Sahil Rajapkar
CIS3120 - Avinash Jairam
10/23/21

Project 1

For Project 1, I found it overall to be very interesting. Some things that were relevant to homework one were incorporating spans of Python topics from web scraping to developing, using software such as Anaconda-Spyder, Sublime, and Google Colab, making use of if statements, for loops, and so on. Hence, I found the overall project to be a bit repetitive, in a good way, but, the only twist to this project was having roster's from other colleges for both the swimming and volleyball teams as well as incorporating the ideology and methodology of creating a data frame.

Now, when I went about starting this project, but more importantly the questions, I continued what Theo had told me at the beginning which was to just stick with one sport, one gender, and one set of rosters. For example, upon creating a data frame for just the men's swimming team for Brooklyn College, Baruch College, and John Jay, I then found the average of all those players from the data frame. But, before even doing this, I created a data frame for each university and then eventually merged them to create one big data frame. From this data frame, I found the average. But, one thing I learned was, aside from the average, you can also get more information such as the "head" of your data which was, in this case, the tallest of athletes, and the "tail" of your data which was the shortest of athletes.

Once again, I stuck to the method of only coding for one distinct area or group of individuals and colleges. From doing this, I simply used .nlargest and .nsmallest to find the top five tallest men's swimmers and the top five smallest's men's swimmers. This gave me two answers whereas, in the end, I came out with a total of eight (four tallest athletes and four smallest athletes) along with four averages.

If we go back to homework one, we can remember that a big insight was "that the average of both the men's and women's volleyball teams would have a greater average height than both the swim teams." Again, we can see that this statement holds. For all the rosters across the many universities, we came out with the following averages: (1) the average height of 39 men's volleyball players was 72 inches (2) the average height of 38 women's volleyball players was 67 inches (3) the average height of 37 men's swimmers was 71 inches and (4) the average height of 34 women's swimmers was 65 inches. So, though not a significant difference, volleyball players on average remain taller than swimmers. Hence, to answer question seven, the average swimmer is not taller than the average volleyball player (as previously mentioned in homework 1). Though true, I was able to continue my findings with a conclusion that gave me some key takeaways for my programming style, the overall thought process in getting successful results, and much more!

In conclusion, to restate some points from the beginning, I used web scraping, tags, classes, list comprehensions, for loops, apps such as Sublime Text and Anaconda-Spyder, all to help grow my knowledge and expertise with Python, But, with that beings said, I was able to learn how to create and manipulate a data frame, starting with one distinct group then moving on to many more. I also learned how to merge data frames and get useful information out of it such as the average. Lastly, one huge takeaway for me is to always seek for insight which in this case was between athletes.