Sujita Mini Project

https://catalog.data.gov/dataset/nchs-death-rates-and-life-expectancy-at-birth/resource/eb20e7ad-2a82-4dce-82df-503a0bdb27be

```
library(ggplot2)
library(dplyr)
```

```
Attaching package: 'dplyr'

The following objects are masked from 'package:stats':

    filter, lag

The following objects are masked from 'package:base':

    intersect, setdiff, setequal, union
```

```
library(maps)
library(plotly)
```

```
Attaching package: 'plotly'

The following object is masked from 'package:ggplot2':

    last_plot

The following object is masked from 'package:stats':

    filter
```

```
The following object is masked from 'package:graphics':

    layout
```

```
library(viridis)
```

```
Loading required package: viridisLite


Attaching package: 'viridis'

The following object is masked from 'package:maps':

    unemp
```

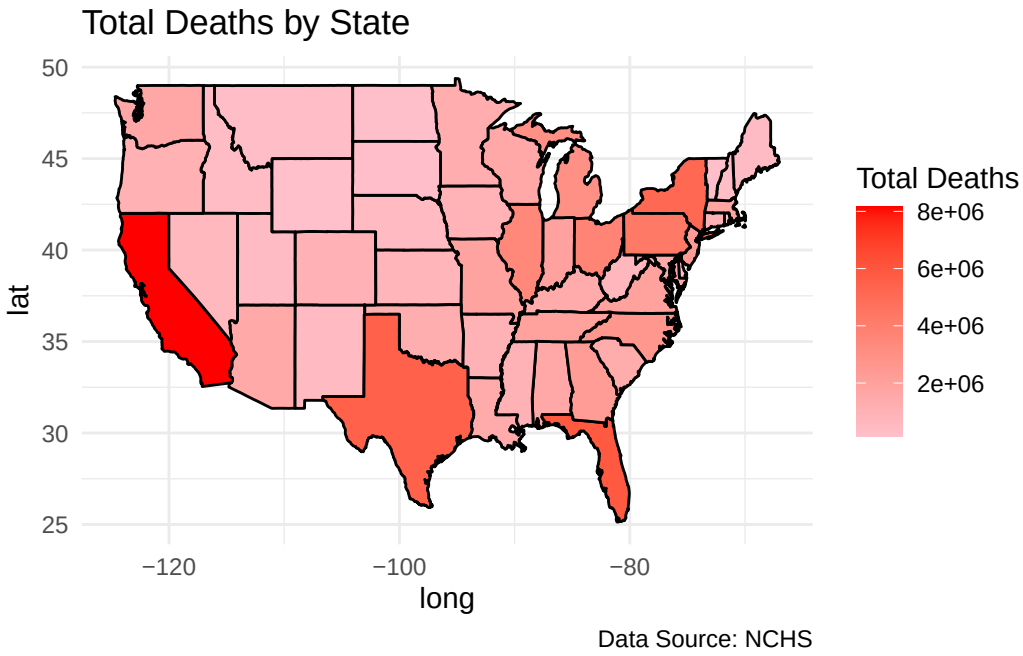
```
library(htmlwidgets)
library(webshot2)
library(readr)

data <- read_csv("NCHS_-_Leading_Causes_of_Death__United_States.csv")
```

```
Rows: 10868 Columns: 6

-- Column specification --------------------------------------------------------
Delimiter: ","
chr (3): 113 Cause Name, Cause Name, State
dbl (3): Year, Deaths, Age-adjusted Death Rate

i Use `spec()` to retrieve the full column specification for this data.
i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
df_state_deaths <- data %>%
  group_by(State) %>%
  summarise(Total_Deaths = sum(Deaths, na.rm = TRUE))

us_states <- map_data("state")

df_state_deaths <- df_state_deaths %>%
  mutate(State = tolower(State))

data_map <- us_states %>%
  left_join(df_state_deaths, by = c("region" = "State"))
```

Static plot for numerical variable

```
static_numerical_plot <- ggplot(data_map, aes(x = long, y = lat, group = group, fill = Total_
  geom_polygon(color = "black") +
  scale_fill_gradient(low = "pink", high = "red", na.value = "blue") +
  labs(title = "Total Deaths by State",
       fill = "Total Deaths",
       caption = "Data Source: NCHS") +
  theme_minimal()

print(static_numerical_plot)
```



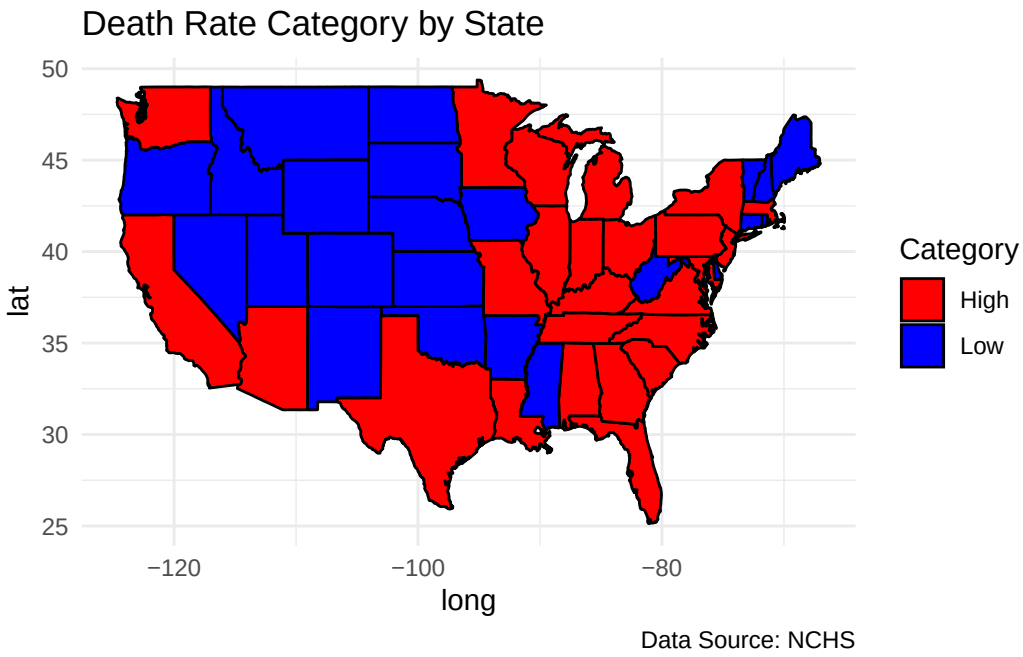
This is a choropleth map of the United States visualizing total deaths by state, with states shaded in varying intensities of red based on death counts. The title at the top reads "Total Deaths by State", and the legend on the right indicates that darker red represents higher total deaths, while lighter shades represent lower numbers. The x-axis represents longitude, and the y-axis represents latitude, positioning the states geographically. A caption at the bottom credits the NCHS (National Center for Health Statistics) as the data source. The map clearly highlights states like California and Texas as having the highest total deaths, whereas states in the Midwest and Mountain West have lower totals.

Interactive plot for numerical variable

```
interactive_numerical_plot <- ggplot(data_map, aes(x = long, y = lat, group = group, fill = 
                                                   text = paste("State:", region, "<br>Total
  geom_polygon(color = "white") +
  scale_fill_viridis_c(option = "plasma", na.value = "gray") +
  labs(title = "Total Deaths by State",
       fill = "Total Deaths",
       caption = "Data Source: NCHS") +
  theme_minimal()

interactive_numerical_plotly <- ggplotly(interactive_numerical_plot, tooltip = c("text"))


print(interactive_numerical_plotly)
```

Static plot for categorical variable

```
df_state_deaths <- df_state_deaths %>%
  mutate(Category = ifelse(Total_Deaths > median(Total_Deaths, na.rm = TRUE), "High", "Low"))

us_states <- map_data("state")

data_map <- us_states %>%
  left_join(df_state_deaths, by = c("region" = "State"))


static_categorical_plot <- ggplot(data_map, aes(x = long, y = lat, group = group, fill = Cate
  geom_polygon(color = "black") +
  scale_fill_manual(values = c("High" = "red", "Low" = "blue"), na.value = "gray") +
  labs(title = "Death Rate Category by State",
       fill = "Category",
       caption = "Data Source: NCHS") +
  theme_minimal()

print(static_categorical_plot)
```



Interactive plot for categorical variable

```
interactive_categorical_plot <- ggplot(data_map, aes(x = long, y = lat, group = group, fill =
                                                     text = paste("State:", region, "<br>Cate
  geom_polygon(color = "white") +
  scale_fill_manual(values = c("High" = "grey", "Low" = "lightgreen")) +
  labs(title = "Death Rate Category by State",
       fill = "Category",
       caption = "Data Source: NCHS") +
  theme_minimal()

interactive_categorical_plotly <- ggplotly(interactive_categorical_plot, tooltip = c("text"))

print(interactive_categorical_plotly)
```