

ETL Pipeline Exercise

Task

Build a small ETL pipeline in **Python** that:

- extracts data from the [WHO GH OData API](#),
- transforms the data in any reasonable way,
- and loads it into a **PostgreSQL** database.

You're free to decide what data to extract and how to model it, as long as the result could plausibly support analysis.

What we're evaluating

We're mainly interested in:

- your backend / integration thinking,
 - how you structure and explain your code,
 - how you reason about data transformation and correctness,
 - how you think about the system as a whole (not just “making it work”).
-

Requirements

- Python backend
 - PostgreSQL
 - Data validation
 - Handling edge cases
 - Explain your approach to testing and debugging
-

Bonus (optional — not required)

- Ability to stop and resume the pipeline
- Re-running the pipeline while only fetching new data
- Being able to resume the **extract** step without starting from scratch
- Clear setup instructions (README)
- Any tests (unit or functional)
- Notes on what you would improve with more time