Sample data: Please refer to the attached files — user_profiles.csv (sample user profile data) and user_events_20231026.json (sample user events data).

Scenario and Problem:
You are working on the analytics layer of an online platform and are responsible for collecting, processing, and preparing user activity data from various sources for analysis by product and marketing teams.

The data comes from two main sources:

1. **User profile data (user_profiles.csv):**

- Contains basic and relatively static information about users.
- General structure: user_id, name, registration_date, location
- This file represents a source that provides profile information (sample file attached).

2. **User event data (user_events_YYYYMMDD.json):**

- Contains real-time user activity events (such as page visits, button clicks, purchases, etc.).
- This data is incremental and usually generated in separate files for each day.
- General structure of each event: user_id, event_type, timestamp, details
- The details field can contain different information (and event_type has a nested and variable structure which may sometimes be empty). (Sample file attached)

Your goal:
Design and implement an ETL/ELT data pipeline that:

- Loads data from input files — JSON for events and CSV for profiles.
- Joins the event and profile data based on user_id.
- Extracts key information needed for analysis from the details field (such as page_url, button_id, item_id, etc.).
- Converts date and time fields from the timestamp field into appropriate formats (e.g., extracting event_date from timestamp).
- Saves the processed data in an optimized format such as Parquet, with possible partitioning (e.g., by event_date) to support analytical querying.

Expected output data:
A consolidated dataset including the following fields (and any other fields extracted from details you find useful):

- event_id (a unique event identifier — you can generate this)
- user_id
- name (from profile)
- location (from profile)
- registration_date (from profile)
- event_type
- timestamp (event timestamp)
- event_date (date extracted from timestamp)
- details_raw (the original details field as string or JSON)
- page_url (if present in details)
- button_id (if present in details)
- item_id (if present in details)