

PROJECT SPECIFICATION

Data Engineering Capstone

Write Up

CRITERIA	MEETS SPECIFICATIONS
Scoping the Project	The write up includes an outline of the steps taken in the project. The purpose of the final data model is made explicit.
Addressing Other Scenarios	The write up describes a logical approach to this project under the following scenarios: <ul style="list-style-type: none"><li>The data was increased by 100x.</li><li>The pipelines would be run on a daily basis by 7 am every day.</li><li>The database needed to be accessed by 100+ people.</li></ul>
Defending Decisions	The choice of tools, technologies, and data model are justified well.

Execution

CRITERIA	MEETS SPECIFICATIONS
Project code is clean and modular.	All coding scripts have an intuitive, easy-to-follow structure with code separated into logical functions. Naming for variables and functions follows the PEP8 style guidelines. The code should run without errors.
Quality Checks	The project includes at least two data quality checks.
Data Model	<ul style="list-style-type: none"><li>The ETL processes result in the data model outlined in the write-up.</li><li>A data dictionary for the final data model is included.</li><li>The data model is appropriate for the identified purpose.</li></ul>
Datasets	The project includes: <ul style="list-style-type: none"><li>At least 2 data sources</li><li>More than 1 million lines of data.</li><li>At least two data sources/formats (csv, api, json)</li></ul>

Suggestions to Make Your Project Stand Out!

To make your project stand out:

- Work with large amounts of data.
- Combine datasets that are difficult to combine.
- Enrich the data from several disparate sources.
- Include recommendations for how to use to data to come up with insights.
- Write a blog post about your project and link to it.