Cluster Analysis of Billionaires

Contribution: This was an individual project and is fully my own work.

Tools: The code for this project was written in R, utlizing ggplot2 and dplyr from the tidyverse stack for data exploration and manipulation. Clustering was performed using the NBClust and dbscan packages.

Performing NLP on Fox News Newsletters

Contribution: The portion of the report titled "Linear Discriminant Analysis", pages 7 - 9; my contribution is detailed more fully on pages 11 - 12.

Tools: The code for this project was written in R. LDA and PCA functions were taken from the MASS and psych packages, respectively. I wasn't yet used to using the tidyverse stack at the time of this analysis, so all manipulation and visualization was done in base R.

Regression Analysis of Chicago's Air Quality During the COVID-19 Pandemic

Contribution: Code for this analysis was written collaboratively as a group via Zoom call, so no individual credit can be claimed for that part of the analysis. I did, however, do a majority of the literature review for the project, and wrote the introduction in full.

Tools: The code for this project was written in SAS.

Note: Regrettably, I no longer have access to the code to run this analysis, nor could I run it again if I did since I also currently lack access to SAS. However, I am still including it in this portfolio to show my skills in scientific writing and literature review.

Visualizing Trends in the Olympic Games Over Time

Contribution: My main contributions to this report are Fig 10, titled "Representation of Top 5 Summer Olympic Sports Over Time" on page 8, and Fig 11, titled "Relative Frequency of Athletes' Ages Over Time" on page 9, as well as the written analysis for both visualizations, which continues onto page 10. I also contributed the two exploratory ridgeline plots titled "BMI by Sex and Era" and "Age by Sex and Era", both found on page 5.

Tools: The code for this project was written in R, using the tidyverse packages dplyr, tidyr, and ggplot2 for manipulating, tidying, and visualizing the data.

Note: There is an appendix in the report PDF containing all code, but for easier reference I also included an R Markdown file that is just my own code.