

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

ARTHUR BARBOSA DE OLIVEIRA SP3046974
GABRIEL NOVATO NASCIMENTO SP3052419
RENAN GOMES SP3081796

Projeto - Análise exploratória de dados

São Paulo - SP - Brasil
2023

Comparação de suicídios entre os anos de 1996 e 2021

Para a primeira parte do trabalho, iremos analisar os suicídios ao decorrer dos anos. Utilizamos os seguintes códigos para ler os arquivos e fazer o levantamento dos dados para calcular o total de suicídios nos anos de 1996, ano que começa a ser registrado o tipo de morte violenta ou não natural (contém o filtro de suicídios), e 2021, último ano com informações registradas de forma completa:

```
# Carregar os dados de 1996
dados_mortalidade_1996 <-
read.csv2("C:\\Users\\renan\\OneDrive\\Documentos\\estatistica\\Mortalidade_Geral_
1996.csv", sep = ";")

# Carregar os dados de 2021
dados_mortalidade_2021 <-
read.csv2("C:\\Users\\renan\\OneDrive\\Documentos\\estatistica\\Mortalidade_Geral_
2021.csv", sep = ";")

# Função para calcular o total de suicídios em um data frame
calcularTotalSuicidios <- function(data) {
  if ("CIRCOBITO" %in% colnames(data)) {
    total_suicidios <- sum(data$CIRCOBITO == 2, na.rm = TRUE)
    return(total_suicidios)
  } else {
    return(NA) # Retorna NA se a coluna CIRCOBITO não estiver presente
  }
}

# Calcular o total de suicídios em cada tabela
total_suicidios_1996 <- calcularTotalSuicidios(dados_mortalidade_1996)
total_suicidios_2021 <- calcularTotalSuicidios(dados_mortalidade_2021)

# Imprimir os resultados
cat("Total de suicídios em 1996:", total_suicidios_1996, "\n")
```

```
cat("Total de suicídios em 2021:", total_suicidios_2021, "\n")
```

Total de suicídios em 1996 = 6836

Total de suicídios em 2021 = 14724

Feito o levantamento total de suicídios nos dois anos, foi gerado um gráfico comparativo para mostrar a diferença entre os dois anos:

```
# Carregar a biblioteca ggplot2
```

```
library(ggplot2)
```

```
# Dados
```

```
anos <- c("1996", "2021")
```

```
total_suicidios <- c(total_suicidios_1996, total_suicidios_2021)
```

```
# Criar um data frame com os dados
```

```
dados_comparativos <- data.frame(Ano = anos, TotalSuicidios = total_suicidios)
```

```
# Criar o gráfico de barras comparativo
```

```
ggplot(dados_comparativos, aes(x = Ano, y = TotalSuicidios, fill = Ano)) +
```

```
  geom_bar(stat = "identity", position = "dodge") +
```

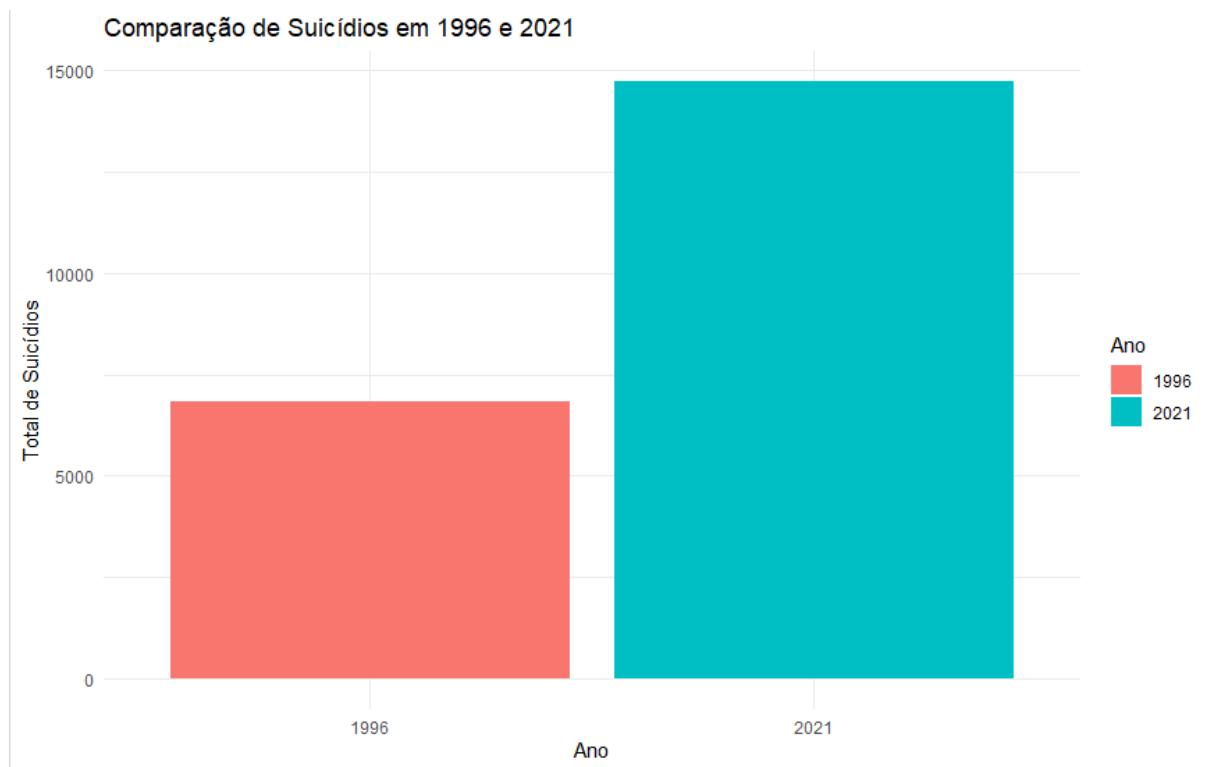
```
  labs(title = "Comparação de Suicídios em 1996 e 2021",
```

```
       x = "Ano",
```

```
       y = "Total de Suicídios") +
```

```
  theme_minimal()
```

Gráfico comparativo entre os anos de 1996 e 2021



Total de suicídios em 1996: 6836

Total de suicídios em 2021: 14724

Comparação entre mortes violentas e suicídio entre os anos de 1996 e 2021

Após ser feita a primeira comparação, comparamos o total de mortes não naturais com suicídios, seguem os códigos do R:

Função para calcular o total de suicídios em um data frame

```
calcularTotalSuicidios <- function(data) {
  if ("CIRCOBITO" %in% colnames(data)) {
    total_suicidios <- sum(data$CIRCOBITO == 2, na.rm = TRUE)
    return(total_suicidios)
  } else {
    return(NA) # Retorna NA se a coluna CIRCOBITO não estiver presente
  }
}
```

```

# Carregar os dados de 1996
dados_mortalidade_1996 <-
read.csv2("C:\\Users\\renan\\OneDrive\\Documentos\\estatistica\\Mortalidade_Geral_
1996.csv", sep = ";")

# Carregar os dados de 2021
dados_mortalidade_2021 <-
read.csv2("C:\\Users\\renan\\OneDrive\\Documentos\\estatistica\\Mortalidade_Geral_
2021.csv", sep = ";")

# Calcular o total de suicídios em cada tabela
total_suicidios_1996 <- calcularTotalSuicidios(dados_mortalidade_1996)
total_suicidios_2021 <- calcularTotalSuicidios(dados_mortalidade_2021)

# Imprimir os resultados
cat("Total de suicídios em 1996:", total_suicidios_1996, "\n")
cat("Total de suicídios em 2021:", total_suicidios_2021, "\n")

# Filtrar apenas mortes violentas (CIRCOBITO diferente de 2 e não vazias)
mortes_violentas_1996 <-
dados_mortalidade_1996[!is.na(dados_mortalidade_1996$CIRCOBITO) &
dados_mortalidade_1996$CIRCOBITO != 2, ]
mortes_violentas_2021 <-
dados_mortalidade_2021[!is.na(dados_mortalidade_2021$CIRCOBITO) &
dados_mortalidade_2021$CIRCOBITO != 2, ]

# Criar um data frame com os dados
dados_grafico <- data.frame(
  Ano = rep(c("1996", "2021"), each = 2),
  Categoria = rep(c("mortes não naturais", "Suicídios"), times = 2),
  Total = c(
    nrow(mortes_violentas_1996), total_suicidios_1996,
    nrow(mortes_violentas_2021), total_suicidios_2021
  )
)

```

```
)
)
```

```
# Criar o gráfico de colunas sobrepostas
```

```
library(ggplot2)
```

```
grafico <- ggplot(dados_grafico, aes(x = Ano, y = Total, fill = Categoria)) +
```

```
  geom_bar(stat = "identity", position = "dodge") +
```

```
  labs(title = "Comparação entre não naturais e Suicídios (1996 vs. 2021)",
```

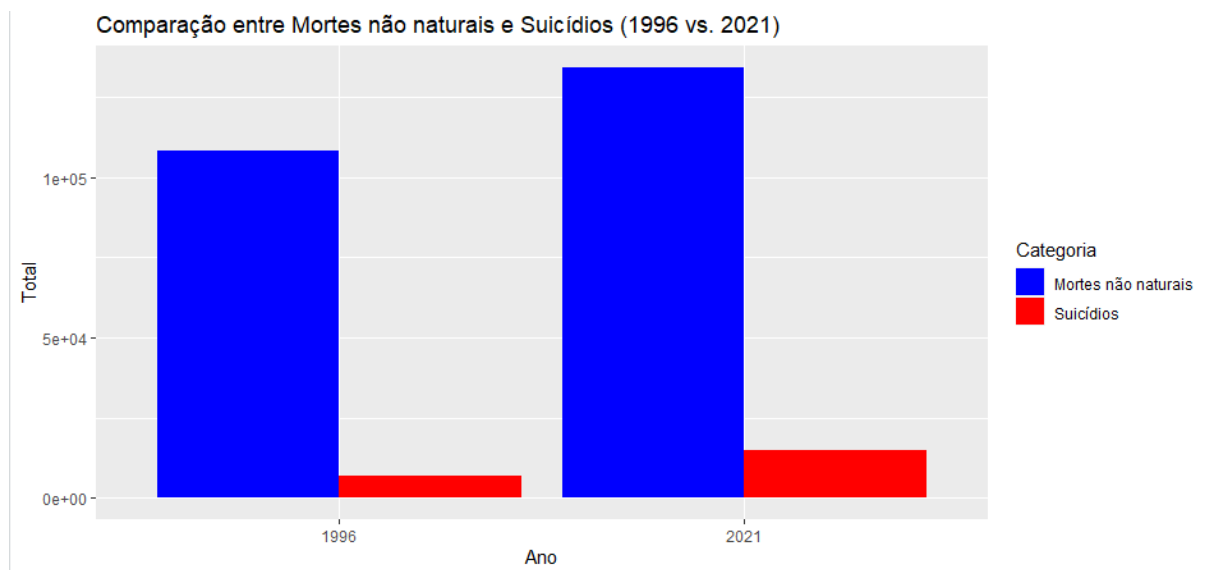
```
        x = "Ano",
```

```
        y = "Total") +
```

```
  scale_fill_manual(values = c("Mortes não naturais" = "blue", "Suicídios" = "red"))
```

```
# Exibir o gráfico
```

```
print(grafico)
```



Crescimento de mortes por suicídio ao longo dos anos:

Foi feito um comparativo do crescimento de mortes por suicídio ao decorrer dos anos, desde 1996 a 2021, sendo que de 1996 a 2020 os dados são comparados em um intervalo de dois anos(1996, 1998, 2000...):

Dados de 1996 a 2021

```
calcularTotalSuicidios <- function(data) {
  tryCatch({
    if ("CIRCOBITO" %in% colnames(data)) {
      total_suicidios <- sum(data$CIRCOBITO == 2, na.rm = TRUE)
      return(total_suicidios)
    } else {
      return(NA)
    }
  }, error = function(e) {
    cat("Erro:", conditionMessage(e), "\n")
    return(NA)
  })
}

# Carregar os dados de 1996
dados_mortalidade_1996 <-
read.csv2("C:\\Users\\renan\\OneDrive\\Documentos\\estatistica\\Mortalidade_Geral_
1996.csv", sep = ";")

# Carregar os dados de 1998
dados_mortalidade_1998 <-
read.csv2("C:\\Users\\renan\\OneDrive\\Documentos\\estatistica\\Mortalidade_Geral_
1998.csv", sep = ";")

# Carregar os dados de 2000
```

```
dados_mortalidade_2000 <-  
read.csv2("C:\\Users\\renan\\OneDrive\\Documentos\\estatistica\\Mortalidade_Geral_  
2000.csv", sep = ";")
```

```
# Carregar os dados de 2002  
#dados_mortalidade_2002 <-  
read.csv2("C:\\Users\\renan\\OneDrive\\Documentos\\estatistica\\Mortalidade_Geral_  
2002.csv", sep = ";")
```

```
# Carregar os dados de 2004  
dados_mortalidade_2004 <-  
read.csv2("C:\\Users\\renan\\OneDrive\\Documentos\\estatistica\\Mortalidade_Geral_  
2004.csv", sep = ";")
```

```
# Carregar os dados de 2006  
dados_mortalidade_2006 <-  
read.csv2("C:\\Users\\renan\\OneDrive\\Documentos\\estatistica\\Mortalidade_Geral_  
2006.csv", sep = ";")
```

```
# Carregar os dados de 2008  
dados_mortalidade_2008 <-  
read.csv2("C:\\Users\\renan\\OneDrive\\Documentos\\estatistica\\Mortalidade_Geral_  
2008.csv", sep = ";")
```

```
# Carregar os dados de 2010  
dados_mortalidade_2010 <-  
read.csv2("C:\\Users\\renan\\OneDrive\\Documentos\\estatistica\\Mortalidade_Geral_  
2010.csv", sep = ";")
```

```
# Carregar os dados de 2012  
dados_mortalidade_2012 <-  
read.csv2("C:\\Users\\renan\\OneDrive\\Documentos\\estatistica\\Mortalidade_Geral_  
2012.csv", sep = ";")
```



```
# Carregar os dados de 2014
dados_mortalidade_2014 <-
read.csv2("C:\\Users\\renan\\OneDrive\\Documentos\\estatistica\\Mortalidade_Geral_
2014.csv", sep = ";")
```

```
# Carregar os dados de 2016
dados_mortalidade_2016 <-
read.csv2("C:\\Users\\renan\\OneDrive\\Documentos\\estatistica\\Mortalidade_Geral_
2016.csv", sep = ";")
```

```
# Carregar os dados de 2018
dados_mortalidade_2018 <-
read.csv2("C:\\Users\\renan\\OneDrive\\Documentos\\estatistica\\Mortalidade_Geral_
2018.csv", sep = ";")
```

```
# Carregar os dados de 2020
dados_mortalidade_2020 <-
read.csv2("C:\\Users\\renan\\OneDrive\\Documentos\\estatistica\\Mortalidade_Geral_
2020.csv", sep = ";")
```

```
# Carregar os dados de 2021
dados_mortalidade_2021 <-
read.csv2("C:\\Users\\renan\\OneDrive\\Documentos\\estatistica\\Mortalidade_Geral_
2021.csv", sep = ";")
```

```
# Calcular o total de suicídios em cada tabela
total_suicidios_1996 <- calcularTotalSuicidios(dados_mortalidade_1996)
total_suicidios_1998 <- calcularTotalSuicidios(dados_mortalidade_1998)
total_suicidios_2000 <- calcularTotalSuicidios(dados_mortalidade_2000)
#total_suicidios_2002 <- calcularTotalSuicidios(dados_mortalidade_2002)
total_suicidios_2004 <- calcularTotalSuicidios(dados_mortalidade_2004)
total_suicidios_2006 <- calcularTotalSuicidios(dados_mortalidade_2006)
total_suicidios_2008 <- calcularTotalSuicidios(dados_mortalidade_2008)
total_suicidios_2010 <- calcularTotalSuicidios(dados_mortalidade_2010)
```

```
total_suicidios_2012 <- calcularTotalSuicidios(dados_mortalidade_2012)
total_suicidios_2014 <- calcularTotalSuicidios(dados_mortalidade_2014)
total_suicidios_2016 <- calcularTotalSuicidios(dados_mortalidade_2016)
total_suicidios_2018 <- calcularTotalSuicidios(dados_mortalidade_2018)
total_suicidios_2020 <- calcularTotalSuicidios(dados_mortalidade_2020)
total_suicidios_2021 <- calcularTotalSuicidios(dados_mortalidade_2021)
```

```
# Imprimir os resultados
```

```
cat("Total de suicídios em 1996:", total_suicidios_1996, "\n")
cat("Total de suicídios em 1998:", total_suicidios_1998, "\n")
cat("Total de suicídios em 2000:", total_suicidios_2000, "\n")
#cat("Total de suicídios em 2002:", total_suicidios_2002, "\n")
cat("Total de suicídios em 2004:", total_suicidios_2004, "\n")
cat("Total de suicídios em 2006:", total_suicidios_2006, "\n")
cat("Total de suicídios em 2008:", total_suicidios_2008, "\n")
cat("Total de suicídios em 2010:", total_suicidios_2010, "\n")
cat("Total de suicídios em 2012:", total_suicidios_2012, "\n")
cat("Total de suicídios em 2014:", total_suicidios_2014, "\n")
cat("Total de suicídios em 2016:", total_suicidios_2016, "\n")
cat("Total de suicídios em 2018:", total_suicidios_2018, "\n")
cat("Total de suicídios em 2020:", total_suicidios_2020, "\n")
cat("Total de suicídios em 2021:", total_suicidios_2021, "\n")
```

```
#####
#####
```

```
#fazer o grafico
```

```
# Instalar e carregar a biblioteca ggplot2
```

```
if (!requireNamespace("ggplot2", quietly = TRUE)) {
  install.packages("ggplot2")
}
library(ggplot2)
```

```
# Criar um data frame com os anos e o total de suicídios
```

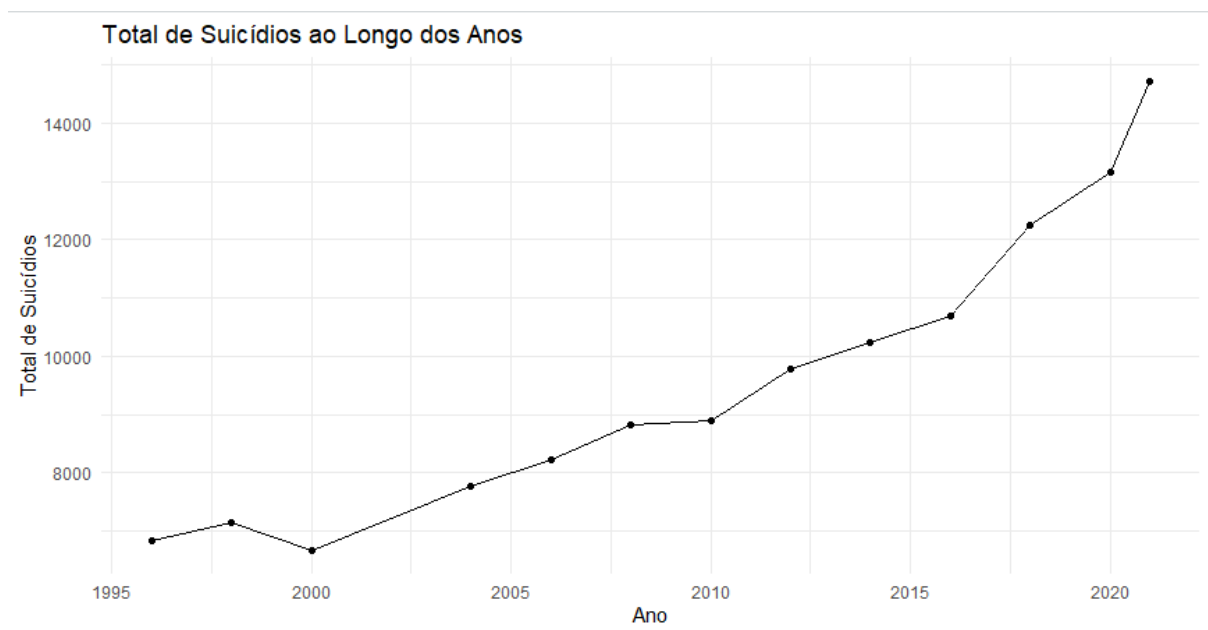
```
dados_grafico <- data.frame(
```

```
Ano = c(1996, 1998, 2000, 2004, 2006, 2008, 2010, 2012, 2014, 2016, 2018, 2020,
2021),
```

```
Total_Suicidios = c(
  total_suicidios_1996, total_suicidios_1998, total_suicidios_2000,
  total_suicidios_2004, total_suicidios_2006, total_suicidios_2008,
  total_suicidios_2010, total_suicidios_2012, total_suicidios_2014,
  total_suicidios_2016, total_suicidios_2018, total_suicidios_2020,
total_suicidios_2021
)
)
```

```
# Criar o gráfico de linhas
```

```
ggplot(dados_grafico, aes(x = Ano, y = Total_Suicidios)) +
  geom_line() +
  geom_point() +
  labs(title = "Total de Suicídios ao Longo dos Anos",
       x = "Ano",
       y = "Total de Suicídios") +
  theme_minimal()
```



Teste de hipótese, se houve uma aumento significativo entre o ano de 2020 e o ano de 2021:

Após verificar na tabela anterior um acentuação muito grande entre os anos de 2020 e 2021, fizemos um teste de hipóteses para avaliar se houve algum aumento significativo de suicídios entre os anos de 2020 e 2021:

Código R:

```
# Definir os totais de suicídios reais para 2020 e 2021
total_suicidios_2020_real <- total_suicidios_2020
total_suicidios_2021_real <- total_suicidios_2021

# Criar amostras com base nos totais reais
amostra_2020 <- rep(2020, total_suicidios_2020_real)
amostra_2021 <- rep(2021, total_suicidios_2021_real)

# Combinar as amostras em um único vetor
dados <- c(amostra_2020, amostra_2021)

# Criar um vetor indicando o ano correspondente para cada observação
anos <- rep(c("2020", "2021"), c(total_suicidios_2020_real,
total_suicidios_2021_real))

# Criar um data frame
dados_teste <- data.frame(Ano = as.factor(anos), Suicidios = as.factor(dados))

# Criar uma tabela de contingência
tabela_contingencia <- table(dados_teste$Ano, dados_teste$Suicidios)

# Realizar o teste qui-quadrado
resultado_teste_qui_quadrado <- chisq.test(tabela_contingencia)

# Exibir os resultados do teste qui-quadrado
```

```
cat("Resultado do Teste Qui-Quadrado:\n")
print(resultado_teste_qui_quadrado)
```

```
Pearson's Chi-squared test with Yates' continuity correction
```

```
data:  tabela_contingencia
X-squared = 27887, df = 1, p-value < 2.2e-16
```

O resultado do teste qui-quadrado indica que há uma diferença estatisticamente significativa nas contagens de suicídios entre os anos de 2020 e 2021. O valor da estatística qui-quadrado é 27887, valor alto que indica diferença muito grande entre as duas comparações e o valor p é bem próximo a 0, o que deixa em evidência a diferença nas contagens de suicídios entre 2020 e 2021, mas não implica necessariamente em significância prática. A interpretação deve levar em consideração o contexto específico e a aplicabilidade dos resultados.

Distribuição de casos de suicídio por escolaridade em 2021:

Para essa análise, utilizamos o filtro de escolaridade em anos totais de escolaridade para verificar os casos de suicídio por faixa de escolaridade:

```
# Carregar os dados de 2021
dados_mortalidade_2021 <-
read.csv2("C:\\Users\\renan\\OneDrive\\Documentos\\estatistica\\Mortalidade_Geral_
2021.csv", sep = ";")

calcularTotalCasosSuicidioPorEscolaridade <- function(data) {
```

```

tryCatch({
  if ("CIRCOBITO" %in% colnames(data) && "ESC" %in% colnames(data)) {
    # Filtrar casos de suicídio (CIRCOBITO = 2) e considerar apenas valores
    permitidos em ESC
    casos_suicidio <- data[data$CIRCOBITO == 2 & data$ESC %in% c(1, 2, 3, 4, 5,
9), ]

    # Calcular o total de casos para cada nível de escolaridade
    total_por_escolaridade <- table(casos_suicidio$ESC)

    # Garantir que todos os níveis de escolaridade estejam presentes no resultado
    total_casos_suicidio_por_escolaridade <- integer(length = max(c(1, 2, 3, 4, 5,
9)))

    # Atribuir os totais para cada nível de escolaridade

    total_casos_suicidio_por_escolaridade[as.numeric(names(total_por_escolaridade))]
<- as.numeric(total_por_escolaridade)

    return(total_casos_suicidio_por_escolaridade)
  } else {
    return(rep(NA, length(c(1, 2, 3, 4, 5, 9)))) # Retorna um vetor de NA se os dados
não estiverem corretos
  }
}, error = function(e) {
  cat("Erro:", conditionMessage(e), "\n")
  return(rep(NA, length(c(1, 2, 3, 4, 5, 9)))) # Retorna um vetor de NA em caso de
erro
})
}

# Função para calcular o total de suicídios em um data frame
calcularTotalSuicidios <- function(data) {
  if ("CIRCOBITO" %in% colnames(data)) {

```

```

total_suicidios <- sum(data$CIRCOBITO == 2, na.rm = TRUE)
return(total_suicidios)
} else {
  return(NA) # Retorna NA se a coluna CIRCOBITO não estiver presente
}
}

# Calcular os totais de casos de suicídios por escolaridade para o ano de 2021
totais_casos_suicidio_2021 <-
calcularTotalCasosSuicidioPorEscolaridade(dados_mortalidade_2021)
total_suicidios_2021 <- calcularTotalSuicidios(dados_mortalidade_2021)

# Imprimir os totais de casos para o ano de 2021
print("Totais de casos de suicídios por escolaridade em 2021:")
print(totais_casos_suicidio_2021)
cat("Total de suicídios em 2021:", total_suicidios_2021, "\n")

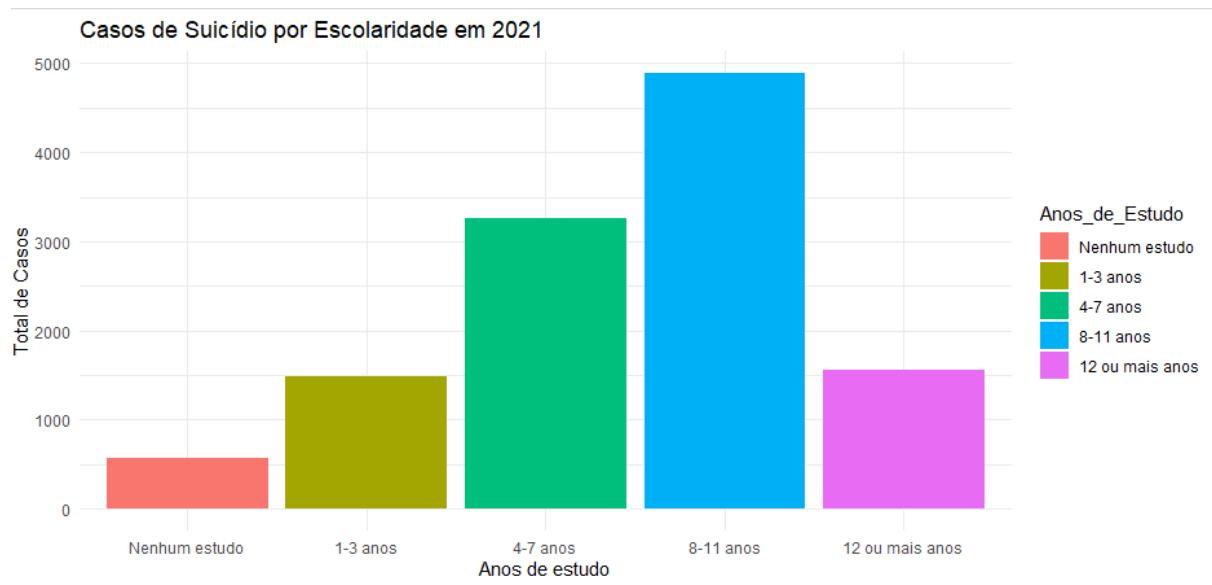
# Totais de casos de suicídios por escolaridade em 2021
totais_casos_suicidio_2021 <- c(574, 1484, 3259, 4891, 1562, 0, 0, 0, 2203)
#Valores colocado na mão pq eu estavapassando raiva tentando ajustar os vetores
# Níveis de escolaridade correspondentes (removendo os valores 0 e "Ignorado")
niveis_escolaridade <- c("Nenhum estudo", "1-3 anos", "4-7 anos", "8-11 anos", "12
ou mais anos")[totais_casos_suicidio_2021[1:5] != 0]

# Totais de casos de suicídios por escolaridade correspondentes (removendo os
valores 0 e "Ignorado")
totais_casos_suicidio_2021 <-
totais_casos_suicidio_2021[totais_casos_suicidio_2021 != 0][1:5]

# Criar um dataframe com os dados
dados_grafico <- data.frame(Niveis_Escolaridade = factor(niveis_escolaridade,
levels = niveis_escolaridade), Total_Casos_Suicidio = totais_casos_suicidio_2021)
# Criar o gráfico de colunas

```

```
grafico <- ggplot(dados_grafico, aes(x = Niveis_Escolaridade, y =
Total_Casos_Suicidio, fill = Niveis_Escolaridade)) +
  geom_bar(stat = "identity") +
  labs(title = "Casos de Suicídio por Escolaridade em 2021",
    x = "Níveis de Escolaridade",
    y = "Total de Casos") +
  theme_minimal()
print(grafico)
```



A distribuição por sexo nos casos de suicídio no ano de 2021:

Para esta análise, foi feita a divisão do total de mortes por suicídio por sexo (masculino e feminino), segue os códigos:

```
# Carregar os dados de 2021
dados_mortalidade_2021 <-
read.csv2("C:\\Users\\renan\\OneDrive\\Documentos\\estatistica\\Mortalidade_Geral_
2021.csv", sep = ";")

# Função para calcular o total de suicídios em um data frame
```



```

calcularTotalSuicidios <- function(data) {
  if ("CIRCOBITO" %in% colnames(data)) {
    total_suicidios <- sum(data$CIRCOBITO == 2, na.rm = TRUE)
    return(total_suicidios)
  } else {
    return(NA) # Retorna NA se a coluna CIRCOBITO não estiver presente
  }
}

# Função para verificar o sexo dos casos de suicídio
verificarSexoSuaicidio <- function(data) {
  if ("CIRCOBITO" %in% colnames(data) & "SEXO" %in% colnames(data)) {
    # Filtrar casos de suicídio (CIRCOBITO == 2)
    suicidios <- data[data$CIRCOBITO == 2, ]

    # Inicializar vetores para contar os casos por sexo
    total_M <- sum(suicidios$SEXO == 1, na.rm = TRUE)
    total_F <- sum(suicidios$SEXO == 2, na.rm = TRUE)
    total_I <- sum(suicidios$SEXO == 3, na.rm = TRUE)
    # Armazenar os totais em um vetor ou lista
    totais <- c(M = total_M, F = total_F, I = total_I)

    # Retornar uma lista com os totais
    return(totais)
  } else {
    stop("As colunas CIRCOBITO e/ou SEXO não estão presentes nos dados.")
  }
}

# Uso das funções
total_suicidios <- calcularTotalSuicidios(dados_mortalidade_2021)
totais_casos_suicidio_por_sexo <- verificarSexoSuaicidio(dados_mortalidade_2021)

```

```
# Imprimir resultados
print("Total de casos de suicídios em 2021:")
print(total_suicidios)

print("Totais de casos de suicídios por sexo em 2021:")
print(totais_casos_suicidio_por_sexo)

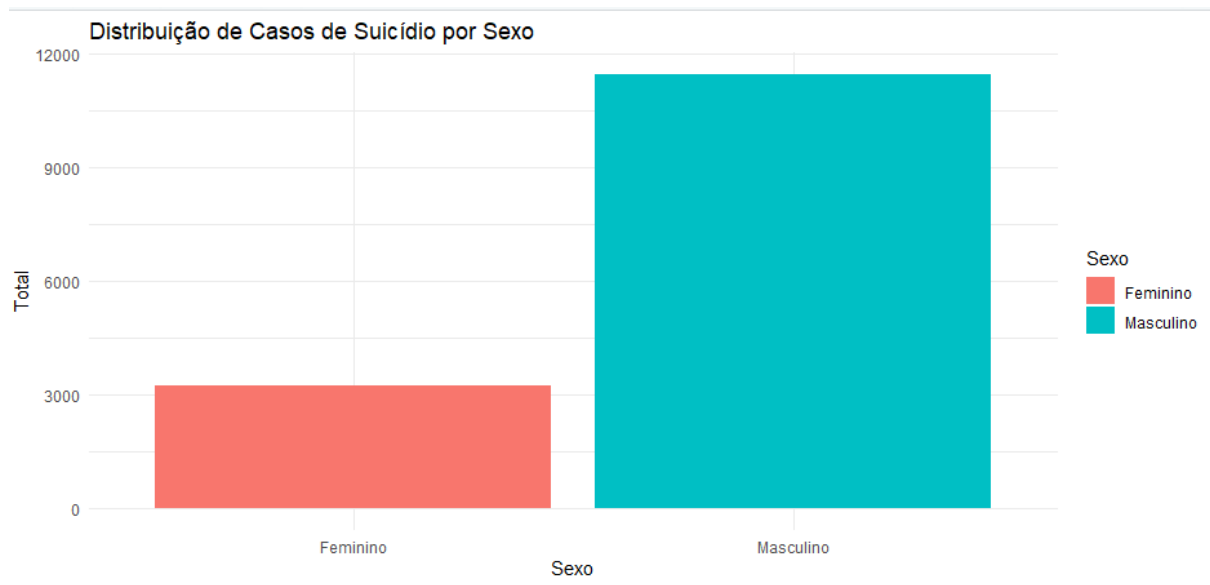
#####
# Carregar a biblioteca ggplot2
library(ggplot2)

# Remover a categoria "Indefinido"
dados_suicidio_por_sexo <- dados_suicidio_por_sexo[-3]

# Criar um data frame com os dados
df <- data.frame(Sexo = c("Masculino", "Feminino"), Total =
dados_suicidio_por_sexo)

# Criar o gráfico de colunas
grafico_colunas <- ggplot(df, aes(x = Sexo, y = Total, fill = Sexo)) +
  geom_col() +
  labs(title = "Distribuição de Casos de Suicídio por Sexo",
        x = "Sexo",
        y = "Total") +
  theme_minimal()

# Exibir o gráfico
print(grafico_colunas)
```



A distribuição por raça/cor

Para esta análise, foi feita a divisão das pessoas por cor para avaliar a distribuição dos casos de suicídio entre raças:

Código R:

```
# Função para verificar a raça/cor dos casos de suicídio
verificarRacaCorSuicidio <- function(data) {
  # Verificar se as colunas necessárias estão presentes nos dados
  if ("CIRCOBITO" %in% colnames(data) & "RACACOR" %in% colnames(data)) {
    # Filtrar casos de suicídio (CIRCOBITO == 2)
    suicidios <- data[data$CIRCOBITO == 2, ]

    # Contar casos por raça/cor
    total_Branco <- sum(suicidios$RACACOR == 1, na.rm = TRUE)
    total_Preta <- sum(suicidios$RACACOR == 2, na.rm = TRUE)
    total_Amarela <- sum(suicidios$RACACOR == 3, na.rm = TRUE)
    total_Pardo <- sum(suicidios$RACACOR == 4, na.rm = TRUE)
    total_Indigena <- sum(suicidios$RACACOR == 5, na.rm = TRUE)
```

```

# Armazenar os totais em um vetor ou lista
totais <- c(Branco = total_Branco, Preta = total_Preta, Amarela = total_Amarela,
Pardo = total_Pardo, Indigena = total_Indigena)

# Retornar um vetor com os totais
return(totais)
} else {
  stop("As colunas CIRCOBITO e/ou RACACOR não estão presentes nos dados.")
}
}

# Uso da função
totais_casos_suicidio_por_racacor <-
verificarRacaCorSuicidio(dados_mortalidade_2021)

# Imprimir resultados
print("Totais de casos de suicídios por raça/cor em 2021:")
print(totais_casos_suicidio_por_racacor)
#####

# Carregar a biblioteca ggplot2
library(ggplot2)

# Criar um data frame com os dados
df_racacor <- data.frame(
  RacaCor = c("Branco", "Preta", "Amarela", "Pardo", "Indigena"),
  Total = totais_casos_suicidio_por_racacor
)

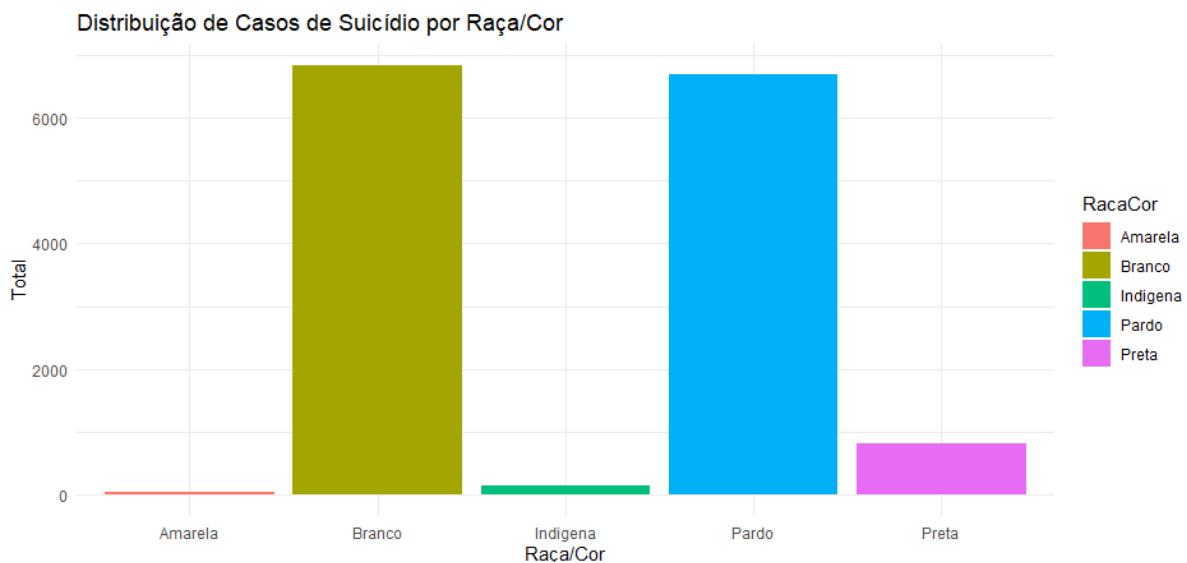
# Criar o gráfico de colunas
grafico_colunas_racacor <- ggplot(df_racacor, aes(x = RacaCor, y = Total, fill =
RacaCor)) +
  geom_col() +

```

```
labs(title = "Distribuição de Casos de Suicídio por Raça/Cor",
     x = "Raça/Cor",
     y = "Total") +
theme_minimal()
```

Exibir o gráfico

```
print(grafico_colunas_racacor)
```



Quantas mortes decorrentes de acidentes de trabalho ocorreram em 2020 e 2021?

Código R:

```
# Carregar os dados de 2020
```

```
dados_mortalidade_2020 <-
```

```
read.csv2("C:\\Users\\renan\\OneDrive\\Documentos\\estatistica\\Mortalidade_Geral_2020.csv", sep = ";")
```

```
# Carregar os dados de 2021
```

```
dados_mortalidade_2021 <-
read.csv2("C:\\Users\\renan\\OneDrive\\Documentos\\estatistica\\Mortalidade_Geral_
2021.csv", sep = ";")
```

```
# Função para verificar acidentes
```

```
verificarAciden <- function(data) {
  # Verificar se a coluna necessária está presente nos dados
  if ("ACIDTRAB" %in% colnames(data)) {
    # Filtrar acidentes
    Ignorado <- sum(data$ACIDTRAB == 9, na.rm = TRUE)
    Sim <- sum(data$ACIDTRAB == 1, na.rm = TRUE)
    Nao <- sum(data$ACIDTRAB == 2, na.rm = TRUE)

    # Armazenar os totais em um vetor
    totais <- c(Ignorado = Ignorado, Sim = Sim, Nao = Nao)

    # Retornar um vetor com os totais
    return(totais)
  } else {
    stop("A coluna ACIDTRAB não está presente nos dados.")
  }
}
```

```
# Uso da função para 2020
```

```
totais_acidentes_2020 <- verificarAciden(dados_mortalidade_2020)
```

```
# Uso da função para 2021
```

```
totais_acidentes_2021 <- verificarAciden(dados_mortalidade_2021)
```

```
# Imprimir resultados para 2020
```

```
print(totais_acidentes_2020)
```

```
# Imprimir resultados para 2021
```

```
print(totais_acidentes_2021)
```

```
> print(totais_acidentes_2020)
```

Ignorado	Sim	Nao
13957	3035	28075

```
> print(totais_acidentes_2021)
```

Ignorado	Sim	Nao
14246	3553	28993

Houve aumento significativo entre os anos ?

```
> print("Resultados do teste qui-quadrado (apenas casos  
'Sim'):")
```

```
[1] "Resultados do teste qui-quadrado (apenas casos 'Sim'):"
```

```
> print(teste_qui_quadrado)
```

Chi-squared test for given probabilities

```
data:  tabela_contingencia
```

```
X-squared = 40.729, df = 1, p-value = 1.749e-10
```

```
> print("Diferença nas proporções entre 2021 e 2020:")
```

```
[1] "Diferença nas proporções entre 2021 e 2020:"
```

```
> print(aumento_proporcao)
```

```
Sim
```

```
0.008587607
```

A análise permite concluir que houve um aumento significativo nos acidentes de trabalho entre os anos de 2020 e 2021.

Qual é a distribuição por raça/cor?

Código R:

```
# Carregar os dados de 2021
```

```
dados_mortalidade_2021 <-
read.csv2("C:\\Users\\renan\\OneDrive\\Documentos\\estatistica\\Mortalidade_Geral_
2021.csv", sep = ";")
```

```
verificarRacaCorAcidentesTrabalho <- function(data) {
  # Verificar se as colunas necessárias estão presentes nos dados
  if ("ACIDTRAB" %in% colnames(data) & "RACACOR" %in% colnames(data)) {
    # Filtrar casos de acidentes de trabalho (ACIDTRAB == 1)
    acidentes_trabalho <- data[data$ACIDTRAB == 1, ]

    # Contar casos por raça/cor
    total_Branco <- sum(acidentes_trabalho$RACACOR == 1, na.rm = TRUE)
    total_Preta <- sum(acidentes_trabalho$RACACOR == 2, na.rm = TRUE)
    total_Amarela <- sum(acidentes_trabalho$RACACOR == 3, na.rm = TRUE)
    total_Pardo <- sum(acidentes_trabalho$RACACOR == 4, na.rm = TRUE)
    total_Indigena <- sum(acidentes_trabalho$RACACOR == 5, na.rm = TRUE)

    # Armazenar os totais em um vetor ou lista
    totais <- c(Branco = total_Branco, Preta = total_Preta, Amarela = total_Amarela,
Pardo = total_Pardo, Indigena = total_Indigena)

    # Retornar um vetor com os totais
    return(totais)
  } else {
    stop("As colunas ACIDTRAB e/ou RACACOR não estão presentes nos dados.")
  }
}
```

```
# Uso da nova função para acidentes de trabalho
totais_acidentes_trabalho_por_racacor <-
verificarRacaCorAcidentesTrabalho(dados_mortalidade_2021)
# Imprimir resultados
print("Totais de casos de acidentes de trabalho por raça/cor em 2021:")
```



```
print(totais_acidentes_trabalho_por_racacor)
```

```
> print(totais_acidentes_trabalho_por_racacor)
```

Branco	Preta	Amarela	Pardo	Indigena
1552	223	12	1720	12