**DSC 540 Final Project Summary**

**Lara Clasen**

*A 250-word paper summarizing your steps and any challenges you ran into during the project. Discuss the importance and relevance of this type of process if you were a data scientist. How often do you think you would have to do this to get the data you need?*

The final project most certainly challenged me to bring together much of what we have learned this term! After several hours of searching for an appropriate source to connect with for the project, I ended up putting together multiple, smaller projects into one. I had a really hard time finding sources that had somewhat organized data, an accessible API or website for scraping, as well as 15 or more variables to choose from. Originally, I had wanted to pull data from a movie website such as IMDB but it proved to be a bit difficult for my level in Python. Thankfully there are resources across the web to help me through it. Because that piece only included5 variables, I knew I needed to complete more than one example for my final project. The most difficult thing I ran into for this was finding a usable site with more than 15 variables. There are so many interesting data sets that I came across but they fell short in that respect. I know that more than anything I will need to understand more about how to recognize all of the various sources and types that are available to us in order to begin accessing data. This will be of the utmost importance in a data science career since I know we will be met with all kinds of data, from big to small and from neat to sloppy, and it will be up to us to know where to start. This project helped me learn a lot about my strengths and weaknesses, and reminded me of why I am so interested in this career – because after struggling and investigating, it feels so good once you finally get things to work and find something interesting that you couldn’t readily see before!