Jessica Owens

DSC530

Final Paper

The overall objective was to find a commonality as to why customers are disconnecting in their first year of service. The data set I used has several variables that I was able to explore. I felt that the most important variables were the tenure, disconnect reason and speed. It was interesting to see that so many customers disconnect within the first month and that monthly costs do not seem to have a large impact on disconnects. If I were to continue analyzing this, I think that this would be something that may be good to explore.

My initial assumptions about the disconnects was to focus on monthly costs and speed. While the monthly cost variable doesn’t seem to be much of a contributor, the speed does. Per the heat map, the correlation between speed and disconnect reason has the most impact for disconnects. Another eye-opening data discovery was that many of the disconnect reasons were “other”. Many of these were for customers who obtained services and disconnected services in one month which makes me wonder if there is a data discrepancy.

One of the challenges I faced is that I’m not experienced with completing an analysis from start to finish. I found myself duplicating efforts and having to re-create things due to different packages, and just not having enough knowledge yet. I feel that I may have limited the outcome due to the data set I chose. I felt that I had a very good grasp of the data by the end, but I struggled to keep a holistic view of the data. I continued going down rabbit holes that didn’t equate too much. While this course was wonderful, and I feel that I’ve got many of the concepts understood, I still lack the ability to put all the concepts together in an informative way. I believe as I continue learning and working with python, that I’ll get a better understanding of how all the concepts relate and when to use them.