

Take-home test for Data Engineer

You will be building a small and simple ETL+API system. **All should be done using Python 3.10**

Tasks:

- Create a scheduled task running every 10 seconds simulating an ETL that will:
 - Load one of the 4 provided data files (included in the `data` folder)
 - Clean and process the data
 - Add the data to a simple SQLite db
- Create a REST API using the SQLite db to provide functionality to serve the following types of queries:
 - Movies between two years
 - Movies from a specific genre
 - Best rated director (in average)
 - Movies from a specific director

Dataset columns:

- **Movies** Movie names
- **Genre** Movie genre
- **Year** Release year of the movie
- **Rating** Audience rating for the movie
- **One-line** Short description about the movie
- **Stars** The movie casting
- **Votes** Number of rating votes for the movie

Key points to consider

- Efficient code structure
- Clean and easy-to-follow code
- Make sure the data in the SQLite database is in good form/quality
- Perform the appropriate checks in the API requests

Submission

Please provide us a zip file with all the files needed and a small document explaining how to run it.