Sean Abernathy Self-Assesment

For our project, the team decided we would work on data surrounding March Madness, which is a collegiate basketball tournament played every year to crown a champion. We had data from all the regular seasons and March Madness tournaments from 2008 to 2022 (with 2020 excluded since the tourney was not played that year).

Our first step was determining what question(s) we wanted to answer. I put together a list and we ultimately decided our machine learning model would answer: Can we use this data to accurately predict the winner?

From there we cleaned up the data so we could use it by removing some columns, dropping any null values, and adding in a solution or y column for our neural network model. I was able to help with step by using Jupyter notebook and our knowledge on Pandas to clean up the data. Next, we looked to run our neural network model that spit out our accuracy score. We ended up at an accuracy score of 98.5% which is much higher than I anticipated due to my previous knowledge of the tournament. The other members of the team ultimately did the work here, however, I was involved with helping to figure out the code and make a decision on how many enochs, neurons and activation functions to use. Then, we worked on our visualizations in Tableau. I put together the story we have used for this project, but the whole team had input on what we wanted to visualize. Finally, we got to the presentation portion of the project. As a group we decided which parts we would discuss and then each individual made their slides.

Throughout the project, I thought the whole team was engaged and willing to put in the work to make us successful. We communicated very well via either Slack or over Zoom calls. Our team also seemed to be proficient in all of the analysis tools we used, such as Jupyter Notebook, Tableau, Pandas, GitHub, Scikit-Learn and SQL.

I don’t think there is anything I would change about our analysis, but I would be interested to dive into the data more to find any more trends or helpful insights as it is a topic that intrigues me. Also, I don’t know how to use our model to make a prediction or if that is even possible, but that would have been awesome to do for the 2023 tourney since it was going on while we were working. Our biggest pain point in the project felt like it was getting the data cleaned up and making a decision on how to proceed, but even that didn’t take us long.