Problem Set 6

Dillon Beake

March 8, 2024

1 Overview

Three different visualizations were created for this project. I used data from the 2007 NFL Draft. The Visualizations are included below. Before I move on, I want to discuss any aspects of the data I had to clean. To clean the data I took 3 approaches. 1) For the Scatter Plot, I excluded any entries that had an "NA", instead of a value that I wanted to include. this was written in the code, as a command line. 2) For the Box Plot I deleted any item that had only 1 example. For example, there were a few positions that were coded wrong, i.e. "S = safety", instead of "SS", or "FS", strong safety and free safety respectively. This was written in the code as a command line. This was done to make the visualization easier to read and it did not take away from the message I wanted it to convey. 3) I noticed in my Data that there were Draft Rounds that did not exist. To correct the draft rounds I went through the data and manually changed the draft round for each player. For future research. I will find data NFL Draft data that at least has the correct draft. that was painstaking for even one year I will want various details about each player for 10-20 years.

2 Scatter Plot

2.1 Scatter Plot

The scatter plot visualization was used to demonstrate the average height and weight of the different offensive positions in the 2007 NFL Draft. I am interested in this visualization because I want to analyze differences in football players. As part of the my research, I want to analyze the difference between the player positions, understand the norms or these positions, and also any trends of how they are changing over time.

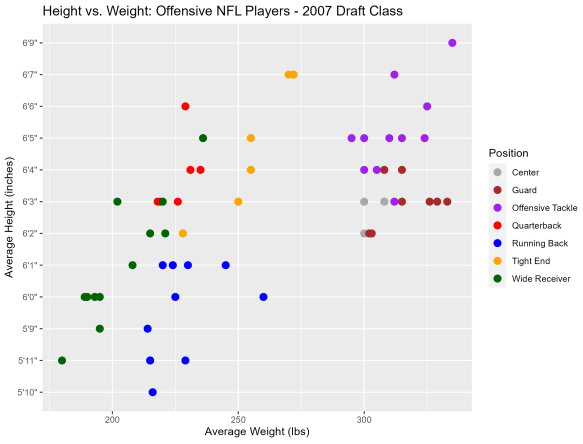

Figure 1: Scatter Plot: Height vs Weight Offensive Players in 2007 NFL Draft.

2.2 Histogram

The Histogram is a visualization of the number of players drafted in the 2007 draft from each of the 7 rounds. The data I used was an incomplete list. So, I was curious about this data. Also, I wanted to learn more about what I could do with the data set and how that could be incorporated in a histogram. For future Research, I might make a variable that separates players by arm length and create a variable. I could set that as the x-axis and draft-round as the y axis. That would be a very helpful visualization. However, I did not have that specific data. I used this part of the assignment to learn what I could do with a histogram.

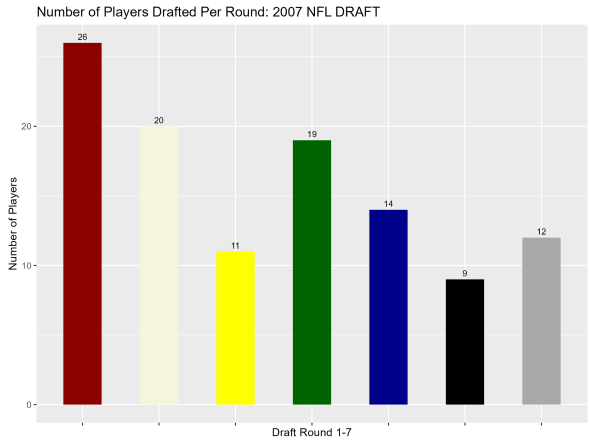

Figure 2: Histogram: Players Drafted Per Round: 2007 NFL Draft

2.3 Box Plot

The Box Plot demonstrates the Weight Range of each position in the 2007 NFL Draft. I found this interesting because it is yet another way I can visualize anthropometric measurements taken from data. This is more for me to visualize the norms of each position. These norms can be weight, height, body fat percentage, etc. If I have access to that data.

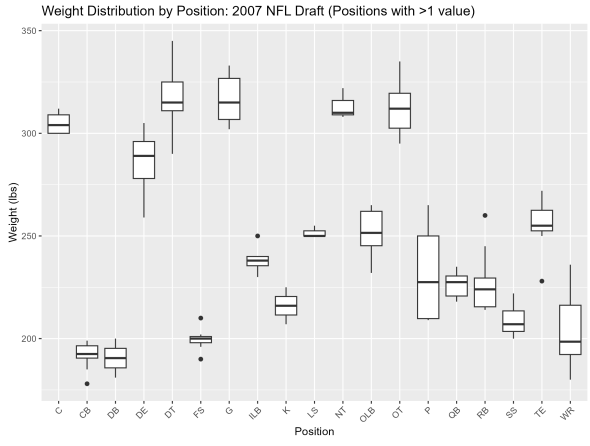

Figure 3: Box Plot: Average weight Range Per Position in the 2007 NFL Draft