

PS6 Finley

Caden Finley

March 2023

1 Process

For my data, I wanted to find the game between the San Antonio Spurs and the Los Angeles Lakers on January 25th, 2023. This data shows a variety of information for the performance of both teams. The data includes Field Goal percentage, Rebounds, and more. In order to isolate this data, I first found the data from the ESPN website of the game. From here I copied the URL and parsed the HTML file with "SelectorGadget". I pulled the table containing the relevant data that was called ".TeamStatsTable". Once I had all this, it was a matter of cleaning the data. I started with using the "rvest" package and assigning the different variables for URL, page, and table. From here I needed to extract the column and header names from the data in order to reproduce it. This was done with the HTML nodes function to isolate the column and row data. From here I needed team headers to order my data with the correct stats. I essentially took my data variable and used the subsections in it to get what was desired. This included Team and Abbreviation for both teams. Once the column and row names were set, I looked to remove any unnecessary ones. There also became an issue with some columns not being named correctly, so I had to manually assign names to them. From here I need to convert numeric columns to numeric format in order to help read the data. The last step was to convert the data to a CSV in order to be read and visualized in other ways.

2 Visualization

The data should visualize what would be the performance of both teams. Looking at shot percentages, fouls, blocks, steals, rebounds, etc. This would give an idea as to what each team succeeds more than the other at. This could also be used to understand how the team typically plays if you look at more of a total distribution for a team. In the case of visualizing this data, it can tell who was better in certain categories and of course who won the game. Visualizing this data could help the viewer understand how good of a rebounding team the Spurs are or how good the Lakers are at shooting 3's for example. I tried to use GGplot as well as a Bar chart to demonstrate how you can isolate certain stats of the game in order to give you a better look at how the team performs

in certain periods or in certain situations. Lastly, I tried to implement a heat map to look at the concentration of where and how many 3 point shots were made by both teams.