Name: Andrea Fox

Date: August 4, 2021

Course: DSC680 – T301 Applied Data Science

Assignment: Project 3 Check-in

1. Any surprises from your domain from these data?

The biggest surprise I’ve had so far was the data files themselves. When I went to do my cleaning, I printed out a list of columns which came out to 8 and all were necessary in my analysis, so no clean up there. I then went to look for null and missing values and found zero in all four of my files. I triple checked this because I didn’t believe it. This meant I really didn’t have to spend ton of time on cleaning up my data for analysis, which was a nice surprise.

1. The dataset is what you thought it was?

The dataset is exactly what I thought it was going to be. It was also very easy to combine all four csv files into one dataset after making a few minor updates. Really the biggest thing I needed to do was add an additional column to all four individual datasets to include type. That way when I combined them, I still knew which wine name and rating went with which wine type.

1. Have you had to adjust your approach or research questions?

I have not had to adjust my research questions although I haven’t fully answered them yet in my analysis thus far. As for my approach I think I am going to have to change it. I continuously forget to run a dataset through the different methods it must be in integer format, which most of mine is not. I can’t even change my values because it wouldn’t make sense to change them to integers, so I had mentioned I wanted to investigate machine learning but not entirely sure that will be possible.

1. Is your method working?

My visual analysis method has worked so far. I was able to look at rating by various methods such as location, price, and wine type. As stated above I may have to rethink how I approach my technical analysis, but I have a couple of references that might be able to help steer me in the right direction.

1. What challenges are you having?

So far, no challenges necessarily. I may have to change my technical methods, but the jury is still out on that for now. This felt like the right time to do a check-in since I was in between my visual and technical analysis.