Data Science Career Track

Capstone: EDA Rubric

Learning Objectives

- Understand the importance of performing EDA on data science projects.
- Apply data wrangling techniques, as laid out in the DSM building data profiles, tables, and figures to evaluate the feature relationships.
- Identify the features that are likely to have the most impact in modeling based on relationships between the features and the response variable.

Criteria	Meets Expectations
Completion	Every feature is investigated using either histograms, bi-plots, or other visual and numeric means.
	Pearson correlation coefficients and other statistical methods were used to identify statistical relationship strengths.
Process and understanding	☐ The submission shows that the student understands how to explore feature relationships in the data.
	☐ The submission demonstrates that the student made data-supported decisions on when to select specific features.
Presentation	Jupyter Notebook with all the applied code steps in working order and with notation or comments as needed.
	☐ The submission is complete and uploaded in full to the Github repo.