 (70.7%). Essas probabilidades indicam que a maioria das pessoas na amostra não se vacinou contra a gripe, já que a probabilidade de não ter se vacinado é maior do que a probabilidade de ter se vacinado. Campanhas de conscientização, é possível identificar que uma proporção significativa de pessoas não se vacinou contra a gripe. Isso pode ser um indicador de baixa adesão à vacinação. As autoridades de saúde podem usar esses dados para direcionar campanhas de conscientização, promovendo os benefícios da vacinação e incentivando as pessoas a se vacinarem.



```
Console Terminal x Background Jobs x
R 4.3.0 · /cloud/project/
> # Exibir as probabilidades
> print(prob_vacinado)
1
0.2941482
> print(prob_nao_vacinado)
2
0.7058518
> |
```

Figura 20 – Estimativa de probabilidade de Vacinação - Resultado

Arquivo: Análise Probabilística.R

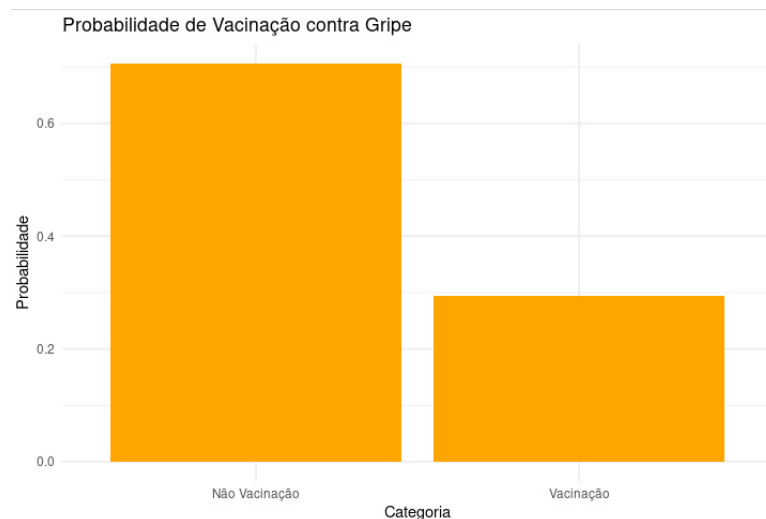


Figura 21 – Estimativa de probabilidade de Vacinação - Gráfico

Arquivo: Análise Probabilística.R

- Estimativa de probabilidade de Vacinação entre pardos e pretos

Variáveis:

- Recebeu vacina contra gripe
- Raça/Cor

Análise: Com base na análise realizada, a probabilidade de uma pessoa preta ou parda ter se vacinado contra a gripe é de aproximadamente 0,261 ou 26,1%. Isso significa que, considerando apenas os casos válidos de pessoas pretas ou pardas na amostra, cerca de 26,1% delas receberam a vacina contra a gripe. Identificação de disparidades, a análise da probabilidade de vacinação entre diferentes grupos étnicos/raciais pode ajudar a identificar possíveis disparidades na cobertura vacinal. Se a probabilidade de vacinação for significativamente menor em um grupo específico, isso pode indicar a necessidade de intervenções direcionadas para aumentar a conscientização e o acesso à vacinação nesse grupo. Planejamento de campanhas de vacinação, compreender a probabilidade de vacinação em diferentes grupos étnicos/raciais pode auxiliar na formulação de estratégias e campanhas de vacinação mais eficazes.

```
Console Terminal x Background Jobs x
R 4.3.0 · /cloud/project/
> # Exibir a probabilidade
> print(prob_vacinado_preta_parda)
[1] 0.2614582
>
```

Figura 22 – Estimativa de probabilidade de Vacinação entre pardos e pretos - Resultado

Arquivo: Análise Probabilística.R

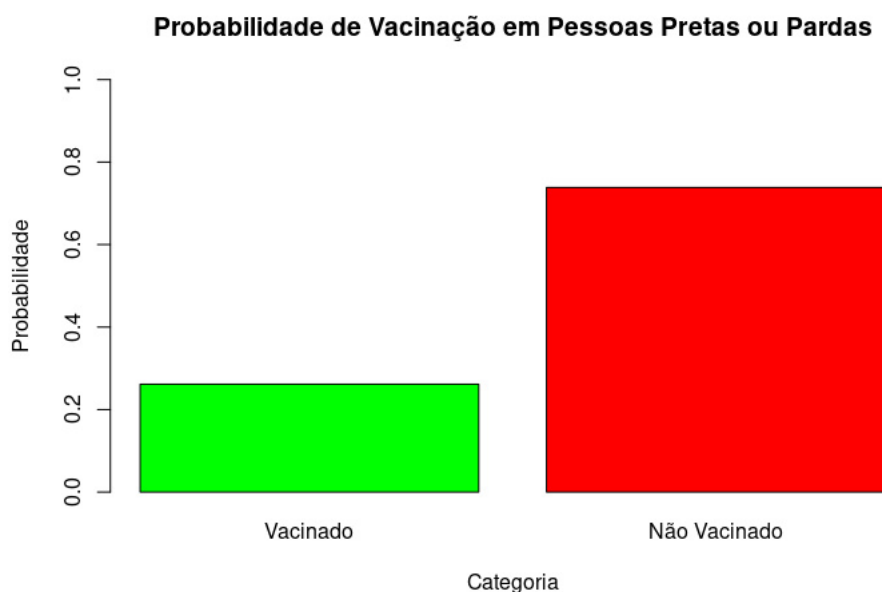


Figura 23 – Estimativa de probabilidade de Vacinação entre pardos e pretos - Gráfico
Arquivo: Análise Probabilística.R

- **Estimativa de probabilidade de Vacinação entre pardos e pretos do sexo feminino**

Variáveis:

- Recebeu vacina contra gripe
- Raça/Cor
- Sexo

Análise: A análise realizada calculou a probabilidade de uma pessoa preta ou parda do sexo feminino não ter se vacinado contra a gripe, considerando apenas os valores válidos (1 e 2) na variável VACINA. O resultado obtido foi 0.7138, o que indica que aproximadamente 71.4% das pessoas pretas ou pardas do sexo feminino não receberam a vacina contra a gripe. Essa informação é relevante para a área de saúde, pois permite identificar uma possível lacuna na imunização das pessoas pretas ou pardas do sexo feminino. Com base nesses dados, medidas podem ser tomadas para aumentar a conscientização e promover a vacinação entre esse grupo específico, visando melhorar a cobertura vacinal e reduzir o risco de doenças relacionadas à gripe. Isso pode incluir campanhas direcionadas, divulgação de informações sobre a importância da vacinação e facilitação do acesso à vacina para essa população.

```
Console Terminal x Background Jobs x
R 4.3.0 · /cloud/project/
> print(prob_ao_vacinado_parda_feminino)
[1] 0.7025624
>
```

Figura 24 – Estimativa de probabilidade de Vacinação entre pardos e pretos do sexo feminino - Resultado

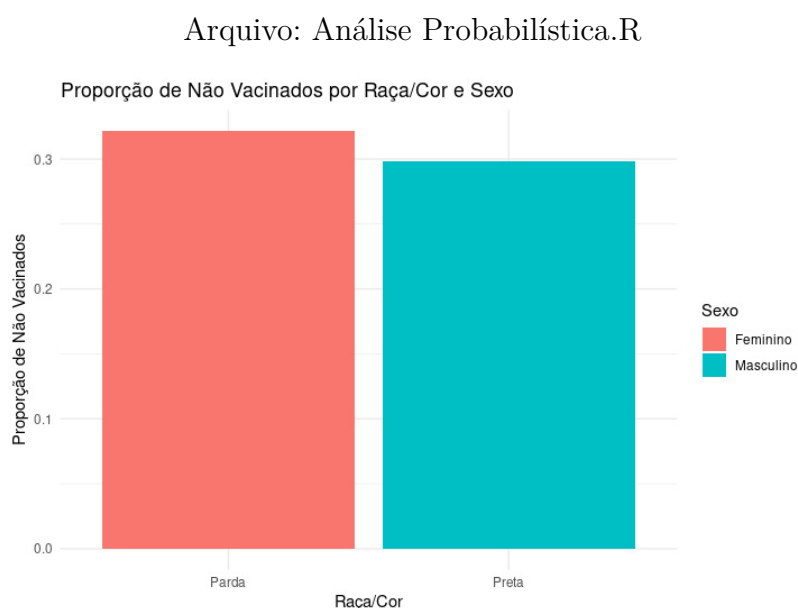


Figura 25 – Estimativa de probabilidade de não vacinados entre pessoas femininas da cor parda e pretos do sexo masculino - Gráfico

Arquivo: Análise Probabilística.R

- **Estimativa de Probabilidade de Sintomas: Febre, Tosse e Dor de Garganta:**

Variáveis:

- Sinais e sintomas - Febre
- Sinais e sintomas - Tosse
- Sinais e sintomas - Dor de garganta

Análise: Com base nos resultados da análise da estimativa de probabilidade de sintomas (febre, tosse e dor de garganta) em sua base de dados, podemos concluir que a febre é relatada por aproximadamente 81.94% dos pacientes, a tosse por cerca de 89.89% dos pacientes e a dor de garganta por cerca de 21.74% dos pacientes. Essas probabilidades podem ser utilizadas para identificar a prevalência desses sintomas, planejar recursos adequados na área de saúde e desenvolver intervenções e políticas direcionadas.

```
Console Terminal x Background Jobs x
R 4.2.2 · ~/
> # Exibindo os resultados
> print(paste("Febre:", format(probabilidade_febre * 100, digits = 4),
"%"))
[1] "Febre: 81.94 %"
> print(paste("Tosse:", format(probabilidade_tosse * 100, digits = 4),
"%"))
[1] "Tosse: 89.89 %"
> print(paste("Dor de Garganta:", format(probabilidade_garganta * 100, d
igits = 4), "%"))
[1] "Dor de Garganta: 21.74 %"
>
```

Figura 26 – Estimativa de Probabilidade de Sintomas: Febre, Tosse e Dor de Garganta - Resultado

Arquivo: Análise Probabilística.R

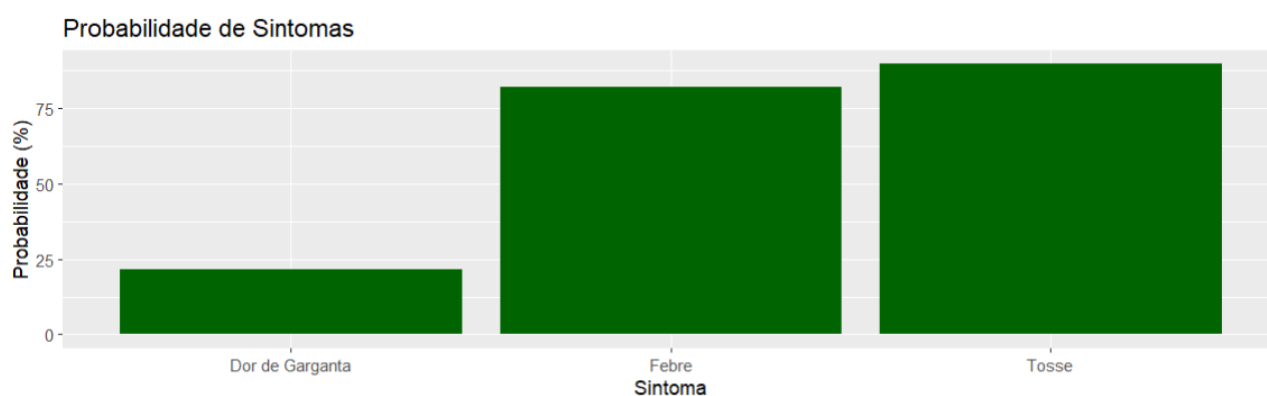


Figura 27 – Estimativa de Probabilidade de Sintomas: Febre, Tosse e Dor de Garganta - Gráfico

Arquivo: Análise Probabilística.R

- **Estimativa de Probabilidade de Vacinação contra a Gripe**

Variáveis:

- Recebeu vacina contra gripe

Análise: Com base na análise dos dados, concluímos que cerca de 23.25% dos pacientes analisados receberam a vacina contra a gripe. Isso indica uma adesão significativa à vacinação, o que é importante para reduzir a propagação da doença e suas complicações. Os profissionais podem utilizar essa informação para avaliar a cobertura vacinal e identificar grupos que possam estar com baixa adesão à vacinação. Com esses dados, podem ser desenvolvidas estratégias para aumentar a conscientização, direcionar campanhas e garantir a disponibilidade adequada de vacinas. Além disso, a análise contínua dessas probabilidades ao longo do tempo

permite monitorar a adesão à vacinação e avaliar a eficácia das intervenções em saúde pública relacionadas à gripe.

```

Console Terminal x Background Jobs x
R 4.2.2 · ~/
> # Exibindo o resultado
> print(paste("Vacinação contra Gripe:", format(probabilidade_vacina * 100, digits = 4), "%"))
[1] "Vacinação contra Gripe: 23.25 %"

```

Figura 28 – Estimativa de Probabilidade de Vacinação contra a Gripe - Resultado

Arquivo: Análise Probabilística.R

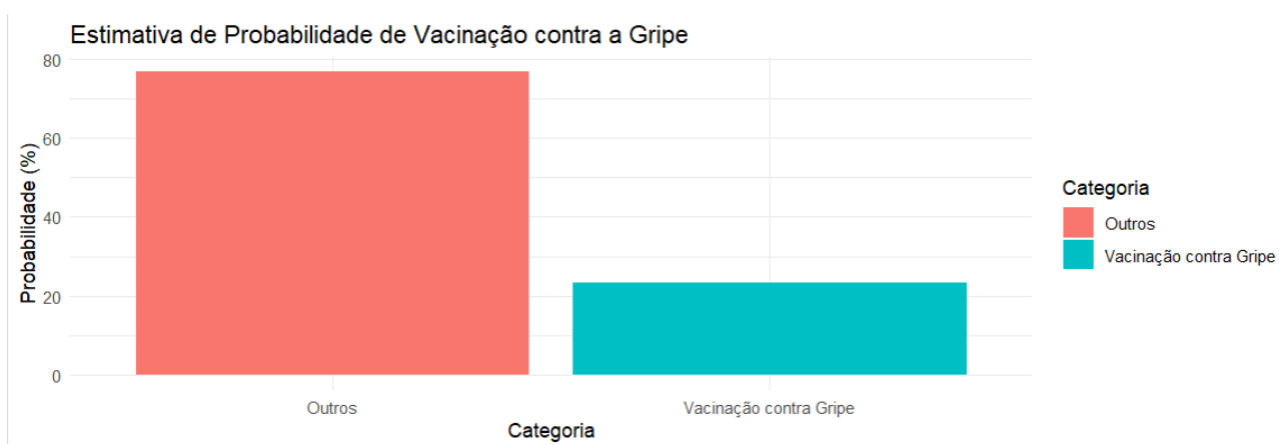


Figura 29 – Estimativa de Probabilidade de Vacinação contra a Gripe - Gráfico

Arquivo: Análise Probabilística.R

• Estimativa de Probabilidade de Categorias de Gestação

Variáveis:

- Gestante

Análise: Com base nos dados analisados, podemos inferir as probabilidades de diferentes categorias de gestação. Aproximadamente 0.5935% dos pacientes estão no 1º trimestre, enquanto cerca de 1.239% estão no 2º trimestre e aproximadamente 1.359% estão no 3º trimestre. A proporção de pacientes com idade gestacional desconhecida é de cerca de 0.0547%. Além disso, cerca de 25.28% dos pacientes não estão em gestação, e aproximadamente 68.61% não se aplicam à categoria de gestação. Elas podem auxiliar no planejamento de recursos, como leitos e profissionais especializados, com base na distribuição das gestações nas diferentes categorias. Além disso, as probabilidades estimadas permitem o monitoramento da saúde materno-infantil, facilitando a identificação de possíveis problemas. Com base nessas conclusões, os profissionais de saúde podem adaptar suas intervenções e cuidados de acordo com

o estágio gestacional, proporcionando suporte personalizado. Também é possível identificar grupos de risco, como gestantes em estágios avançados da gravidez ou com idade gestacional desconhecida, para a adoção de medidas preventivas específicas.

```

Console Terminal x Background Jobs x
R 4.2.2 · ~/
> # Exibindo os resultados
> print(paste("Gestação no 1º Trimestre:", format(probabilidade_trimestr
e1 * 100, digits = 4), "%"))
[1] "Gestação no 1º Trimestre: 0.5935 %"
> print(paste("Gestação no 2º Trimestre:", format(probabilidade_trimestr
e2 * 100, digits = 4), "%"))
[1] "Gestação no 2º Trimestre: 1.239 %"
> print(paste("Gestação no 3º Trimestre:", format(probabilidade_trimestr
e3 * 100, digits = 4), "%"))
[1] "Gestação no 3º Trimestre: 1.359 %"
> print(paste("Gestação com Idade Gestacional Ignorada:", format(probabi
lidade_gestacional_ignorado * 100, digits = 4), "%"))
[1] "Gestação com Idade Gestacional Ignorada: 0.0547 %"
> print(paste("Sem Gestação:", format(probabilidade_sem_gestacao * 100,
digits = 4), "%"))
[1] "Sem Gestação: 25.28 %"
> print(paste("Não se Aplica:", format(probabilidade_sem_aplicacao * 10
0, digits = 4), "%"))
[1] "Não se Aplica: 68.61 %"

```

Figura 30 – Estimativa de Probabilidade de Categorias de Gestação - Resultado

Arquivo: Análise Probabilística.R

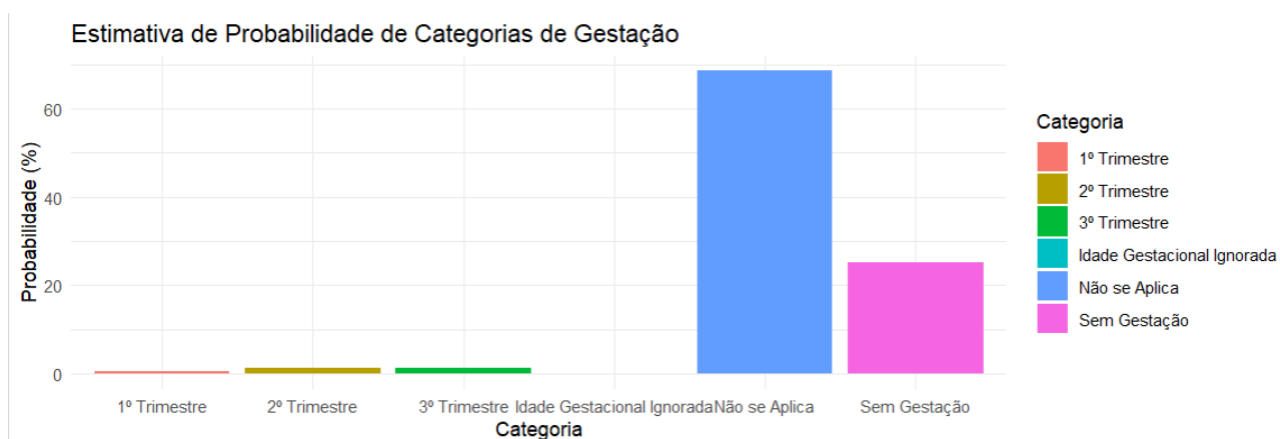


Figura 31 – Estimativa de Probabilidade de Categorias de Gestação - Gráfico

Arquivo: Análise Probabilística.R

2.5 Análise de Vacinação

A análise de vacinação desempenha um papel crucial na exploração de dados para profissionais da saúde. Ela fornece insights valiosos sobre a eficácia e a cobertura vacinal em determinada população, permitindo uma compreensão mais completa dos padrões de imunização e suas implicações para a saúde pública. Para as análises de vacinação foi utilizado as bibliotecas **dplyr** e **ggplot2**. Funções como **group_by**, **summarize**, **filter**, **mutate** e **complete.cases** do pacote **dplyr** são utilizadas para agrupar, resumir e filtrar os dados. Além disso, são calculadas taxas de vacinação, intervalos de tempo e proporções usando as funções **prop.table** e **sum** do R. Para a visualização dos resultados, são empregadas técnicas como gráficos de barras **geom_bar** e histogramas **geom_histogram** com o uso da biblioteca **ggplot2**. Os códigos e funções relacionados a Análise de Vacinação podem ser encontrados no arquivo **Análise de Vacinação.R**.

- **Data dos primeiros sintomas / Diagnóstico:** Observamos que nas datas dos primeiros sintomas a taxa de vacinação geral é baixa, com apenas alguns pacientes tendo recebido a vacina contra a gripe. Isso ressalta a necessidade de aumentar a conscientização sobre a importância da vacinação, especialmente para aqueles que apresentam a SRAG. Ao analisar a distribuição da vacinação por faixa etária, notamos que a vacinação é mais comum em pacientes mais jovens. Isso pode indicar que a conscientização sobre a importância da vacinação é maior entre os grupos mais jovens. Além disso, ao considerar a presença de sintomas respiratórios, como febre, tosse e dor de garganta, observamos que alguns pacientes vacinados relataram a presença desses sintomas. Isso pode sugerir a ocorrência de infecções respiratórias que não estão relacionadas à gripe ou uma possível falta de eficácia da vacina em alguns casos. Com base nesses insights, é importante que os profissionais de saúde reforcem a importância da vacinação contra a gripe para pacientes com SRAG, independentemente da idade. Além disso, é fundamental investigar as razões por trás da baixa taxa de vacinação e desenvolver estratégias de conscientização direcionadas. Também é recomendado avaliar a eficácia da vacinação em pacientes que apresentam sintomas respiratórios e investigar outras possíveis causas de infecções respiratórias.

```

> # Calcular a taxa de vacinação por faixa etária
> taxa_vacinacao <- INFLUD13%>%
+   group_by(NU_IDADE_N) %>%
+   summarize(Total = n(), Vacinados = sum(VACINA == 1), Taxa = Vacinados / Total * 100)
> # Carregar o pacote necessário
> library(dplyr)
>
> # Calcular a taxa de vacinação por faixa etária
> taxa_vacinacao <- INFLUD13%>%
+   group_by(NU_IDADE_N) %>%
+   summarize(Total = n(), Vacinados = sum(VACINA == 1), Taxa = Vacinados / Total * 100)
>
> # Imprimir os resultados
> print(taxa_vacinacao)
# A tibble: 160 x 4
  NU_IDADE_N Total Vacinados Taxa
  <dbl> <int> <int> <dbl>
1     1000     1         0  0
2     1008     1         0  0
3     2000    125        NA  NA
4     2001     6         1 16.7
5     2002     7         NA  NA
6     2003     6         1 16.7
7     2004    12         0  0
8     2005    11         2 18.2
9     2006    21         NA  NA
10    2007    21         1  4.76

```

Figura 32 – Data dos Primeiros / Diagnósticos - Resultado

Arquivo: Analise de Vacinação.R

- **Gestação / Vacinação:** Com base nos resultados retornados, podemos obter alguns insights relevantes para profissionais da saúde lidando com SRAG. O gráfico de barras mostra a distribuição da vacinação entre as diferentes categorias de gestação. Podemos observar o número de pacientes vacinados e não vacinados em cada categoria. Essa informação pode ajudar os profissionais a entenderem a cobertura vacinal e identificar possíveis lacunas na vacinação em grupos específicos, como gestantes. Observamos que a categoria "5" (gestantes) apresenta um número significativo de pacientes vacinados. Isso ressalta a importância da vacinação em gestantes para prevenir complicações da SRAG, especialmente durante a gravidez, quando as mulheres estão mais suscetíveis a infecções respiratórias. Observamos baixa cobertura vacinal em outras categorias como "1" (1º trimestre), "2" (2º trimestre) e "3" (3º trimestre) apresentando um número menor de pacientes vacinados em comparação com a categoria "5" (gestantes) podendo indicar a necessidade de esforços adicionais para aumentar a conscientização e a cobertura vacinal nessas categorias de gestação. Esses insights podem ajudar os profissionais de saúde a avaliar a eficácia da campanha de vacinação e direcionar esforços para aumentar a cobertura vacinal em grupos específicos, como gestantes e mulheres em diferentes estágios da gravidez.

```

Console Terminal x Background Jobs x
R 4.3.0 · /cloud/project/
> # Contar o número de vacinados e não vacinados por categoria de gestação
> vaccinacao_gestante <- INFLUD13 %>%
+   group_by(CS_GESTANT) %>%
+   summarize(Vacinados = sum(VACINA == 1, na.rm = TRUE),
+             Nao_Vacinados = sum(VACINA == 2))
>
> print(vaccinacao_gestante)
# A tibble: 8 x 3
  CS_GESTANT Vacinados Nao_Vacinados
  <dbl>      <int>      <int>
1         1         61          NA
2         2        209          NA
3         3        258          NA
4         4          2         13
5         5       2349          NA
6         6       5467          NA
7         9        153          NA
8        NA          1           0
> |

```

Figura 33 – Gestação / Vacinação contra Gripe - Resultado

Arquivo: Analise de Vacinação.R

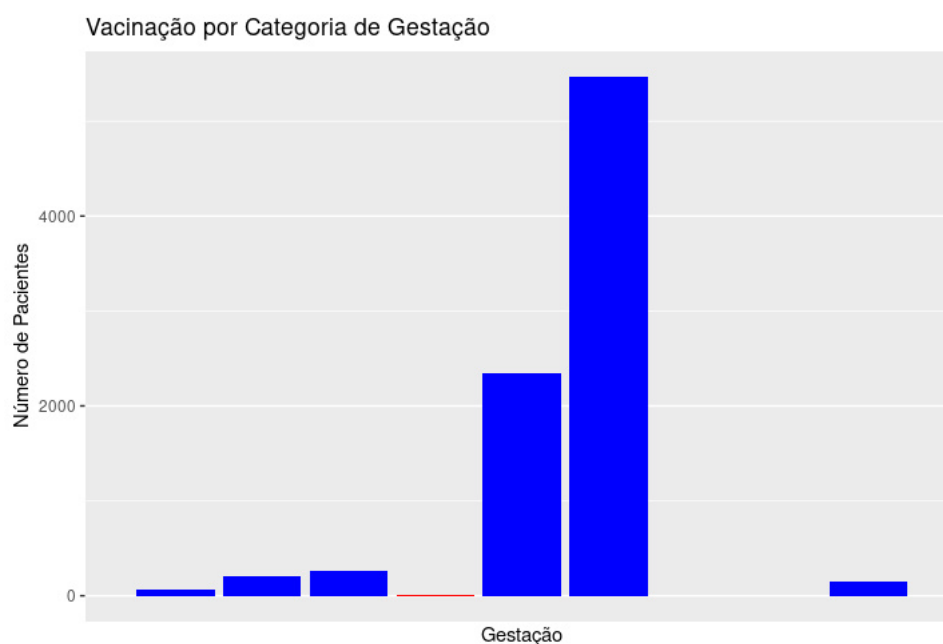


Figura 34 – Gestação / Vacinação contra Gripe - Gráfico

Arquivo: Analise de Vacinação.R

- **Intervalo de tempo / Vacinação contra Gripe:** Ao analisar os resultados apresentados, podemos obter insights relevantes para um profissional da saúde lidando com a SRAG. A data de notificação dos casos permite identificar variações temporais e possíveis surtos da doença. A data dos primeiros sintomas ou diagnóstico ajuda a entender a temporalidade da doença e o tempo de detecção e notificação dos casos. A idade dos pacientes pode indicar grupos etários mais suscetíveis ou com maior gravidade da SRAG. A distribuição por sexo pode revelar diferenças na incidência

e gravidade da doença entre homens e mulheres. A variável de gestação destaca a importância de proteger a saúde das gestantes. A raça/cor permite identificar disparidades étnico-raciais na ocorrência da SRAG. A escolaridade indica o acesso à informação e a necessidade de estratégias educativas adequadas. Esses insights auxiliam na compreensão da epidemiologia da SRAG, na identificação de grupos de risco e na implementação de medidas preventivas e de controle direcionadas.

R 4.3.0 · /cloud/project/ ↗												
> print(intervalo_tempo)												
# A tibble: 35,684 × 109												
	DT_NOTIFIC	ID_MUNICIP	SEM_NOT	NU_ANO	SG_UF	DT_SIN_PRI	DT_NASC	NU_IDADE_N	CS_SEXO	CS_GESTANT	CS_RACA	CS_ESCOL_N
	<chr>	<dbl>	<dbl>	<dbl>	<dbl>	<date>	<chr>	<dbl>	<chr>	<dbl>	<dbl>	<dbl>
1	04/02/2013	110020	201306	2013	11	2013-02-01	14/02/...	4026	M	6	1	3
2	08/04/2013	110004	201315	2013	11	2013-03-05	16/02/...	4030	F	5	4	2
3	30/03/2013	110004	201313	2013	11	2013-03-30	14/11/...	4057	F	5	1	0
4	26/03/2013	110020	201313	2013	11	2013-03-21	28/10/...	4001	M	6	NA	10
5	11/03/2013	110020	201311	2013	11	2013-03-07	14/11/...	4001	M	6	5	10
6	08/03/2013	110020	201310	2013	11	2013-03-02	02/03/...	4002	M	6	5	10
7	08/03/2013	110020	201310	2013	11	2013-03-01	02/09/...	3005	F	6	5	10
8	08/03/2013	110020	201310	2013	11	2013-02-28	02/01/...	3001	F	6	5	10
9	02/04/2013	110020	201314	2013	11	2013-03-29	18/02/...	4023	M	6	4	3
10	22/04/2013	110020	201317	2013	11	2013-04-18	13/12/...	4027	F	5	4	2

Figura 35 – Intervalo de tempo / Vacinação contra Gripe - Resultado

Arquivo: Analise de Vacinação.R

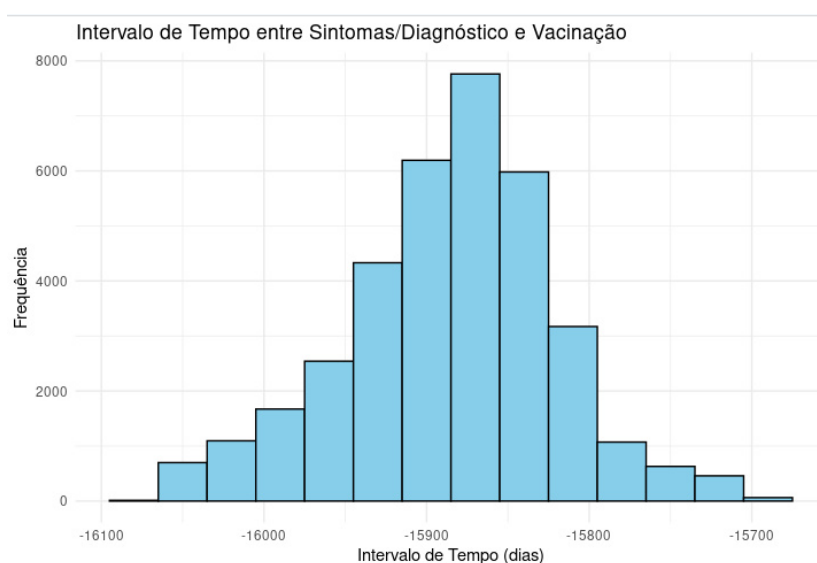


Figura 36 – Intervalo de tempo / Vacinação contra Gripe - Gráfico

Arquivo: Analise de Vacinação.R

- **Taxa de vacinação para homens e mulheres separadamente:** Com base nos resultados da taxa de vacinação por sexo, é possível observar que as mulheres apresentaram uma taxa de vacinação (26,2%) maior do que os homens (21,5%). Esse dado indica uma possível diferença na adesão à vacinação entre os sexos, o que pode ser relevante para um profissional da saúde lidando com a SRAG. É importante que o profissional utilize essas informações para direcionar estratégias de conscientização e

educação sobre a importância da vacinação, buscando aumentar a adesão entre ambos os sexos. Além disso, é essencial monitorar continuamente os dados de vacinação por sexo e avaliar o impacto das ações implementadas, visando garantir uma cobertura vacinal adequada e o controle da síndrome.

```

Console Terminal x Background Jobs x
R 4.3.0 · /cloud/project/
> print(taxa_vacinacao_sexo)
# A tibble: 3 × 4
  CS_SEXO Total Vacinados Taxa
  <chr>    <int>    <int> <dbl>
1 F      17910     4684  26.2
2 I         8         1  12.5
3 M      17766     3815  21.5
> |

```

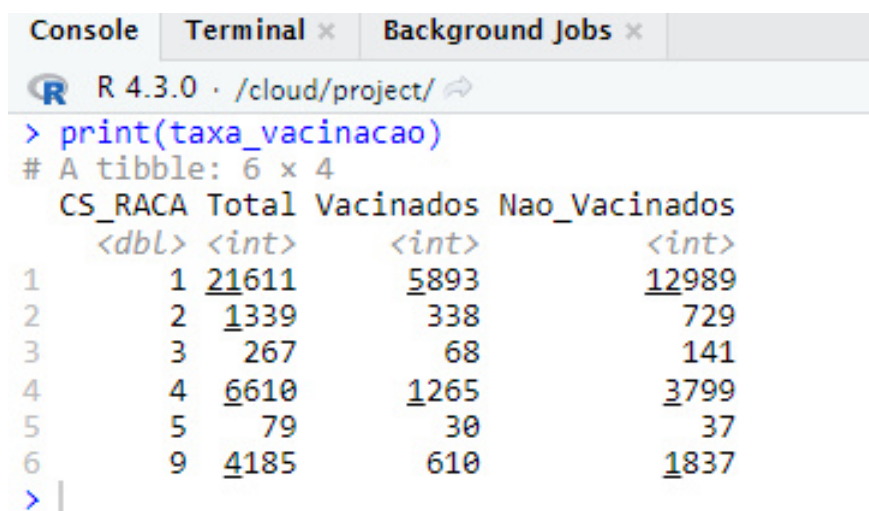
Figura 37 – Taxa de vacinação para homens e mulheres separadamente - Resultado
Arquivo: Analise de Vacinação.R



Figura 38 – Taxa de vacinação para homens e mulheres separadamente - Gráfico
Arquivo: Analise de Vacinação.R

- **Vacinação contra a gripe entre diferentes grupos raciais/étnicos - Resultado:** Com base nos resultados da análise da relação entre raça/cor (CS_RACA) e a vacinação contra a gripe (VACINA), podemos obter insights relevantes para profissionais de saúde lidando com a SRAG. Observamos que a maioria dos pacientes pertence às raças/cor 1 e 4, seguidas por outras raças/cor com menor representação. Ao calcular a taxa de vacinação em cada grupo racial/cor, podemos identificar possíveis disparidades na adesão à vacinação. A análise comparativa entre o número de indivíduos vacinados e não vacinados em cada grupo pode revelar grupos com menor cobertura vacinal. Com essas informações, os profissionais de saúde podem direcionar

esforços para aumentar a conscientização e a adesão à vacinação em populações específicas, visando prevenir e controlar a SRAG.



```

R 4.3.0 · /cloud/project/
> print(taxa_vacinacao)
# A tibble: 6 × 4
  CS_RACA Total Vacinados Nao_Vacinados
  <dbl> <int> <int> <int>
1 1 21611 5893 12989
2 2 1339 338 729
3 3 267 68 141
4 4 6610 1265 3799
5 5 79 30 37
6 9 4185 610 1837
> |

```

Figura 39 – Vacinação contra a gripe entre diferentes grupos raciais/étnicos - Resultado
Arquivo: Analise de Vacinação.R

- **Paciente recebeu a vacina contra a gripe:** Com base nos resultados da análise da variável "VACINA" que indica se o paciente recebeu a vacina contra a gripe, podemos obter insights relevantes para profissionais de saúde lidando com a SRAG. Observamos que a maioria dos pacientes não recebeu a vacina (VACINA=2) e apenas uma proporção relativamente baixa foi vacinada (VACINA=1). Além disso, houve um número significativo de casos em que a informação sobre a vacinação estava faltando (VACINA=NA). Isso sugere a necessidade de uma maior conscientização e promoção da vacinação contra a gripe entre os pacientes com SRAG. Os profissionais de saúde podem se concentrar em fornecer informações sobre os benefícios da vacinação, esclarecer dúvidas e garantir o acesso adequado à vacina, visando aumentar a taxa de vacinação e reduzir a incidência e gravidade da doença.

```

Console Terminal x Background Jobs x
R 4.3.0 · /cloud/project/
> print(contagem_vacina)
# A tibble: 4 × 2
  VACINA Contagem
  <dbl>   <int>
1     1     8500
2     2    20397
3     9     6787
4    NA      879
> print(taxa_vacinacao)
# A tibble: 4 × 3
  VACINA Contagem Taxa
  <dbl>   <int> <dbl>
1     1     8500  23.2
2     2    20397  55.8
3     9     6787  18.6
4    NA      879   2.40
> |

```

Figura 40 – Paciente recebeu a vacina contra a gripe - Resultado
Arquivo: Analise de Vacinação.R

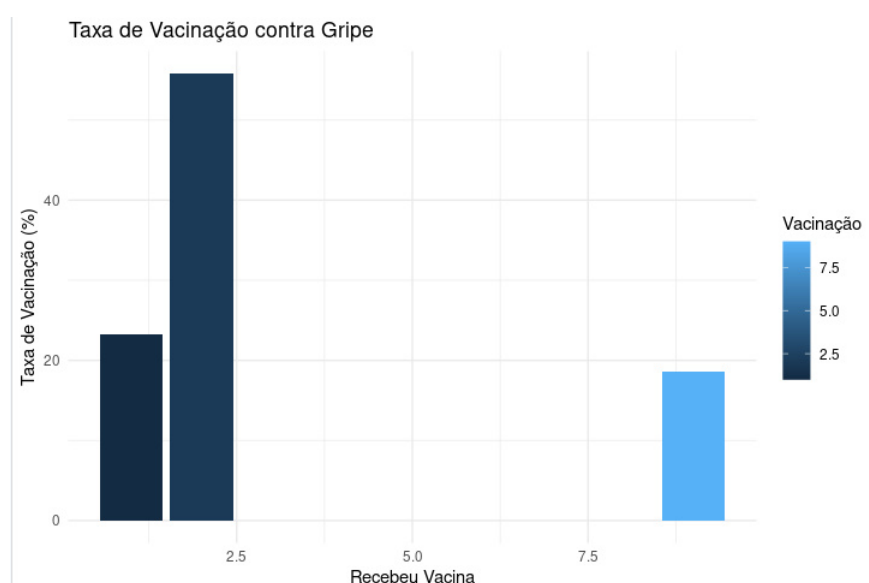


Figura 41 – Paciente recebeu a vacina contra a gripe - Gráfico
Arquivo: Analise de Vacinação.R

- **Vacinação entre pacientes com e sem febre:** Com base nos resultados da tabela de taxa de vacinação por presença de febre, podemos obter alguns insights relevantes para profissionais de saúde que lidam com SRAG. Observamos que a taxa de vacinação contra a gripe é mais alta para os pacientes sem febre (57.94%) em comparação com aqueles que apresentaram febre (23.49%). Isso sugere que a presença de febre pode influenciar negativamente a decisão de receber a vacina. Os

profissionais de saúde podem usar esses resultados para conscientizar os pacientes sobre a importância da vacinação contra a gripe, especialmente entre aqueles com febre. Estratégias direcionadas, informações claras sobre os benefícios da vacinação e pontos de vacinação convenientes para pacientes com SRAG podem aumentar a taxa de vacinação.

```
R 4.3.0 · /cloud/project/
> # SÉTIMA ANÁLISE
> # Calcular a taxa de vacinação entre pacientes com e sem febre
> tabela <- table(INFLUD13$FEBRE, INFLUD13$VACINA)
> taxa_vacinacao <- prop.table(tabela, margin = 1) * 100
> print(taxa_vacinacao)
```

	1	2	9
1	23.49363	57.93773	18.56865
2	26.35686	55.43610	18.20704
9	15.22310	30.70866	54.06824

Figura 42 – Vacinação entre pacientes com e sem febre - Resultado

Arquivo: Analise de Vacinação.R

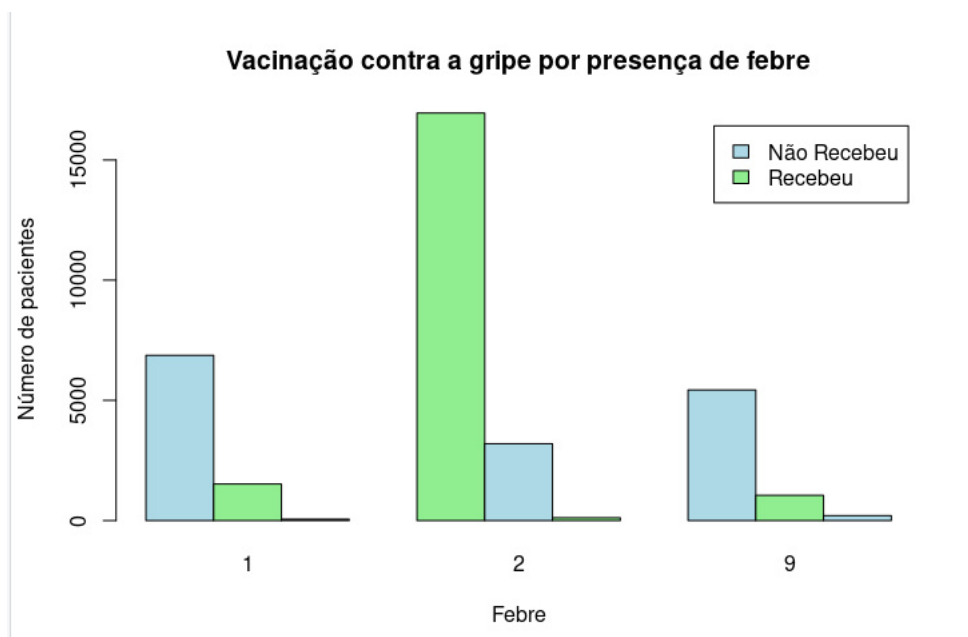


Figura 43 – Vacinação entre pacientes com e sem febre - Gráfico

Arquivo: Analise de Vacinação.R

- **Vacinação entre pacientes com e sem tosse:** Com base nos resultados da tabela de taxa de vacinação por presença de tosse, podemos obter insights relevantes para profissionais de saúde que lidam com SRAG. Os resultados revelam que a taxa de vacinação contra a gripe é mais baixa entre os pacientes com tosse (24.26%) em comparação com aqueles sem tosse (57.89%). Isso pode indicar uma possível relação entre a presença de tosse e a decisão de receber a vacina contra a gripe. Os profissionais de saúde podem utilizar esses resultados como um alerta para a importância de

intensificar as campanhas de vacinação entre os pacientes com SRAG que apresentam tosse. É fundamental fornecer informações claras sobre os benefícios da vacinação e esclarecer eventuais preocupações ou dúvidas dos pacientes. Além disso, estratégias direcionadas, como o estabelecimento de pontos de vacinação específicos para pacientes com sintomas respiratórios, podem ser adotadas para aumentar a adesão à vacinação. Esses insights também reforçam a necessidade de uma vigilância contínua dos sintomas, incluindo a tosse, como indicador importante para o monitoramento e a prevenção da propagação da gripe e de outras doenças respiratórias. Ao personalizar o atendimento e oferecer informações e suporte adequados, os profissionais de saúde podem desempenhar um papel crucial na promoção da vacinação contra a gripe entre os pacientes com SRAG, visando a proteção individual e coletiva contra essa doença sazonal.

```

Console Terminal x Background Jobs x
R 4.3.0 · /cloud/project/
> print(taxa_vacinacao)

      1      2      9
1 24.26374 57.89179 17.84447
2 21.03517 54.21367 24.75116
9 11.67665 24.55090 63.77246
>

```

Figura 44 – Vacinação entre pacientes com e sem tosse - Resultado

Arquivo: Analise de Vacinação.R

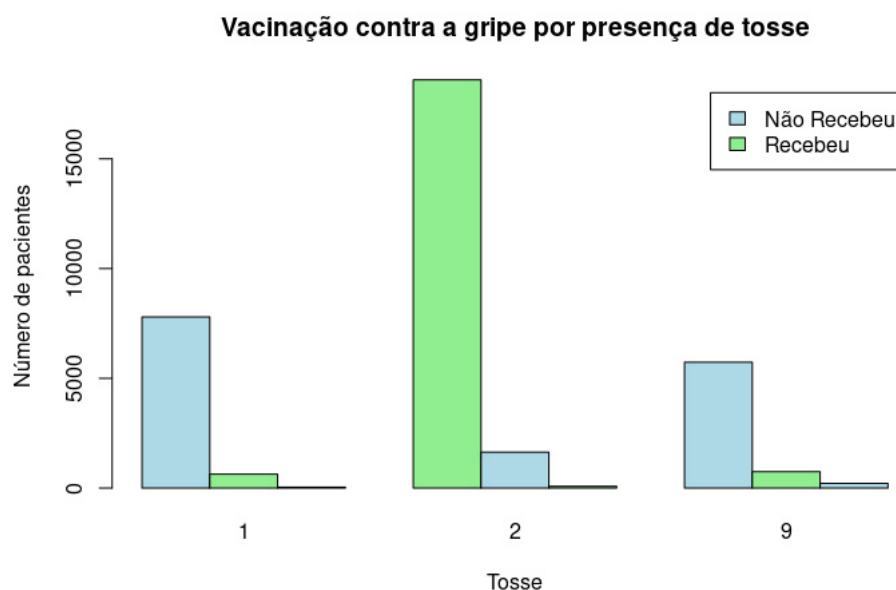


Figura 45 – Vacinação entre pacientes com e sem tosse - Gráfico

Arquivo: Analise de Vacinação.R

- **Vacinação entre pacientes com e sem dor de garganta:** Com base nos resultados da tabela de proporções de vacinação por presença de dor de garganta, podemos obter insights relevantes para profissionais de saúde que lidam com SRAG. Os resultados revelam que a taxa de vacinação contra a gripe é mais baixa entre os pacientes com dor de garganta (27.13%) em comparação com aqueles sem dor de garganta (62.75%). Isso sugere uma possível relação entre a presença de dor de garganta e a decisão de receber a vacina contra a gripe. Os profissionais de saúde devem estar cientes dessa associação e dedicar esforços adicionais para incentivar a vacinação entre os pacientes com SRAG que apresentam dor de garganta. Os insights obtidos indicam a importância de uma abordagem personalizada para a promoção da vacinação. Estratégias como a identificação proativa de pacientes com dor de garganta e a oferta de informações específicas sobre a vacinação podem ser eficazes para aumentar a adesão à vacinação nesse grupo.

```

Console Terminal x Background Jobs x
R 4.3.0 · /cloud/project/
> proporcoes

      1      2      9
1 0.2712644 0.6274585 0.1012771
2 0.2423660 0.5629723 0.1946617
9 0.1602238 0.5353510 0.3044252
> |

```

Figura 46 – Vacinação entre pacientes com e sem dor de garganta - Resultado

Arquivo: Analise de Vacinação.R

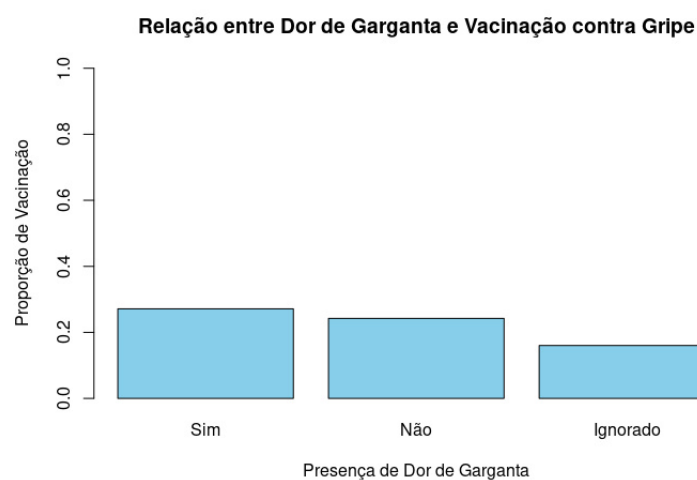


Figura 47 – Vacinação entre pacientes com e sem dor de garganta - Gráfico

Arquivo: Analise de Vacinação.R

2.6 Análise de Sintomas

A análise de sintomas é uma ferramenta essencial para profissionais da saúde entenderem a ocorrência e a relação entre diferentes sintomas em uma população. Utilizando bibliotecas como **dplyr** e **ggplot2**, é possível agrupar, filtrar e calcular proporções de sintomas, como dor de garganta, febre e tosse, em diferentes grupos de gestantes. Para visualizar os resultados, são aplicados gráficos de barras e histogramas da biblioteca **ggplot2**. Para as análises de sintomas foram utilizado uma amostra de **1000** pessoas. Essas representações gráficas ajudam a compreender visualmente as proporções e padrões identificados. Os códigos e funções utilizadas podem ser encontradas no arquivo **Análise de Sintomas.R**.

- **Relação entre os sintomas (febre, tosse e dor de garganta) e a Raça/Cor:**

Variáveis:

- Raça/Cor
- Sinais e sintomas - Febre
- Sinais e sintomas - Tosse
- Sinais e sintomas - Dor de Garganta

Análise: Nesta análise, examinamos a relação entre as variáveis GARGANTA, FEBRE, TOSSE e Raça/Cor. Com base nos dados fornecidos, podemos observar algumas informações relevantes:

- **Relação entre Dor de Garganta e Raça/cor:** Em relação a Dor de Garganta e a Raça/Cor, foram analisados os diferentes valores de Dor de Garganta (1, 2 e 9) para cada grupo racial (Raça/Cor 1 a 5, 9). Ao observar as proporções de GARGANTA para cada valor de Raça/Cor, identificamos a distribuição das condições de garganta em cada grupo racial. Para Raça/Cor, as proporções de Dor de Garganta foram: GARGANTA = 1: 0.124597207, GARGANTA = 2: 0.296455424 e GARGANTA = 9: 0.047261010. Com base nos dados analisados, podemos observar variações significativas na prevalência de diferentes condições de dor de garganta em relação aos diferentes grupos raciais. Essa informação ressalta a importância de considerar fatores raciais ao avaliar e tratar problemas relacionados à garganta, em busca de proporcionar um cuidado mais equitativo e personalizado aos pacientes de diferentes origens raciais.
- **Relação entre Febre e Raça/Cor:** Em relação a FEBRE e Raça/Cor, foram analisadas as proporções de Febre(1, 2 e 9) para cada grupo racial. Ao examinar essas proporções, podemos identificar a distribuição da gravidade da febre em relação à raça. Para Raça/Cor, as proporções de Febre foram: FEBRE = 1:

0.124597207, FEBRE = 2: 0.007518797 e FEBRE = 9: 0.001074114. Esses números sugerem a prevalência de diferentes níveis de gravidade de febre em diferentes grupos raciais. Os resultados sugerem que certos grupos raciais podem apresentar uma maior predisposição a desenvolver febres mais graves em comparação com outros grupos. Essa informação é relevante para os profissionais de saúde, pois destaca a necessidade de abordagens de manejo da febre personalizadas, levando em consideração os diferentes grupos raciais. Compreender essas variações pode auxiliar na identificação precoce, monitoramento e tratamento adequado da febre, visando melhores resultados clínicos e promovendo a saúde equitativa para todos os pacientes, independentemente de sua raça.

- **Relação entre Tosse e Raça/Cor:** Devido à natureza dos dados analisados, não foi possível identificar uma relação clara entre tosse e Raça/Cor. Embora tenhamos informações detalhadas sobre a frequência de tosse para diferentes valores de GARGANTA e FEBRE, os dados não nos permitem tirar conclusões diretas sobre a associação entre tosse e raça.

```

R 4.2.2 · ~/
, , GARGANTA = 1, CS_RACA = 1

      TOSSE
FEBRE    1      2      9
  1 0.124597207 0.007518797 0.001074114
  2 0.018259936 0.001074114 0.000000000
  9 0.000000000 0.000000000 0.000000000

, , GARGANTA = 2, CS_RACA = 1

      TOSSE
FEBRE    1      2      9
  1 0.296455424 0.034371643 0.002148228
  2 0.088077336 0.021482277 0.000000000
  9 0.003222342 0.000000000 0.000000000

, , GARGANTA = 9, CS_RACA = 1

      TOSSE
FEBRE    1      2      9
  1 0.047261010 0.002148228 0.000000000
  2 0.003222342 0.000000000 0.001074114
  9 0.001074114 0.000000000 0.000000000

```

Figura 48 – Relação entre os sintomas (febre, tosse e dor de garganta) e a raça/cor - Resultado

Arquivo: Analise de Sintomas.R

Gráfico: O gráfico de barras apresenta, por meio de cores, a proporção dos três sintomas em relação à raça/cor das pessoas. Os códigos utilizados são:

- 1: Indica presença do sintoma
- 2: Indica ausência do sintoma
- 3: Indica informação ignorada

Portanto, o código "1.1.1" indica a proporção de indivíduos que possuem todos os sintomas, enquanto o código "2.2.2" indica a proporção de indivíduos que não apresentam nenhum sintoma.

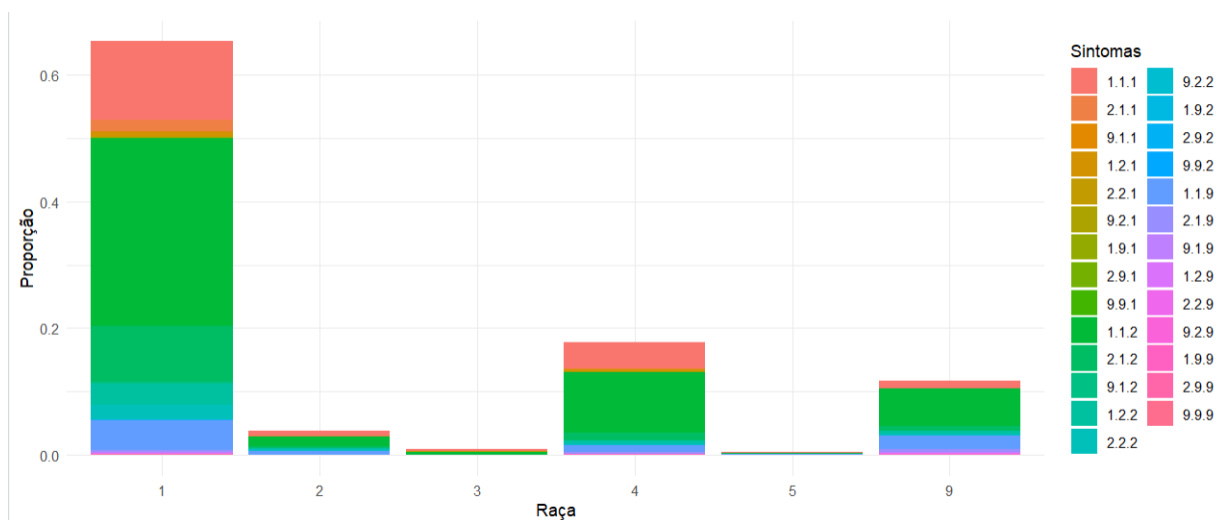


Figura 49 – Relação entre os sintomas (febre, tosse e dor de garganta) e a raça/cor - Gráfico

Arquivo: Analise de Sintomas.R

• Relação entre os sintomas (febre, tosse e dor de garganta) e Sexo:

Variáveis:

- Sexo
- Sinais e sintomas - Febre
- Sinais e sintomas - Tosse
- Sinais e sintomas - Dor de Garganta

Análise:

- **Relação entre Dor de Garganta e Sexo:** A análise dos dados não revela uma diferença clara nas proporções de problemas de dor de garganta entre os sexos. No entanto, observa-se que a ausência de dor de garganta, tende a

apresentar proporções mais altas em geral, tanto para homens quanto para mulheres. Embora não haja uma associação clara entre garganta e sexo, a ausência de dor de garganta, pode auxiliar na triagem e no direcionamento de recursos e tratamentos adequados. Isso permite que os profissionais de saúde foquem em outras possíveis causas ou sintomas relacionados à saúde do paciente, além da dor de garganta.

- **Relação entre Febre e Sexo:** A análise revela que mulheres têm maior proporção de febre, independentemente da garganta, indicando maior prevalência entre elas. Ambos os sexos apresentam proporções mais baixas de ausência de febre. A baixa proporção de dados indisponíveis para febre indica que profissionais de saúde possuem informações suficientes para avaliar febre em pacientes de ambos os sexos e categorias de garganta. Essa análise auxilia profissionais ao considerarem o sexo dos pacientes, direcionando recursos e investigações adicionais para mulheres com febre. Também enfatiza a importância de avaliar sintomas relacionados à febre considerando o sexo do paciente, proporcionando abordagens personalizadas e eficazes no diagnóstico e tratamento, aprimorando a qualidade da assistência médica.
- **Relação entre Tosse e Sexo:** Homens apresentam uma maior proporção de tosse em comparação com mulheres, indicando maior prevalência de tosse entre os homens. A ausência de tosse é mais comum para ambos os sexos. A proporção de casos sem informação sobre tosse é baixa para ambos os sexos. Essa análise fornece insights para profissionais de saúde ao avaliar pacientes com sintomas de tosse, permitindo direcionar recursos e tratamentos adequados para homens com tosse. A disponibilidade de dados suficientes sobre tosse é crucial para um diagnóstico preciso e tratamento adequado. Considerar o sexo do paciente ao avaliar sintomas de tosse leva a uma abordagem personalizada e eficaz no diagnóstico e tratamento de condições relacionadas, melhorando a qualidade do cuidado de saúde oferecido pelos profissionais.

```
, , GARGANTA = 1, CS_SEXO = F

      TOSSE
FEBRE      1      2      9
1 0.102590674 0.004145078 0.000000000
2 0.012435233 0.001036269 0.000000000
9 0.000000000 0.000000000 0.000000000

, , GARGANTA = 2, CS_SEXO = F

      TOSSE
FEBRE      1      2      9
1 0.223834197 0.013471503 0.000000000
2 0.051813472 0.011398964 0.000000000
9 0.000000000 0.000000000 0.000000000

, , GARGANTA = 9, CS_SEXO = F

      TOSSE
FEBRE      1      2      9
1 0.045595855 0.002072539 0.004145078
2 0.003108808 0.000000000 0.000000000
9 0.001036269 0.000000000 0.003108808
```

Figura 50 – Relação entre os sintomas (febre, tosse e dor de garganta) e o Sexo - Resultado
Arquivo: Analise de Sintomas.R

Gráfico: O gráfico de barras apresenta, por meio de cores, a proporção dos três sintomas em relação ao Sexo das pessoas. Os códigos utilizados são:

- 1: Indica presença do sintoma
- 2: Indica ausência do sintoma
- 3: Indica informação ignorada

Portanto, o código "1.1.1" indica a proporção de indivíduos que possuem todos os sintomas, enquanto o código "2.2.2" indica a proporção de indivíduos que não apresentam nenhum sintoma.

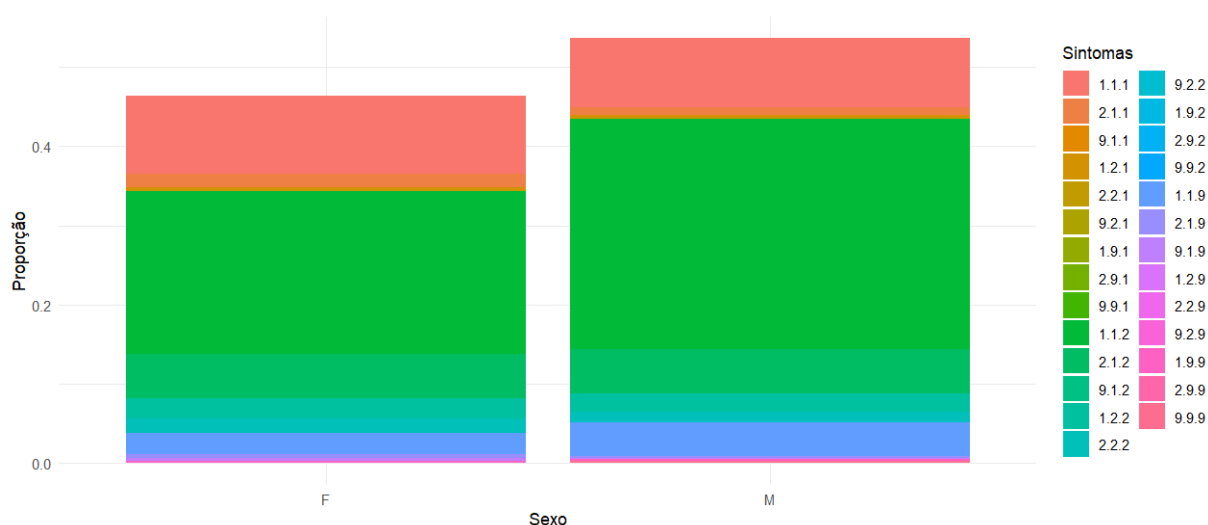


Figura 51 – Relação entre os sintomas (febre, tosse e dor de garganta) e a Sexo - Gráfico
Arquivo: Analise de Sintomas.R

- **Relação entre os sintomas (febre, tosse e dor de garganta) e Gestante:**

Variáveis:

- Gestante
- Sinais e sintomas - Febre
- Sinais e sintomas - Tosse
- Sinais e sintomas - Dor de Garganta

Análise:

- **Relação entre Dor de Garganta e Gestante:** Observando as proporções, é possível notar que a presença de dor de garganta é baixa em todas as categorias de gestante. Entre as gestantes analisadas, algumas apresentaram dor de garganta, enquanto outras não. Não há nenhum caso registrado de dor de garganta entre as gestantes do 1º trimestre. A presença de dor de garganta é mais significativa em gestantes dos grupos do 2º trimestre e do 3º trimestre, representando cerca de 0,7% e 0,1% das amostras selecionadas, respectivamente. Essa relação sugere uma possível associação entre a gravidez e a presença de dor de garganta, que pode ser explorada em estudos futuros.
- **Relação entre Febre e Gestante:** Analisando as proporções, observa-se que a ocorrência de febre é baixa em todas as categorias de gestante. Entre as gestantes analisadas, algumas apresentaram febre, enquanto outras não. Não foram registrados casos de febre entre as gestantes do 1º trimestre. Gestantes dos grupos do 2º trimestre e do 3º trimestre apresentam proporções mais elevadas de febre, representando cerca de 7% e 32,5% das amostras selecionadas, respectivamente. Esses resultados sugerem que a febre pode estar mais presente em gestantes de maior idade gestacional.
- **Relação entre Tosse e Gestante:** Analisando as proporções, observa-se que a tosse está presente em todas as categorias de gestante. A ocorrência de tosse é baixa entre gestantes do 1º trimestre, representando menos de 0,5% das amostras selecionadas. No entanto, gestantes dos grupos do 2º trimestre e do 3º trimestre apresentam proporções mais elevadas de tosse, correspondendo a aproximadamente 10,9% e 64,2% das amostras selecionadas, respectivamente. Esses resultados sugerem que a tosse pode ser mais comum em gestantes de maior idade gestacional.


```

Console Terminal x Background Jobs x
R 4.2.2 · ~/
, , GARGANTA = 1, CS_GESTANT = 1

TOSSE
FEBRE      1      2      9
1 0.003105590 0.000000000 0.000000000
2 0.000000000 0.000000000 0.000000000
9 0.000000000 0.000000000 0.000000000

, , GARGANTA = 2, CS_GESTANT = 1

TOSSE
FEBRE      1      2      9
1 0.002070393 0.000000000 0.000000000
2 0.003105590 0.000000000 0.000000000
9 0.000000000 0.000000000 0.000000000

, , GARGANTA = 9, CS_GESTANT = 1

TOSSE
FEBRE      1      2      9
1 0.000000000 0.000000000 0.000000000
2 0.000000000 0.000000000 0.000000000
9 0.000000000 0.000000000 0.000000000

```

Figura 52 – Relação entre os sintomas (febre, tosse e dor de garganta) e Gestante - Resultado

Arquivo: Analise de Sintomas.R

Gráfico: O gráfico de barras apresenta, por meio de cores, a proporção dos três sintomas em relação a pessoas gestantes. Os códigos utilizados são:

- 1: Indica presença do sintoma
- 2: Indica ausência do sintoma
- 3: Indica informação ignorada

Portanto, o código "1.1.1" indica a proporção de indivíduos que possuem todos os sintomas, enquanto o código "2.2.2" indica a proporção de indivíduos que não apresentam nenhum sintoma.

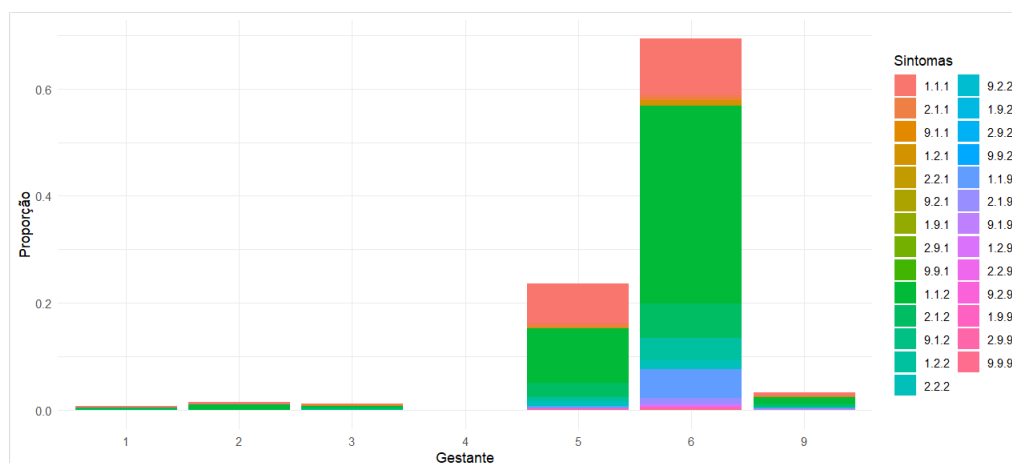


Figura 53 – Relação entre os sintomas (febre, tosse e dor de garganta) e Gestante - Gráfico

Arquivo: Analise de Sintomas.R

2.7 Testes de Hipótese

Os Testes de Hipótese são ferramentas importantes na pesquisa científica, incluindo a investigação da SRAG. Esses testes permitem a avaliação e exame de suposições, relações e diferenças entre variáveis relacionadas à SRAG. Para os testes de hipótese foram utilizados os testes **Qui-Quadrado**, teste **t de Student**, Teste **ANOVA** e uma amostra de **1000** pessoas. Para a visualização dos resultados, a biblioteca **ggplot2** foi utilizada para a criação de gráficos. Os códigos e funções relacionados aos testes de hipótese podem ser encontrados no arquivo **Testes de Hipotese.R**.

- **Hipótese N°1:** Existe diferença na média de idade (NU_IDADE_N) entre pacientes com febre (FEBRE) e sem febre (FEBRE)?

Hipótese nula (H0): Não há diferença na média de idade entre pacientes com febre e sem febre.

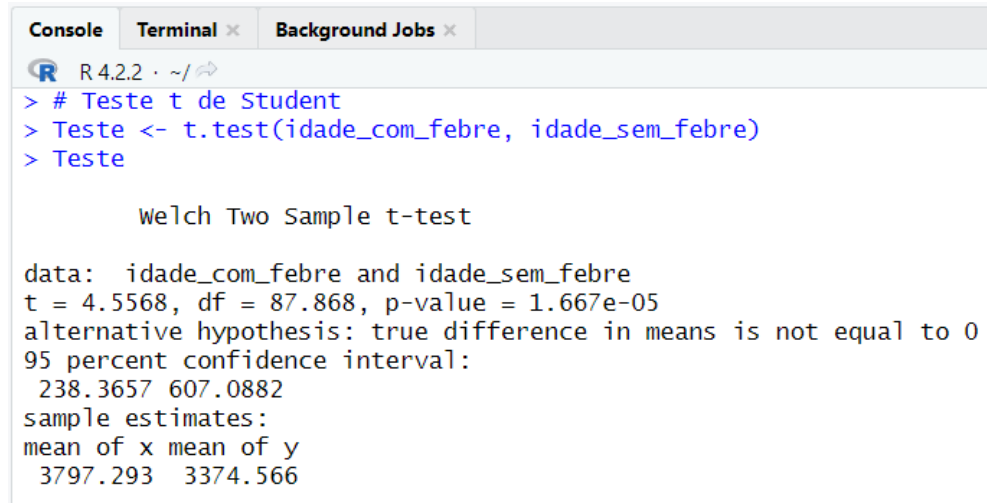
Hipótese alternativa (Ha): Existe diferença na média de idade entre pacientes com febre e sem febre.

Variáveis:

- Idade
- Sinais e sintomas - Febre

Teste: Teste t de Student. O teste t de Student é a melhor opção para comparar as médias de idade entre pacientes com febre e sem febre nesta amostra. Isso se deve ao fato de que a idade é uma variável numérica, tornando o teste t adequado para avaliar as diferenças nas médias. Além disso, o teste t é apropriado para comparar as médias entre dois grupos independentes, como é o caso dos pacientes com febre e sem febre.

Análise: Com base no resultado do teste t de Student, podemos concluir que há uma diferença estatisticamente significativa na média de idade entre pacientes com febre e sem febre. O p-value obtido (1.667e-05) é menor que um nível de significância de 0.05, o que nos leva a rejeitar a hipótese nula de que não há diferença na média de idade. Essa associação pode auxiliar na identificação precoce de casos suspeitos, no diagnóstico e no tratamento adequado da síndrome respiratória, considerando a faixa etária como um fator relevante a ser considerado. Além disso, pode contribuir para a estratificação de risco e a adoção de medidas preventivas específicas para diferentes grupos etários.



```
R 4.2.2 · ~/
> # Teste t de Student
> Teste <- t.test(idade_com_febre, idade_sem_febre)
> Teste

Welch Two Sample t-test

data:  idade_com_febre and idade_sem_febre
t = 4.5568, df = 87.868, p-value = 1.667e-05
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 238.3657 607.0882
sample estimates:
mean of x mean of y
 3797.293  3374.566
```

Figura 54 – Resultado do Teste: Existe Diferença na média de idade entre pacientes com febre e sem febre? com uma amostra de 1000 pessoas

Arquivo: Testes de Hipotese.R.

- **Hipótese Nº2:** Existe uma associação entre a presença de febre (FEBRE) e a presença de tosse (TOSSE)?

Hipótese nula (H0): Não há associação entre a presença de febre (FEBRE) e a presença de tosse (TOSSE).

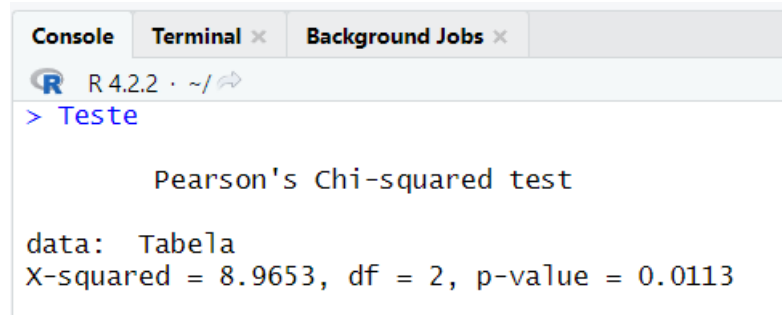
Hipótese alternativa (Ha): A presença de febre (FEBRE) está associada à presença de tosse (TOSSE).

Variáveis:

- Sinais e sintomas - Febre
- Sinais e sintomas - Tosse

Teste: Teste qui-quadrado. O teste qui-quadrado é ideal para analisar a relação entre febre e tosse, pois ambas são variáveis categóricas. Ele avalia se há associação estatisticamente significativa entre essas variáveis.

Análise: Com o p-value abaixo de 0.05 podemos concluir que há uma associação estatisticamente significativa entre a presença de febre e a presença de tosse. Portanto, rejeitamos a hipótese nula de que não há associação entre essas variáveis. A presença conjunta de febre e tosse pode indicar uma maior gravidade da doença ou estar relacionada a determinados subtipos de síndrome respiratória. Com base nessa associação identificada, os profissionais de saúde podem adotar estratégias de diagnóstico e tratamento mais eficazes, monitorando atentamente os pacientes com esses sintomas e tomando medidas preventivas adequadas para evitar complicações respiratórias graves.



```
Console Terminal x Background Jobs x
R 4.2.2 · ~/
> Teste

Pearson's Chi-squared test

data: Tabela
X-squared = 8.9653, df = 2, p-value = 0.0113
```

Figura 55 – Resultado do Teste: Existe uma associação entre a presença de febre e a presença de tosse? com uma amostra de 1000 pessoas

Arquivo: Testes de Hipotese.R

- **Hipótese Nº3:** A média de idade (NU_IDADE_N) varia de acordo com a raça/cor (CS_RACA)?

Hipótese nula (H₀): Não há diferença na média de idade entre as diferentes categorias de raça/cor.

Hipótese alternativa (H_a): Pelo menos uma das médias de idade difere das outras categorias de raça/cor.

Variáveis:

- Idade
- Raça/Cor

Teste: Teste ANOVA. O teste ANOVA é essencial para analisar se a média de idade (variável numérica contínua) varia de acordo com a raça/cor (variável categórica) e para determinar se há uma diferença significativa na média de idade entre os grupos raciais/correspondentes, proporcionando uma análise estatística adequada.

Análise: Com base no resultado do teste de ANOVA, com um p-value de 0.0436, podemos concluir que existe evidência estatística para rejeitar a hipótese nula de que não há diferença na média de idade entre as diferentes categorias de raça/cor. Isso significa que há indícios de que pelo menos uma das médias de idade difere das outras categorias. Saber que a média de idade varia de acordo com a raça/cor pode auxiliar na identificação de grupos populacionais que possam estar mais suscetíveis ou apresentar características diferentes relacionadas à SRAG. Isso pode direcionar estratégias de prevenção, diagnóstico e tratamento específicas para cada grupo, visando uma abordagem mais eficaz e personalizada.

```

R 4.2.2 ~/>
> # Visualizando Resultado
> summary(Teste)
      Df    Sum Sq Mean Sq F value Pr(>F)
amostra_raca  5    3604680   720936   2.295  0.0436 *
Residuals   981   308169921   314139

```

Figura 56 – Resultado do Teste: A média de idade varia de acordo com a raça/cor? com uma amostra de 1000 pessoas

Arquivo: Testes de Hipotese.R

- **Hipótese Nº4:** A proporção de pacientes com tosse (TOSSE) é diferente entre os grupos de gestantes (CS_GESTANT)?

Hipótese nula (H_0): A proporção de pacientes com tosse é igual entre os grupos de gestantes.

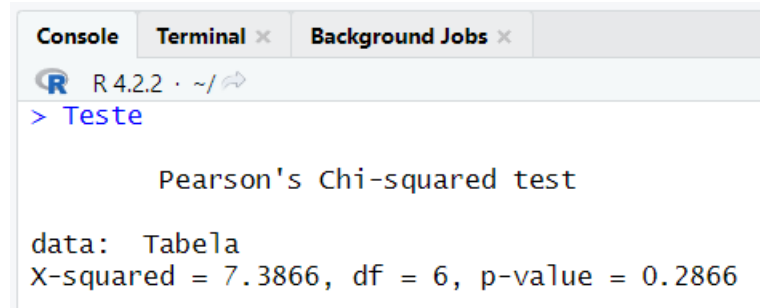
Hipótese alternativa (H_a): A proporção de pacientes com tosse é diferente entre os grupos de gestantes.

Variáveis:

- Sinais e sintomas - Tosse
- Gestante

Teste: Teste qui-quadrado. O teste do qui-quadrado é ideal para essa hipótese por lidar com duas variáveis categóricas (tosse e gestante), verificar associações independentes, comparar proporções e testar hipóteses específicas. No caso da proporção de pacientes com tosse entre grupos de gestantes, o qui-quadrado permite determinar se há diferenças significativas. Portanto, é a melhor opção para avaliar se a proporção de pacientes com tosse varia entre os grupos de gestantes

Análise: Com base no teste do qui-quadrado realizado, com um p-value de 0.2866, não temos evidências suficientes para rejeitar a hipótese nula. Isso significa que não há diferença significativa na proporção de pacientes com tosse entre os grupos de gestantes. Essa informação pode ser útil para um profissional de saúde ao avaliar a SRAG, pois indica que a presença de tosse pode não ser influenciada pelo fato de uma paciente estar grávida ou não.



```
Console Terminal x Background Jobs x
R 4.2.2 · ~/
> Teste

Pearson's Chi-squared test

data: Tabela
X-squared = 7.3866, df = 6, p-value = 0.2866
```

Figura 57 – Resultado do Teste: A proporção de pacientes com tosse é diferente entre os grupos de gestantes? com uma amostra de 1000 pessoas

Arquivo: Testes de Hipotese.R

- **Hipótese N^o5:** Existe uma associação entre a presença de dor de garganta (GARGANTA) e a presença de febre (FEBRE)?

Hipótese nula (H₀): Não há associação entre a presença de dor de garganta (GARGANTA) e a presença de febre (FEBRE).

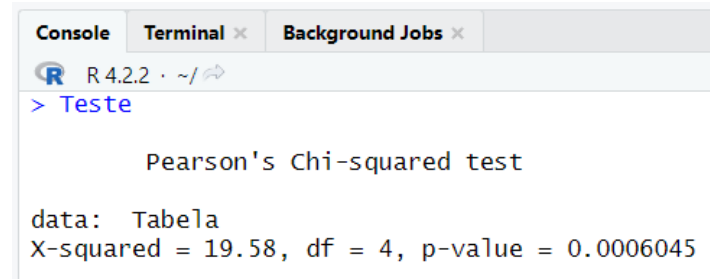
Hipótese alternativa (H_a): A presença de dor de garganta (GARGANTA) está associada à presença de febre (FEBRE).

Variáveis:

- Sinais e sintomas - Dor de garganta
- Sinais e sintomas - Febre

Teste: Teste qui-quadrado. O teste do qui-quadrado é essencial para essa análise porque permite verificar se há uma associação significativa entre duas variáveis categóricas, como a presença de dor de garganta e a presença de febre. Esse teste estatístico avalia se as diferenças observadas entre as frequências esperadas e as frequências observadas são estatisticamente significantes.

Análise: Com base no resultado do qui-quadrado, temos o valor do p-value abaixo de 0.05, então podemos concluir que há uma associação estatisticamente significativa entre a presença de dor de garganta e a presença de febre. Portanto, rejeitamos a hipótese nula de que não há associação entre essas variáveis. A presença conjunta de dor de garganta e febre pode ser um indicativo de infecções do trato respiratório superior, como faringite ou amigdalite, que são comuns na síndrome respiratória. Com base nessa associação identificada, os profissionais de saúde podem direcionar suas investigações clínicas e adotar medidas de tratamento apropriadas para aliviar os sintomas, controlar a infecção e evitar complicações respiratórias mais graves.



```
Console Terminal x Background Jobs x
R 4.2.2 · ~/
> Teste

Pearson's Chi-squared test

data:  Tabela
X-squared = 19.58, df = 4, p-value = 0.0006045
```

Figura 58 – Resultado do Teste: Existe uma associação entre a presença de dor de garganta e a presença de febre?

Arquivo: Testes de Hipotese.R

- **Hipótese N^o6:** Existe diferença na proporção de pessoas que apresentaram dor de garganta (GARGANTA) entre pessoas de raça/cor branca (CS_RACA) e pessoas de raça/cor preta (CS_RACA)?

Hipótese nula (H₀): Não há diferença na proporção de pessoas que apresentaram dor de garganta entre pessoas de raça/cor branca e pessoas de raça/cor preta.

Hipótese alternativa (H_a): Existe diferença na proporção de pessoas que apresentaram dor de garganta entre pessoas de raça/cor branca e pessoas de raça/cor preta.

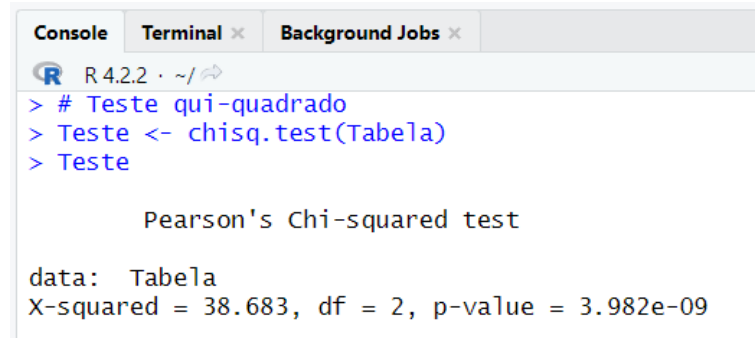
Variáveis:

- Raça/Cor
- Sinais e sintomas - Dor de Garganta

Teste: Teste qui-quadrado. O teste qui-quadrado é utilizado para avaliar a associação entre duas variáveis categóricas. Neste caso, queremos verificar se existe uma associação entre a variável Raça/Cor e a proporção de pessoas que apresentaram dor de garganta.

Análise: Com base no resultado do teste qui-quadrado, em que o p-value é extremamente baixo, rejeitamos a hipótese nula de que não há diferença na proporção de pessoas que apresentaram dor de garganta entre pessoas de raça/cor branca e pessoas de raça/cor preta. Portanto podemos concluir que há uma diferença significativa na proporção de pessoas que apresentaram dor de garganta entre esses dois grupos raciais/cor. Esse resultado pode ser relevante para profissionais de saúde que lidam com a SRAG, pois indica que a raça/cor pode estar associada a uma maior ou menor incidência de dor de garganta nesses casos. Isso pode ajudar na identificação de possíveis fatores de risco ou características específicas de determinados grupos raciais/cor que possam influenciar na manifestação da doença ou nos sintomas apresentados.

- **Hipótese N^o7:** Existe uma relação entre o sexo (CS_SEXO) da pessoa e a vacinação contra gripe (VACINA)?



```
Console Terminal x Background Jobs x
R 4.2.2 · ~/
> # Teste qui-quadrado
> Teste <- chisq.test(Tabela)
> Teste

Pearson's Chi-squared test

data: Tabela
X-squared = 38.683, df = 2, p-value = 3.982e-09
```

Figura 59 – Resultado do Teste: Existe diferença na proporção de pessoas que apresentaram dor de garganta entre pessoas de raça/cor branca e pessoas de raça/cor preta? com uma amostra de 1000 pessoas

Arquivo: Testes de Hipotese.R

Hipótese nula (H_0): Não há relação entre o sexo da pessoa e a vacinação contra gripe.

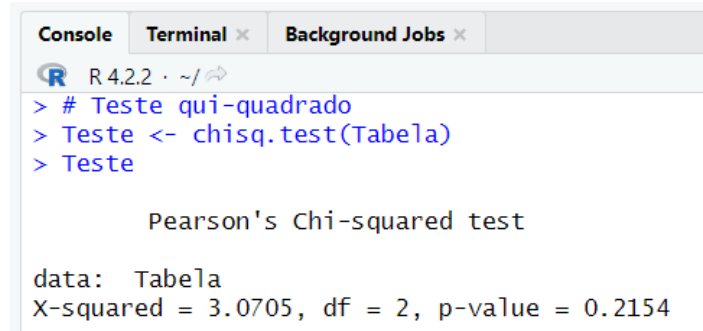
Hipótese alternativa (H_a): Existe uma relação entre o sexo da pessoa e a vacinação contra gripe.

Variáveis:

- Sexo
- Recebeu vacina contra gripe

Teste: Teste qui-quadrado. O teste qui-quadrado é adequado para essa hipótese porque estamos comparando duas variáveis categóricas, o sexo da pessoa e a vacinação contra gripe. O teste qui-quadrado de independência é comumente utilizado para verificar se existe uma associação entre duas variáveis categóricas em uma amostra. Ele permite avaliar se a distribuição das frequências observadas é significativamente diferente da distribuição esperada caso as duas variáveis sejam independentes.

Análise: Com base no resultado do teste qui-quadrado, com um p-value de 0.2154, não temos evidências suficientes para rejeitar a hipótese nula. Isso significa que não há uma relação significativa entre o sexo da pessoa e a vacinação contra gripe. Ao saber que não há uma associação entre o sexo da pessoa e a vacinação contra gripe, o profissional pode direcionar seus esforços e conscientização da vacina para ambos os sexos, sem necessidade de priorização com base no sexo.



```
Console Terminal x Background Jobs x
R 4.2.2 · ~/
> # Teste qui-quadrado
> Teste <- chisq.test(Tabela)
> Teste

Pearson's Chi-squared test

data:  Tabela
X-squared = 3.0705, df = 2, p-value = 0.2154
```

Figura 60 – Resultado do Teste: Existe uma relação entre o sexo da pessoa e a vacinação contra gripe? com uma amostra de 1000 pessoas

Arquivo: Testes de Hipotese.R

- **Hipótese N°8:** Existe uma relação entre a vacinação contra gripe (VACINA) e a presença de tosse (TOSSE)?

Hipótese nula (H0): Não há relação entre a vacinação contra gripe (VACINA) e a presença de tosse (TOSSE)

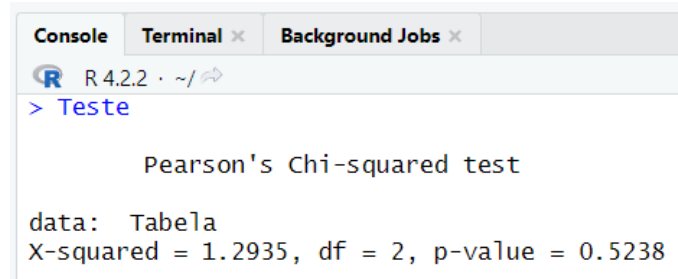
Hipótese alternativa (Ha): A vacinação contra gripe (VACINA) está relacionada à presença de tosse (TOSSE)

Variáveis:

- Recebeu vacina contra gripe
- Sinais e sintomas - Tosse

Teste: Teste qui-quadrado. O teste qui-quadrado é uma escolha adequada para essa hipótese, pois ela envolve a análise da relação entre duas variáveis categóricas: a vacinação contra a gripe (VACINA) e a presença de tosse (TOSSE). O teste qui-quadrado permite avaliar se há uma associação estatisticamente significativa entre essas duas variáveis.

Análise: Com o p-value acima de 0.05, não temos evidências estatisticamente significativas para rejeitar a hipótese nula (H0). Portanto, não há relação entre a vacinação contra gripe e a presença de tosse. Essa análise pode ajudar os profissionais de saúde na identificação de fatores de risco adicionais que podem contribuir para a gravidade dos sintomas respiratórios. Através desses dados, os profissionais de saúde podem adotar medidas preventivas mais eficazes e tomar decisões clínicas mais fundamentadas e também permite a implementação de estratégias mais assertivas para mitigar os riscos associados à SRAG, contribuindo assim para o cuidado e bem-estar dos pacientes.



```

Console Terminal x Background Jobs x
R 4.2.2 · ~/
> Teste

Pearson's Chi-squared test

data: Tabela
X-squared = 1.2935, df = 2, p-value = 0.5238

```

Figura 61 – Resultado do Teste: Existe uma relação entre a vacinação contra gripe e a presença de tosse? com uma amostra de 1000 pessoas

Arquivo: Testes de Hipotese.R

- **Hipótese N^o9:** Há relação entre o PIB per capita da UF(unidade da federação) e o número de evoluções nas internações naquele estado?

Hipótese nula (H₀): Não há relação significativa.

Hipótese alternativa (H_a): Há relação significativa.

Variáveis:

- SG_UF_NOT
- EVOLUCAO
- UF_PIBPC
- UF_CATEGORIA

Teste: Neste teste, testamos a hipótese de a evolução dos quadros dos pacientes estar ligada às condições financeiras da unidade da federação em que se encontra. Isso pode implicar em diversas outras análises, como quais fatores financeiros podem ajudar pacientes a melhorar seus quadros para cura, e outros que podem agravar seu quadro levando à óbito. Porém, em primeiro lugar, é necessário que se faça uma relação entre essas evoluções e a condição financeira. Para isso, usaremos a base de dados "INFLUD13.csv", além dos dados das UF(unidades da federação) retirados diretamente do site do IBGE, como os códigos para cada UF e o PIB per capita delas.

Primeiramente, obtemos os dados do IBGE para as unidades da federação. Em seguida, transformamos esses dados em um dataframe utilizável, com a coluna contendo o código da unidade da federação igual ao dataframe da base de dados "INFLUD13.csv".

Para realizar um teste qui-quadrado para variáveis categóricas, é necessário transformar esses valores de PIB per capita em uma variável categórica. Por isso, é criada a variável "UF_CATEGORIA", realizando uma operação com os quartis da variável

SG_UF_NOT	uf_nome	uf_pibpc
12	Acre	18420.26
27	Alagoas	18657.69
16	Amapá	21431.53
13	Amazonas	27572.96
29	Bahia	20449.29
23	Ceará	18168.35
53	Distrito Federal	87016.16
32	Espírito Santo	34065.98

Figura 62 – Dataframe PIB per Capita - IBGE

Arquivo: Testes de Hipótese.R

35	São Paulo	51364.73	3
28	Sergipe	19583.07	1
17	Tocantins	27448.43	2

Figura 63 – Categorização do PIB

Arquivo: Testes de Hipótese.R

"UF_PIBPC", e gerando novas colunas para os valores, que são categorizados de 1 à 3, sendo 1 considerado um PIB per capita baixo, e 3 um PIB per capita alto.

Então, esses valores são fundidos com a base de dados "INFLUD13.csv" usando o código da UF "SG_UF_NOT" como parâmetro, gerando uma amostra nova contendo dados financeiros de cada UF. Também são filtrados apenas registros contendo dados sobre a evolução "EVOLUCAO" dos pacientes, já que dados nulos poderiam interferir nos testes.

Então, por fim, é realizado um teste qui-quadrado relacionando as variáveis categóricas "EVOLUCAO" e "UF_CATEGORIA" dessa nova amostra de dados, além de gerado um gráfico de barras utilizando a função ggplot, no qual é representada a distribuição da evolução dos casos de síndrome respiratória aguda.

Análise: Com base nos dados obtidos no teste qui-quadrado, é possível inferir a relação entre as duas variáveis: "EVOLUCAO" e "UF_CATEGORIA". Apesar não provar uma relação direta, existe uma correlação entre esses dados, que podem agora ser estudados mais profundamente.

```
> resultado_teste <- chisq.test(merge_evolucao$EVOLUCAO, merge_evolucao$categoria)
> print(resultado_teste)

Pearson's Chi-squared test

data: merge_evolucao$EVOLUCAO and merge_evolucao$categoria
X-squared = 80.263, df = 4, p-value < 2.2e-16
```

Figura 64 – Relação entre PIB per capita e evolução - Resultado

Arquivo: Testes de Hipótese.R

Esses resultados , ao sugerirem a relação entre esses dois fatores, ilustram uma possível influência de fatores econômicos na evolução dos quadros dos pacientes dessa amostra. Possíveis causas como a precariedade dos sistemas hospitalares em regiões mais debilitadas financeiramente, o acesso à educação de qualidade e a conscientização da população sobre medidas preventivas, além da qualidade das moradias e instalações dos pacientes podem ser consideradas.

Porém, é importante considerar que diversos outros fatores possuem relações semelhantes com essa evolução. Logo, é necessário realizar estudos mais complexos para averiguar melhor essa relação e poder tirar conclusões mais assertivas.

Portanto, apesar da correlação provada, é extremamente necessário que hajam mais estudos nesse sentido para que medidas reais possam ser tomadas.

3 Conclusão

Neste relatório, realizamos uma análise exploratória dos dados da SRAG, utilizando bases de dados fornecidas pelo governo do Brasil. Através das análises básicas de estatística, pudemos obter insights valiosos sobre a incidência de síndromes respiratórias e sua evolução para SRAG.

A importância dos dados obtidos não pode ser subestimada. Essas informações fornecem uma visão detalhada da ocorrência e evolução de doenças respiratórias graves no país, permitindo que pesquisadores, profissionais de saúde e autoridades tomem decisões embasadas em evidências.

Durante a análise exploratória, utilizamos técnicas estatísticas para examinar diferentes aspectos dos dados, como frequência, distribuição e evolução das doenças, fatores demográficos dos afetados e outras variáveis relevantes. Isso nos permitiu identificar tendências, padrões e possíveis relações entre os dados.

No entanto, é importante ressaltar que as análises exploratórias de dados são apenas o primeiro passo. Podemos usar testes inferenciais hipotéticos para realizar inferências estatísticas e responder a questões mais específicas. Esses testes nos ajudam a avaliar a significância estatística de certas relações ou diferenças entre grupos, possibilitando a validação de hipóteses e a obtenção de conclusões mais robustas.

Ao conduzir testes inferenciais, podemos explorar hipóteses, como a relação entre fatores de risco específicos e a evolução para SRAG, ou a incidência de determinadas condições no agravamento e obtenção das síndromes. Essas análises fornecem uma base sólida para tomadas de decisão, auxiliando na formulação de políticas de saúde e na implementação de medidas preventivas e de tratamento mais eficazes.

Em suma, a análise exploratória dos dados da SRAG revelou informações valiosas sobre a ocorrência e evolução das síndromes respiratórias graves no Brasil. Através de análises básicas de estatística e testes inferenciais hipotéticos, podemos obter insights mais profundos como a relação entre a idade dos pacientes e a incidência do sintoma febre, a presença dos sintomas de febre e tosse simultaneamente, a relação entre a raça e idade dos infectados, e a presença de tosse entre gestantes. Com uma base de dados robusta e um grande número de dados, foi possível inferir diversas relações, além de presenciar eventos raros apontados por essas hipóteses. Essa abordagem integrada é crucial para compreendermos melhor a SRAG, suas consequências bem como para o desenvolvimento de estratégias eficazes de saúde pública e assistência médica.