A time in which I performed exploratory data analysis was this past semester in DSC495-002 to determine whether or not awareness was used to influence Madden overalls. This analysis was inspired by a passion for the game Madden, and from my own personal experiences playing, that left me questioning of overall was important. I suspect that awareness can inflate a players overall if that player is popular or highly paid. This can keep the fans happy as they would not like to see their favorite player with just an alright overall even if that player may be overrated.

When summarizing the findings, an important place to start is the correlations between awareness and both overall and total salary. In Graph 1 as seen in the index, there is a clear strong, positive correlation between awareness and total salary. This graph, while showing a strong correlation, was not nearly as telling as the graph between awareness and overall. As seen in Graph 2, the trendline almost perfectly fits the points of the plot, showing a very strong, positive correlation. These initial correlations were very promising in regards to my hypothesis that awareness was used to predict overall.

To create adjusted overall, I created a 21 part conditional logic chain that assigned a custom formula to each position group to calculate their overall with awareness removed. These formulas were pulled from the EA.com website where the formula was provided. My initial plan was to create a scatterplot of overall and adjusted overall, Graph 3, but it was hard to see any trends, as the points were so close together. I tried to look at the spread between each players overall and adjusted overall in Graph 4, however that was also unsuccessful. In Graph 5 however, I was able to make a big conclusion that adjusted overalls were more grouped together than real Madden overalls, which were much more spread out. I did this by sorting by position group and plotting the points on top of each other in different colors. The graph shows how all of the top and bottom ratings, mostly above 90 and below 70, are the real Madden overall ratings while the adjusted overall ratings are much closer together. From there I thought a side by side histogram of both the real overalls and adjusted overalls could provide some insight and it reiterated my point of the greater spread with the real overalls. In Graph 6 it can be seen that the left graph looks more like a normal distribution, with points peaking around 65-70 overall with most points between 55-95. The adjusted overall on the right graph peaks around 75 overall but has a ton of points scattered between 65-90 and very few outside of that range.

After reviewing my analysis, I determined that there were some ethical concerns to be addressed. This data could be used to make accusations against the company EA for inflating Madden overalls. If this were public players could argue they missed out on endorsements due to their publicity not being the same as other players for perceived weaknesses in their game. EA could face a loss of their exclusive license with the NFL. It could be argued that this is a good thing since EA has a monopoly on NFL games and a player getting what he deserves is good, however the employees and stakeholders at EA would stand to lose a lot of money.

The final conclusions from this exploratory data analysis are that awareness is likely used by EA to influence Madden overalls to separate the star players from the above average players, as well as keeping the bad players lower than the average players. Awareness is not an attribute that is defined through athletic testing or on field play and EA took advantage of its blanket definition to tweak player overalls without having to inflate a player in false regards such as increasing speed. Future work could include figuring out how exactly what inflates awareness. I identified that total salary is correlated however that was it. It would be interesting to see if other factors such as social media following, or fantasy points could influence it. I am most proud of this project this semester as I had no prior experience graphing in Python and I picked it up fast and did well in my opinion.

**Index**

Graph 1: Chart, scatter chart

Description automatically generated

Graph 2:

Chart, scatter chart

Description automatically generated

Graph 3:

Chart, scatter chart

Description automatically generated

Graph 4:

Chart

Description automatically generated

Graph 5:

A picture containing text, pencil

Description automatically generated

Graph 6:

Chart, histogram

Description automatically generated