ISTE 782: Visual Analytics

Assignment 2

Due date and place: October 22, 2024, 3:30 PM

Part 1:

Question 1. Descriptive Statistics (10 points):

Use the **statedata.csv** and provide a table of descriptive statistics using the **stargazer** package in R. Include summary statistics such as mean, median, standard deviation, and range for the critical variables in the dataset. Use sample code in **stargazer.pdf**

Question 2. Life Expectancy Analysis (15 points):

With the **tidyr** and **dplyr** packages, analyze how life expectancy varies across states. Identify which provinces have the highest and lowest life expectancy.

Question 3. Correlation Heatmap (20 points):

With ggplot2 and **geom_tile(),** create a heatmap showing the correlations between key variables in the statedata dataset. Which pairs of variables exhibit the strongest positive or negative correlations? Are any of these correlations surprising to you? Use sample code in **correlation_heatmap.pdf**

Part 2:

Question 1. Winners (10 Points)

Use the **1976-2020-president.csv** dataset, **tidyr** and **dplyr** packages to calculate which party won presidential elections in each year. (10 points)

Question 2. Election Map (10 Points)

Using the **usmap** library in R, create a map showing the political parties that won the elections in each state in the 2020 elections. Color the Republican party red and the Democratic party blue.

Question 3. Facet plot maps (20 Points)

3.1. Create a facet plot showing the political parties that won the elections in each state in the 2004, 2008, 2012, and 2016 elections. Use **facet_wrap()** in ggplot2 to visualize **(15 points).**

3.2. Comment on the maps: are any swing states changing between Democrats and Republicans in different election years? Which states are they? **(5 points)**

Bonus Question 1: (2 points) Create a Beamer presentation in R Markdown that includes all the plots and analysis from Parts 1 and 2.

Bonus Question 2: (5 points) Provide the link to the Rshiny app and the R code of your app, which has a slider for 2004, 2008, 2012, and 2016. When you slide, it shows presidential election maps for the given year.

Part 3: There are three bugs in the R Shiny code provided below. Can you find and fix them? (5 points)

```
library(shiny)
library(ggplot2)

datasets <- c("economics", "faithfuld", "seals")
ui <- fluidPage(
  selectInput("dataset", "Dataset", choices = datasets),
  verbatimTextOutput("summary"),
  tableOutput("plot")
)

server <- function(input, output, session) {
  dataset <- reactive({
    get(input$dataset, "package:ggplot2")
  })
  output$summmry <- renderPrint({
    summary(dataset())
  })
  output$plot <- renderPlot({
    plot(dataset)
  }, res = 96)
}

shinyApp(ui, server)
```

General rules:

- You will be using **R** for this assignment.
- Save your codes in a single file with the **.R** extension. Your code needs to be divided into segments and tagged with comments to designate different questions.
- Save your tables, images, and explanations in a .doc or .pdf file.
- Name your tables and images appropriately (for instance, Table 1: Correlation Table, All Data)
- Please use the following as the title :

MyFirstName_MyLastName_MyIDNumber_Assignment2