

Data Science Career Track

Capstone: Data Wrangling Rubric

Learning Objective

- Clean the dataset and address issues like missing values and duplicate values.
- Apply data wrangling techniques, as laid out in the DSM of; Data Collection, Data Organization, Data Definition, Data Cleaning.

Criteria	Meets Expectations
Completion	<ul style="list-style-type: none"><input type="checkbox"/> All data wrangling steps required for their dataset are applied.<input type="checkbox"/> Data and code are uploaded to Github as a Jupyter notebook or Python script files.
Process and understanding	<ul style="list-style-type: none"><input type="checkbox"/> The submission shows that the student understands how to acquire, organize, define and clean data.<input type="checkbox"/> The submission includes data sets that were well-chosen and relevant to the problem<input type="checkbox"/> The submission demonstrates that the student made data-supported decisions on how to handle missing values, duplicates and outliers as needed
Presentation	<ul style="list-style-type: none"><input type="checkbox"/> Jupyter Notebook with all the applied code steps in working order and with notation or comments as needed.<input type="checkbox"/> The submission is complete and uploaded in full to the Github repo.