Homework 4: Data Ethics and Visualization

Student Giovanny Joseph
Professor Monica Morales

October 2024

Critique of Graph

The original graph provided by The New York Times has several issues that may make it confusing or misleading to viewers. One major problem is the graph's layout. Pagination causes some titles to be cut off in certain views, which creates clutter and prevents viewers from easily understanding the data. A more compact design or a reduction in the amount of information displayed could make the graphic clearer. Another issue is the overlap of circles surrounding the numbers. This overlap does not represent any meaningful aspect of the data and only adds visual noise, detracting from the clarity of the presentation. The circles do not convey additional insights and instead clutter the graphic unnecessarily.

Furthermore, the graph lacks proper context for interpreting the data. While it shows how often candidates used certain words, it does not explain the significance of this frequency. Without providing trends or analyzing the tone or meaning behind the word usage, the graph could mislead viewers into drawing superficial conclusions. For example, one might incorrectly assume that the frequent use of a certain word by a candidate reflects a deeper value or stance, when in fact it may not.

Redesign of Graph

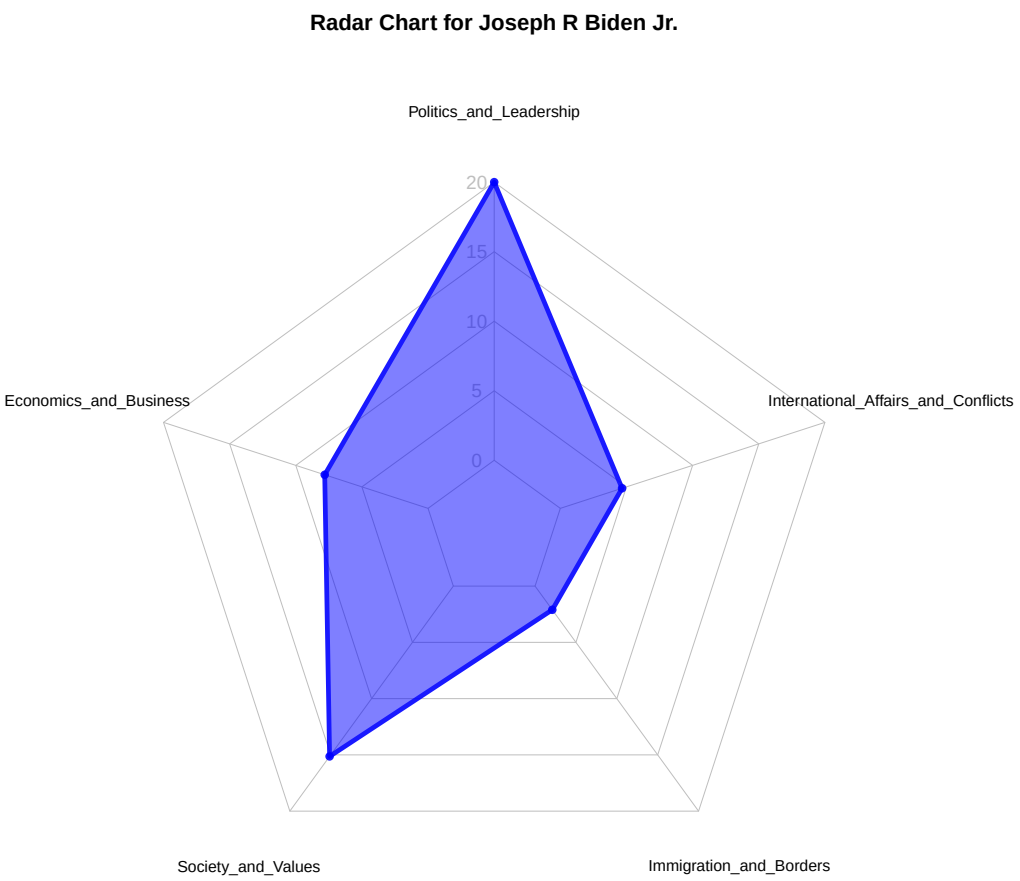

Figure 1: Radar Chart for Joe Biden

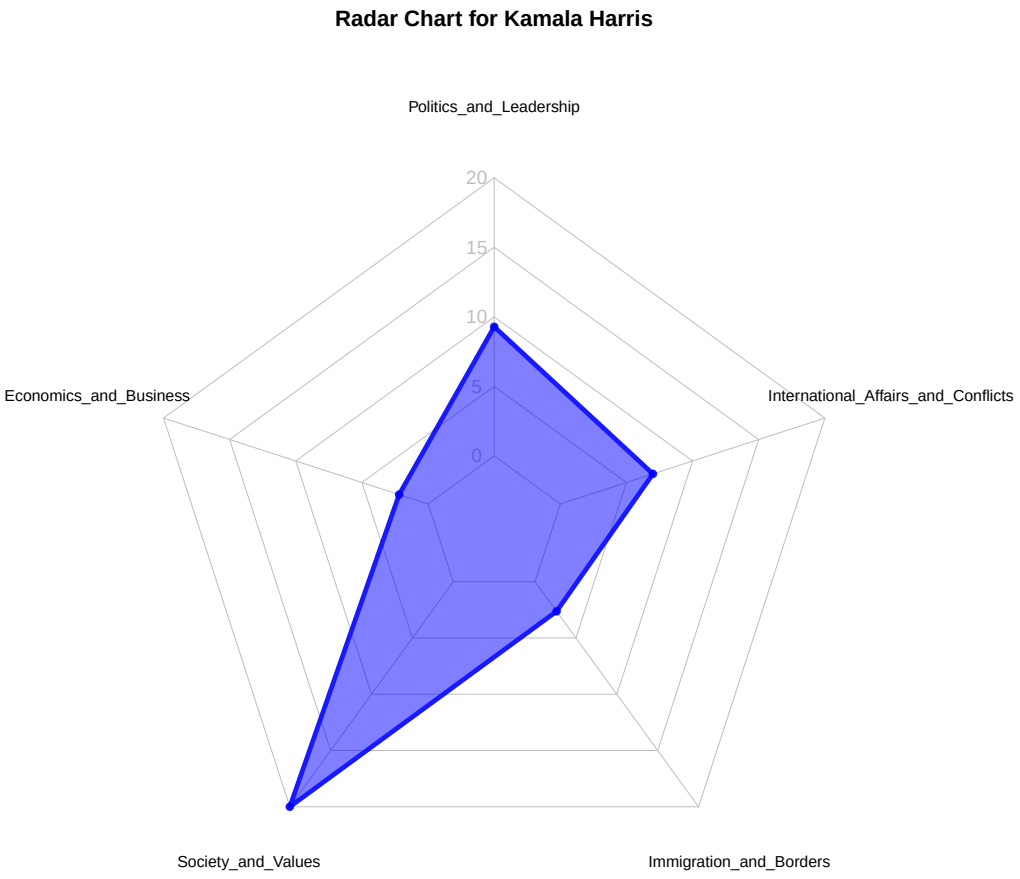

Figure 2: Radar Chart for Kamala Harris

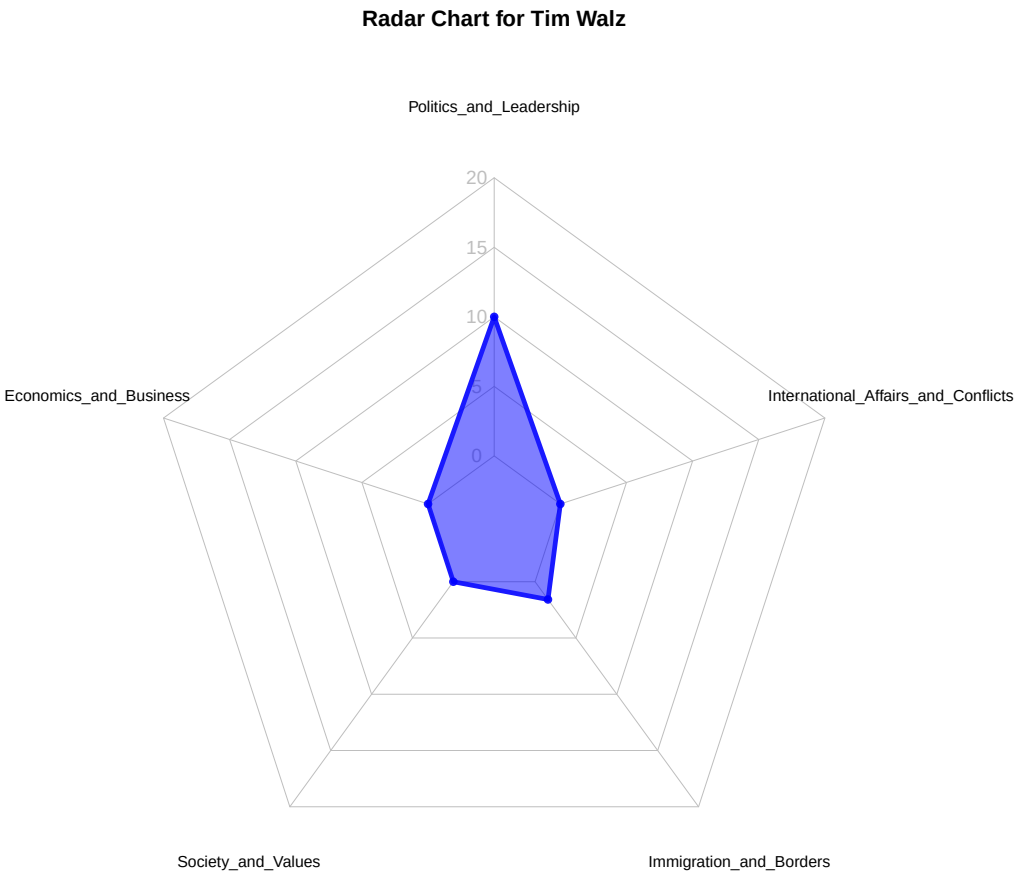

Figure 3: Radar Chart for Tim Walz

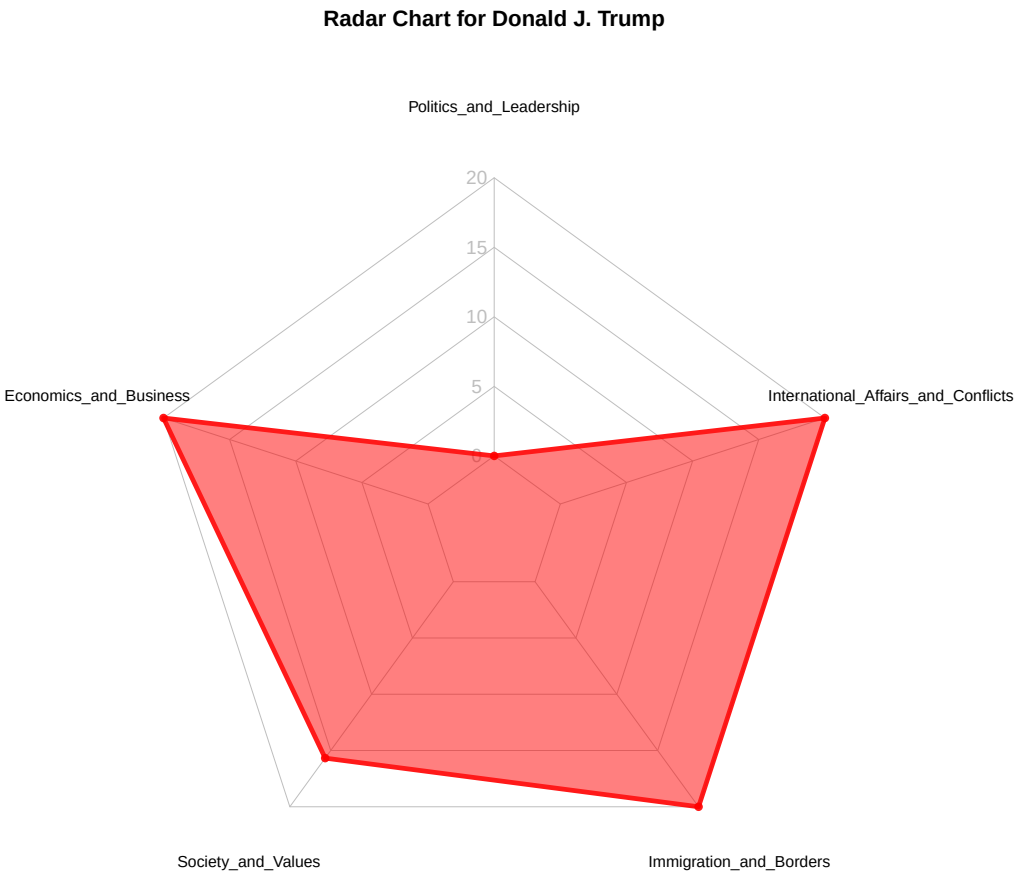

Figure 4: Radar Chart for Donald Trump

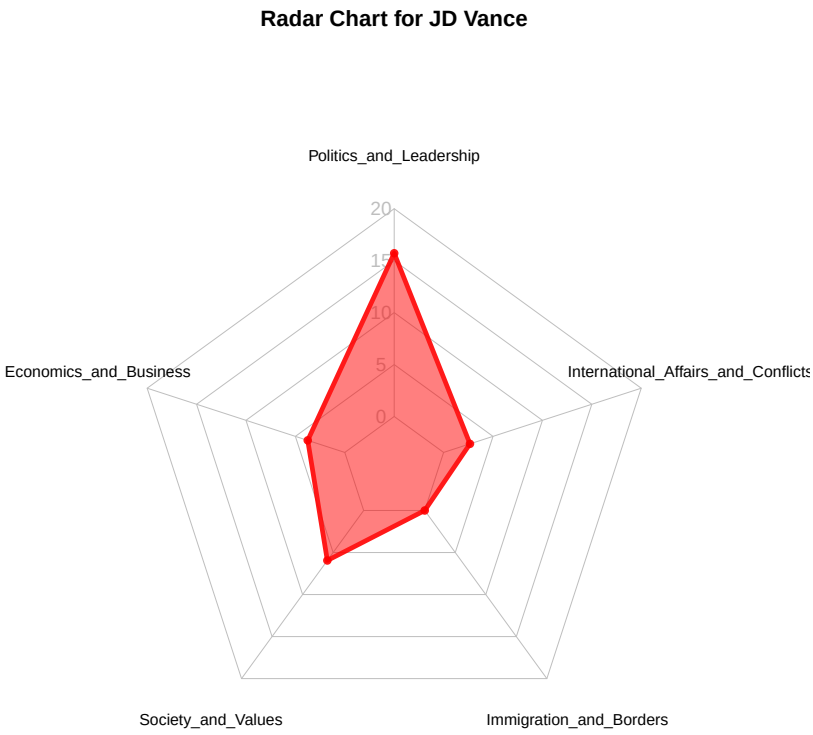

Figure 5: Radar Chart for JD Vance

Explanation

The graphics I created use radar charts to significantly improve upon the original graphic from The New York Times. The compact design of these charts enhances clarity and focus, allowing readers to easily interpret the data without the distraction of visual noise. Unlike the original graphic, which was cluttered with overlapping circles and pagination issues that obscured information, the radar charts offer a clean and organized presentation. Each chart is designed to fit neatly on a single page, ensuring that titles and context remain intact, making it easier for readers to absorb all necessary information without scrolling through multiple pages.

The simplified structure of the radar charts allows for quick analysis of key topics while focusing on individual candidates, making it easier to interpret their word usage in a meaningful way. By organizing the data into five distinct categories—such as Politics and Leadership or Immigration and Borders—the charts provide clear context, reducing ambiguity and helping viewers quickly grasp the values and focus of each candidate. The drastic reduction of visual noise, by eliminating unnecessary elements, highlights the most important points of discussion, creating a more elegant and effective visual that enhances the overall understanding of the data.

Appendix A: R Code

```r
# Load necessary libraries
library(fmsb)
library(readr)

# Load the CSV file and inspect
words_data <- read_csv("words.csv")

# Categories for words in the CSV
person_categories <- data.frame(
  person = words_data$Name,
  Politics_and_Leadership = rowSums(words_data[, c("Trump", "Biden", "Kamala", "Harris")], na.rm = TRUE),
  Economics_and_Business = rowSums(words_data[, c("Economy, economic", "Business", "Job(s)", "Tax(es)", "Inflation")], na.rm = TRUE),
  Society_and_Values = rowSums(words_data[, c("Democacy", "Freedom", "Law", "God", "Love", "Family", "Father", "Mother", "Abortion")], na.rm = TRUE),
  Immigration_and_Borders = rowSums(words_data[, c("Border", "Immigrant, immigration", "Invasion", "Illegal Aliens", "Neighbor, neighborhood")], na.rm = TRUE),
  International_Affairs_and_Conflicts = rowSums(words_data[, c("War", "Ukraine", "Russia", "Putin", "China, Chinese", "Israel", "Gaza", "Hamas", "Terrorist", "Iran", "Afganistan")], na.rm = TRUE)
)

# Find the maximum and minimum for each category to normalize radar chart
max_values <- apply(person_categories[, -1], 2, max)
min_values <- apply(person_categories[, -1], 2, min)

# Add max and min rows to the dataset for radar chart scaling
person_categories_scaled <- rbind(max_values, min_values, person_categories[, -1])

# Function to create radar chart for each person
create_radar_chart <- function(person_data, person_name) {
  if (person_name %in% c("Donald J. Trump", "JD Vance")) {
    polygon_color <- rgb(1, 0, 0, 0.9)  # Red polygon
    fill_color <- rgb(1, 0, 0, 0.5)     # Semi-transparent red fill
  } else {
    polygon_color <- rgb(0, 0, 1, 0.9)  # Blue polygon
    fill_color <- rgb(0, 0, 1, 0.5)     # Semi-transparent blue fill
  }
  par(mar = c(5, 5, 5, 5))

  radarchart(
    person_data,
    axistype = 1,
    pcol = polygon_color,
    pfcol = fill_color,
    plwd = 4,
    cglcol = "grey",
    cglty = 1,
    axislabcol = "grey",
    caxislabels = seq(0, 20, 5),
    cglwd = 0.8,
    vlcex = 0.8,
    title = paste("Radar Chart for", person_name)
  )
}

# Loop through each person to create radar charts
for (i in 3:nrow(person_categories_scaled)) {
  person_name <- person_categories$person[i - 2]
  person_data <- rbind(max_values, min_values, person_categories_scaled[i, ])

  create_radar_chart(person_data, person_name)
}
```