Data Engineer Assessment

The goal of this test is to assess your data engineering and architectural skills.

Requirements

The data attached consists of daily stock trading data for a company, say MSFT from 2017 to 2022.

1. Schema
    - Design an RDBMS table schema to store the CSV data
2. Calculation You are required to use this data and calculate the following using SQL
    - Weekly average of High, Low and Volume
    - Monthly average of High, Low and Volume
    - Quarterly average of High, Low and Volume
    - Yearly average of High, Low and Volume
3. System Design
    - Design a Streaming pipeline where the source will be stock trading data for multiple stocks fed every second. The job of the pipeline is to calculate rolling average of trading data for each stock for the past 20 days, 50 days, and 200 days.
    - The calculations have to be persisted on a database of your choice.
    - The calculated data will be used by APIs on the Frontend side. The API should be able to return
        - the rolling average/averages of a given stock
        - whether a stock is above or below a rolling average
    - You can make assumptions on the incoming and outgoing data formats.
    - The pipeline has to be reliable and robust enough to take care of missing data.
    - You are expected to provide a detailed architecture of the pipeline with justification for the architecture
    - You are also expected to provide the different technologies chosen for each stage of the pipeline, and explain the reason for choosing those over the alternatives.
    - You do not need to provide end to end code for this section, but you should provide and be prepared to explain the relevant code snippets used in different stages of the pipelines

Evaluation criteria

- Clean code
- Meets the requirements
- Explain the assumptions and choices made in the System design

The following earn you bonus points:

- Clear, concise git history
- User friendly README

Submission Guidelines

- Please include a Readme.md file with instructions to run
- Submit as a zip/tarball