ACC Basketball Homework Assignment Reflection

Ashley Nguyen

I learned the importance of thoroughly understanding and cleaning data before extracting and manipulating it. Working with CSV files and pandas dataframes, I encountered challenges like replacing headers, converting data types, and ensuring the data was organized correctly before I could analyze it. These experiences taught me that a strong foundation in understanding the structure of the data is crucial for any analysis. For future projects, I will prioritize grasping the framework and organization of the data table before moving forward with analyzing it.

The most challenging part of this assignment was learning how to effectively use pandas for filtering and extracting information as I’ve never used the pandas library before. I had to familiarize myself with pandas syntax, like using single brackets [] to select one column and double brackets [[]] to select two, and chaining multiple operations in one line of code like .groupby and .sort\_values. To overcome these challenges I read the pandas documentation and was able to understand how they worked. Additionally, plotting the data was difficult because I kept getting error codes. For the players\_500\_minutes I kept getting “'>' not supported between instances of 'str' and 'int'” and for correlation plot I kept getting “TypeError: 'value' must be an instance of str or bytes, not a float.” After some googling, I was able to figure out that the values in the data table were type string not integers. By using the pd.to\_numeric operation to change the values in the columns to integers, I was able to fix the error and move forward with my analysis.

While analyzing ACC basketball statistics, I was able to quickly pull distinct information from a huge data frame in order to answer questions of interest. Additionally, I was also able to identify and visualize key patterns, like correlations, that aren’t obvious a first glance. These techniques can be applied to different datasets, like healthcare or business, to visualize statistics and make meaning out of numbers. What might appear as a table of seemingly random array of numbers, data science can show that’s there more than what the naked eye can see.