Stock Data Web Scrapper ETL Project

Case-study:

Jazzy Investment, one of our clients is a stockbroking firm that deals in issuance and trading of stocks on behalf of its customers in the Nigeria Stock Exchange. For the analytics team to analyze market trends and place the best bet on company stocks, they require the daily stock exchange data published on here: https://afx.kwayisi.org/ngx/

Your task:

As part of your data engineering training program, you're required to build a web scrapper (data pipeline) to Extract, Transform and load the listed companies/securities data from https://afx.kwayisi.org/ngx/ to a Postgresql database.

The data required is under the heading **Listed companies/securities**

This data will be consumed by our client (Jazzy Investment) in the database for their stock trading analytics.

Requirements:

1. Assuming the data of company stocks listed is updated on the website every day and your pipeline will run once every day, ensure newly extracted data is integrated with existing data in the database each time your pipeline run.

2. The first page of the website shows only 100 of 157 listings (click on the Next>> button at the bottom of the table) to confirm this. You are expected to get data for the remaining 57 listing and add to the first 100 list.

3. Your transformation should include a column called **Date** which shows the current date for which the data was scrapped.

4. For companies that does not have a value for volume traded for a particular day, then the value for volume should be the mean volume of all companies that traded that day.

Submission:

You're required to submit one python file (e.g etl.py).