

Senior Data Engineer Candidate Take Home Test

Overview

The focus of this take home is around ingesting a CSV file which contains a mock example of a Trustpilot Review dataset. The goal is loading the data into a database (of your choice) to be used downstream by the rest of the business via an API. Requirements:

1. Ingest the Reviews CSV file provided into a table in a database.
2. Programmatically clean up the data if needed.
3. Create an API that enables our users to perform CRUD operations.
4. Submit your solution via **Github**.

What Are We Looking For?

We want to give you the opportunity to show off your software engineering skills and coding best practices.

Examples of areas we care about:

- Unit testing
- Readability (Code comments, etc)
- Thought Process
- Design Choices
- Performance Considerations
- Productionisation
- Documentation/knowledge sharing
- Cost

The submission is **language agnostic** so please pick the language that you feel most comfortable with, we use a selection of languages (Python, Typescript, Java) and we are open for you to choose other languages too.

Be prepared to discuss your solution and how you could extend it, how you may run the application in a production environment and also how you could replicate the data to a datalake of your choice

How Long Should It Take?

This should take somewhere between 2 and 4 hours. Please be prepared to explain and discuss your solution in the technical interview.