Program Evaluation

Program objectives:

1. Prepare datasets for solving problems.
2. Recommend appropriate data modeling techniques to test hypotheses.
3. Communicate data science results into answers for domain challenges.
4. Identify ethical considerations in dataset preparation and modeling.

Evaluation:

1. This was the easiest objective for me to master. While data preparation does take the most time out of a data science project, I quickly learned how to identify what needs to be done to each dataset to "clean" it. I was also able to identify what needs to be done to a dataset depending on the data types (concatenating strings, removing missing data, etc).
2. Each data science problem requires a dataset, and each dataset has a target variable whose values are modeled and predicted by the other variables. Each target variable lends itself towards one type of modeling technique, and the various projects taught me how to easily identify what models need to be used for a given dataset. For example, regression is used on numerical target variables, but not all types of regression are suitable for the same problem. I feel like I have mastered this objective quickly as well.
3. This objective took me the longest to master. The results of a data science project generally are an evaluation of a data model but translating those results into answers was challenging. I learned how to effectively communicate the results of my projects in a way that answers the questions posed by the business problem.
4. Identifying ethical considerations is something that I came into the program knowing how to do. However, the program increased my ability to locate ethical problems for more specific issues. For example, movie prediction is a niche issue with ethical concerns that weren't immediately evident, but I was able to locate an ethical consideration for it.

In the end, I feel that the program did a great job of preparing me for a future data science career. I mastered all four objectives it set out to teach me, and I learned a lot about how a data science project is developed from start to finish. However, the program wasn't without its weak points. For one, it could have been stronger in teaching us more machine learning projects in R. I feel like 80% of our projects were done in Python, whereas having a background in both Python and R would increase our chances of landing a data science career in the future. A suggestion I have for improving this is to have a class where a project can be done in both programming languages. Nonetheless, with the data science and programming knowledge I've gained, I've made a plan for after I graduate from this MSDS program. I plan to switch my career from biotech to data science while taking a course in AWS to improve m chances of landing a job in the industry.