CS432: Final Project Report

Analysis of COVID-19 Infection in comparison to 2020 Presidential Election Results

Samson Grimm

I. Problem

I chose to analyze any possible correlation between COVID-19 infection numbers per state and how each state voted in the 2020 Presidential election. I found the total percentage of COVID-19 infections and COVID-19 related deaths in the country as of November 3rd, 2020, and then I found how those percentages differed when just looking at states that voted red and states that voted blue. For the data, I used a public data set of COVID-19 data for each state (Washington DC included) that had not been update since election day 2020, and added what party each state voted for to the dataset.

II. Software Design and Implementation

I chose MongoDB as my database. I was able to download the dataset as a .CSV file and easily import it using MongoDB Compass, which I also used to host it on a local server. I implemented my program in Java, using my program from Project 2 as a basis to build off of. I also used the MongoDB Java Drivers library to integrate my database into the Java program.

Firstly, while testing my program I found the log messages that MongoDB would output to be quite annoying, and I felt that they would be detrimental to the user experience. I found the Java library Logback, itself requiring the library SLF4J, which was able to suppress those messages. When the program starts, the user is given the option whether or not they would like to see the log messages. If they choose not to, Logback is used to disable them,

Following that choice, the program finds the total number of infections and deaths for all states, as well as the totals for just red and blue states. These are stored ahead of time so that they don't have to be calculated each time the user wishes to view this information. The user is then asked if they would like to view the information for a specific state (where they are also given the option to view the information for the entire country) or view the data for states split among political parties. The program will wait for user input, display the requested information, and then return to the main menu once the user presses the 'ENTER' key.

III. Project Outcome

	Red States	Blue States	Total
Total Infections	4,531,333	4,629,627	9,160,960
Infections as a % of Population	3.19%	2.44%	2.76%
Total Deaths	83,257	138,988	222,245
Deaths as a % of Population	0.0587%	0.0734%	0.0671%

By the time of the 2020 Presidential Elction, states that voted for Trump on average had a higher rate of COVID-19 infection, while states that voted for Biden had a higher rate of COVID-19 related deaths. My main take away from this is that red states likely were not taking COVID-19 as seriously as blue states, and the higher infection rate in red states could be a result of less safety measures in those states. Meanwhile, the lower rate of COVID-19 deaths in those states would also contribute to them not taking it as seriously, as they would view it as less of a threat. This would explain why these states would vote for the candidate that was not as worried about the pandemic.

In blue states, while the average rate of infection was lower, the average rate of death was higher. Again, the lower infection rates could be from these states taking the pandemic more seriously, and the higher rate of death could have motivated them to take it seriously. Like the red states, this would explain why these states would vote for the candidate that saw COVID-19 as an issue that needed to be dealt with.

References

[1] Ranger, Night. "Covid-19 State Data." *Kaggle*, 3 Nov. 2020, https://www.kaggle.com/datasets/nightranger77/covid19-state-data.

[2] "2020 Presidential Election Results." *CNN*, Cable News Network, https://www.cnn.com/election/2020/results/president.