

**NATÁLIA APARECIDA GADELHA (SP3070727)**  
**VINICIUS DE PINHO SANTOS (SP3072851)**

**Projeto - Análise Exploratória De Dados De Estatística E Probabilidade**  
**Script dos códigos utilizados para gerar as análises**

Bases de dados escolhida para analisar: **Notificações de Síndrome Gripal - 2022**

<https://opendatasus.saude.gov.br/dataset/notificacoes-de-sindrome-gripal-leve-2022>

**Lote de Dados 10 SP - 19/11: Para adicionar ao R Studio Cloud**

<https://s3.sa-east-1.amazonaws.com/ckan.saude.gov.br/SGL/2022/uf=SP/lote=10/part-00000-cf98759e-05e9-44fe-9450-8eba122f0160.c000.csv>

```
# Criar uma amostra aleatória de 5.000 linhas sem repetição
```

```
amostra <-
```

```
part_00000_cf98759e_05e9_44fe_9450_8eba122f0160_c000[sample(nrow(part_00000_cf98759e_05e9_44fe_9450_8eba122f0160_c000), 5000, replace = FALSE), ]
```

```
# Visualize os primeiros registros da amostra
```

```
head(amostra)
```

```
View(amostra)
```

**Cálculo porcentagem pacientes assintomáticos e sintomáticos:**

```
# Calcular a porcentagem de pacientes Assintomáticos e pacientes != Assintomáticos
```

```
tabela_sintomas <- table(amostra$sintomas)
```

```
porcentagem_assintomaticos <- (tabela_sintomas["Assintomático"] / sum(tabela_sintomas)) * 100
```

```
porcentagem_nao_assintomaticos <- (sum(tabela_sintomas!="Assintomático") / sum(tabela_sintomas)) * 100
```

```
porcentagem_outros_sintomas <- (sum(tabela_sintomas[-which(names(tabela_sintomas) == "Assintomático")]) / sum(tabela_sintomas)) * 100
```

```
# Imprimir os resultados
```

```
cat("Porcentagem de pacientes Assintomáticos:", porcentagem_assintomaticos, "%\n")
```

```
#Porcentagem de pacientes Assintomáticos: 17.36 %
```

```
cat("Porcentagem de pacientes != Assintomáticos:", porcentagem_outros_sintomas, "%\n")
```

```
#Porcentagem de pacientes != Assintomáticos: 82.64 %
```

### **Infectados assintomáticos por idade:**

```
# Filtrar dados para incluir apenas os assintomáticos
dados_assintomaticos
dados_assintomaticos <- subset(amostra, sintomas == "Assintomático")

# Criar um histograma da idade para os assintomáticos
ggplot(dados_assintomaticos, aes(x = idade, fill = sintomas)) +
  geom_histogram(binwidth = 3, position = "dodge", color = "black", alpha = 0.7, bins = 30) +
  labs(title = "Idade dos Pacientes Assintomáticos",
       x = "Idade",
       y = "Quantidade") +
  scale_fill_manual(values = c("Assintomático" = "skyblue")) +
  theme_minimal() +
  scale_x_continuous(breaks = seq(0, 90, by = 5)) # Adiciona etiquetas de eixo x a cada 5
anos
```

### **Infectados assintomáticos por Laboratório considerando a primeira dose da vacina:**

```
# Histograma da idade para os assintomáticos, incluindo laboratório
ggplot(dados_assintomaticos, aes(x = idade, fill = factor(codigoLaboratorioPrimeiraDose,
levels = c("PFIZER", "SINOVAC/BUTANTAN", "ASTRAZENECA/FIOCRUZ", "Pendente
Identificação")))) +
  geom_histogram(binwidth = 5, position = "dodge", color = "black", alpha = 0.7, na.rm =
TRUE) +
  labs(title = "Quantidade dos Pacientes Assintomáticos por Laboratório",
       x = "Idade",
       y = "Contagem",
       subtitle = "Considerando a primeira dose da vacina") +
  theme_minimal() +
  scale_fill_manual(values = c("PFIZER" = "skyblue", "SINOVAC/BUTANTAN" = "green",
"ASTRAZENECA/FIOCRUZ" = "yellow", "Pendente Identificação" = "red", na.value =
"black"),
                 name = "Laboratório",
                 na.value = "red") +
  guides(fill = guide_legend(title = "Laboratório")) +
  scale_x_continuous(breaks = seq(0, 90, by = 10))
```

### **Infectados assintomáticos por Laboratório considerando a segunda dose da vacina:**

```
# Criar um histograma da idade para os assintomáticos, incluindo laboratório e visualizando NAs
ggplot(dados_assintomaticos, aes(x = idade, fill = factor(codigoLaboratorioSegundaDose,
levels = c("PFIZER", "SINOVAC/BUTANTAN", "ASTRAZENECA/FIOCRUZ", "Pendente
Identificação")))) +
  geom_histogram(binwidth = 5, position = "dodge", color = "black", alpha = 0.7, na.rm =
TRUE) +
  labs(title = "Quantidade dos Pacientes Assintomáticos por Laboratório",
    x = "Idade",
    y = "Contagem",
    subtitle = "Considerando a segunda dose da vacina") +
  theme_minimal() +
  scale_fill_manual(values = c("PFIZER" = "skyblue", "SINOVAC/BUTANTAN" = "green",
"ASTRAZENECA/FIOCRUZ" = "yellow", "Pendente Identificação" = "red", na.value =
"black"),
    name = "Laboratório",
    na.value = "red") +
  guides(fill = guide_legend(title = "Laboratório")) +
  scale_x_continuous(breaks = seq(0, 90, by = 10))
```

### **Infectados sintomáticos com base no número de doses da vacina: 0 doses**

```
# Carregar a biblioteca ggplot2
library(ggplot2)

amostra <-
part_00000_cf98759e_05e9_44fe_9450_8eba122f0160_c000[sample(nrow(part_00000_cf9
8759e_05e9_44fe_9450_8eba122f0160_c000), 5000, replace = FALSE), ]
# Criar uma nova variável indicando se a pessoa é sintomática ou não
amostra$Sintomatico <- ifelse(amostra$sintomas != "Assintomático", "Sintomático",
"Assintomático")
# Criar uma nova variável indicando o número de doses que a pessoa tomou
amostra$numeroDoses <- ifelse(amostra$dataPrimeiraDose == "None", 0,
  ifelse(amostra$dataSegundaDose == "None", 1, 2))
# Criar o histograma incluindo apenas quem não tomou vacina
histograma <- ggplot(amostra[amostra$numeroDoses == 0, ], aes(x = numeroDoses, fill =
Sintomatico)) +
  geom_histogram(position = "dodge", color = "black", alpha = 0.7) +
  labs(title = "Relação entre Número de Doses e Sintomas",
    x = "Número de Doses",
    y = "Contagem") +
  scale_fill_manual(values = c("Assintomático" = "blue", "Sintomático" = "red")) +
  theme_minimal() +
  scale_x_continuous(breaks = seq(0, 90, by = 1)) # Adiciona etiquetas de eixo x a cada 5
anos

# Mostrar o histograma
print(histograma)
```

### **Infectados sintomáticos com base no número de doses da vacina: 1 dose**

```
# Criar o histograma incluindo apenas quem tomou 1 dose
histograma <- ggplot(amostra[amostra$numeroDoses == 1, ], aes(x = numeroDoses, fill =
Sintomatico)) +
  geom_histogram(position = "dodge", color = "black", alpha = 0.7) +
  labs(title = "Relação entre Número de Doses e Sintomas",
    x = "Número de Doses",
    y = "Contagem") +
  scale_fill_manual(values = c("Assintomático" = "blue", "Sintomático" = "red")) +
  theme_minimal() +
  scale_x_continuous(breaks = seq(0, 90, by = 1)) # Adiciona etiquetas de eixo x a cada 5
anos

# Mostrar o histograma
print(histograma)
```

### **Infectados sintomáticos com base no número de doses da vacina: 2 doses**

```
# Criar o histograma incluindo apenas quem tomou 2 doses
histograma <- ggplot(amostra[amostra$numeroDoses == 2, ], aes(x = numeroDoses, fill =
Sintomatico)) +
  geom_histogram(position = "dodge", color = "black", alpha = 0.7) +
  labs(title = "Relação entre Número de Doses e Sintomas",
    x = "Número de Doses",
    y = "Contagem") +
  scale_fill_manual(values = c("Assintomático" = "blue", "Sintomático" = "red")) +
  theme_minimal() +
  scale_x_continuous(breaks = seq(0, 90, by = 1)) # Adiciona etiquetas de eixo x a cada 5
anos

# Mostrar o histograma
print(histograma)
```

### **Porcentagem da população que tomou 0, 1 e 2 doses da vacina:**

```
# Criar uma nova variável indicando o número de doses que a pessoa tomou
amostra$numeroDoses <- ifelse(amostra$dataPrimeiraDose == "None", 0,
  ifelse(amostra$dataSegundaDose == "None", 1, 2))

# Calcular a porcentagem da população que tomou 0, 1 e 2 doses da vacina
porcentagem_0_doses <- (sum(amostra$numeroDoses == 0) / nrow(amostra)) * 100
porcentagem_1_dose <- (sum(amostra$numeroDoses == 1) / nrow(amostra)) * 100
porcentagem_2_doses <- (sum(amostra$numeroDoses == 2) / nrow(amostra)) * 100

# Visualizar as porcentagens
```

```

cat("Porcentagem da população que tomou 0 doses da vacina:",
round(porcentagem_0_doses, 2), "%\n")
#Porcentagem da população que tomou 0 doses da vacina: 38.92 %
cat("Porcentagem da população que tomou 1 dose da vacina:",
round(porcentagem_1_dose, 2), "%\n")
#Porcentagem da população que tomou 1 dose da vacina: 16.46 %
cat("Porcentagem da população que tomou 2 doses da vacina:",
round(porcentagem_2_doses, 2), "%\n")
#Porcentagem da população que tomou 2 doses da vacina: 44.62 %

# Calcular a porcentagem da população sintomática que tomou 0, 1 e 2 doses da vacina
porcentagem_sintomaticos_0_doses <- (sum(amostra$numeroDoses[amostra$sintomas !=
"Assintomático"] == 0) / sum(amostra$sintomas != "Assintomático")) * 100
porcentagem_sintomaticos_1_dose <- (sum(amostra$numeroDoses[amostra$sintomas !=
"Assintomático"] == 1) / sum(amostra$sintomas != "Assintomático")) * 100
porcentagem_sintomaticos_2_doses <- (sum(amostra$numeroDoses[amostra$sintomas !=
"Assintomático"] == 2) / sum(amostra$sintomas != "Assintomático")) * 100

# Calcular a porcentagem da população assintomática que tomou 0, 1 e 2 doses da vacina
porcentagem_assintomaticos_0_doses <- (sum(amostra$numeroDoses[amostra$sintomas
== "Assintomático"] == 0) / sum(amostra$sintomas == "Assintomático")) * 100
porcentagem_assintomaticos_1_dose <- (sum(amostra$numeroDoses[amostra$sintomas
== "Assintomático"] == 1) / sum(amostra$sintomas == "Assintomático")) * 100
porcentagem_assintomaticos_2_doses <- (sum(amostra$numeroDoses[amostra$sintomas
== "Assintomático"] == 2) / sum(amostra$sintomas == "Assintomático")) * 100

# Visualizar as porcentagens
cat("Porcentagem da população sintomática que tomou 0 doses da vacina:",
round(porcentagem_sintomaticos_0_doses, 2), "%\n")
#Porcentagem da população sintomática que tomou 0 doses da vacina: 39.11 %
cat("Porcentagem da população sintomática que tomou 1 dose da vacina:",
round(porcentagem_sintomaticos_1_dose, 2), "%\n")
#Porcentagem da população sintomática que tomou 1 dose da vacina: 16.67 %
cat("Porcentagem da população sintomática que tomou 2 doses da vacina:",
round(porcentagem_sintomaticos_2_doses, 2), "%\n")
#Porcentagem da população sintomática que tomou 2 doses da vacina: 44.22 %

cat("Porcentagem da população assintomática que tomou 0 doses da vacina:",
round(porcentagem_assintomaticos_0_doses, 2), "%\n")
#Porcentagem da população assintomática que tomou 0 doses da vacina: 38.02 %
cat("Porcentagem da população assintomática que tomou 1 dose da vacina:",
round(porcentagem_assintomaticos_1_dose, 2), "%\n")
#Porcentagem da população assintomática que tomou 1 dose da vacina: 15.44 %
cat("Porcentagem da população assintomática que tomou 2 doses da vacina:",
round(porcentagem_assintomaticos_2_doses, 2), "%\n")
#Porcentagem da população assintomática que tomou 2 doses da vacina: 15.44 %

```

**Gráfico da relação entre número de doses e sintomas (em percentual):**

```

library(ggplot2)

# Calcula as porcentagens para cada combinação de doses e sintomas
porcentagens <- data.frame(
  numeroDoses = rep(c(0, 1, 2), each = 2),
  Sintomatico = rep(c("Assintomatico", "Sintomatico"), times = 3),
  porcentagem = c(
    porcentagem_assintomaticos_0_doses,
    porcentagem_sintomaticos_0_doses,
    porcentagem_assintomaticos_1_dose,
    porcentagem_sintomaticos_1_dose,
    porcentagem_assintomaticos_2_doses,
    porcentagem_sintomaticos_2_doses
  )
)

# Cria o histograma
histograma <- ggplot(porcentagens, aes(x = factor(numeroDoses), y = porcentagem, fill =
Sintomatico)) +
  geom_bar(stat = "identity", position = "dodge", color = "black", alpha = 0.7) +
  labs(title = "Relação entre Número de Doses e Sintomas",
    x = "Número de Doses",
    y = "Porcentagem") +
  scale_fill_manual(values = c("Assintomatico" = "blue", "Sintomatico" = "red")) +
  theme_minimal() +
  scale_x_discrete(labels = c("0" = "0 doses", "1" = "1 dose", "2" = "2 doses"))

# Exibe o histograma
print(histograma)

```