Sahil Rajapkar
CIS3120 - Avinash Jairam
12/21/21

Homework 3

For this homework assignment, I personally found it to be less challenging because it reminded me of the content I have learned in CIS4400, a class that dealt with SQL, which focused on a concept called JOIN. Nevertheless, before I dive into how I was able to use "JOIN" (or merge) through Python programming on Google Colab, I'd like to point out that I will be continuing to use the NBA API and NBA-related data. I choose to do so for a few reasons such as developing my interest in basketball and technology as well as trying to aid others in solving a problem with my analysis. With that being said, I'd like to briefly speak about how I began and the processes I took to get better information and juices flowing for some accurate data frames.

So, as per usual, I installed the NBA API, then I went ahead to import pandas and began adding segments of my last assignment to this one. I did so until I reached a point where I had two frames. Then, I joined the two frames on the column, "full_name" which was the top 50 NBA career scoring leaders. After this, I got rid of unneeded columns and left columns such as total points, points per game, field goals made, and many more. Next, I had to combine the data frame from the start to this new data frame to get dataframe three.

One of the first things I did was immediately use the .describe() operation to get the description statistics of dataframe three. It came out with some stats involving the count, mean, standard deviation, min, max, and others as well. I found this to be something that we have gotten familiar with throughout this semester and along the homeworks and projects we've done but more importantly, what did I do next? I exported the result from dataframe three to a CSV file which concluded homework three for me. Next, I can talk about how this analyses can aid others problem-solving, as promised before.

I want to mention that in the sport of basketball, the landscape is rooted for its competitiveness. Coaches and organizations are constantly thrown with problems that can be both internal and external. But, one major problem is how do they go about solving these problems to help see their team succeed. Well, I am here to tell you that they can use analytics to do just that.

So, one problem that I have noticed in these games is that stakeholders of a basketball team are lacking the information to be confident in their processes, actions, and statements. For example, a coach may not know the exact average PPG of all the basketball players who lead in career points so, one thing, my analysis says is that their average (or mean) is 21.86 PPG. This is a very important number because it solves the problem or expectation of what the coach should expect to be the average amount of points their star player scores per game. Hence, they see a player's performance!