Hevo Exercise - Assessment I

This exercise is designed to assess your ability to set up data pipelines, apply transformations, and deliver results in Snowflake using **Hevo Data**. The instructions are intentionally left somewhat vague, candidates are expected to research and resolve gaps using documentation and their own problem-solving skills.

Prerequisites (candidate must have)

- Access to GitHub (public repo for deliverables)
- Basic database knowledge and SQL (Postgres)
- Snowflake free trial account
- Hevo free trial (via **Snowflake Partner Connect**, recommended)
- Ability to run a local PostgreSQL instance (Docker suggested, but any setup is fine)
- Basic networking knowledge to expose local DB to Hevo

Schema & ER Diagram

Postgres source schema (local database)

customers

- id (PK, integer)
- first_name (varchar)

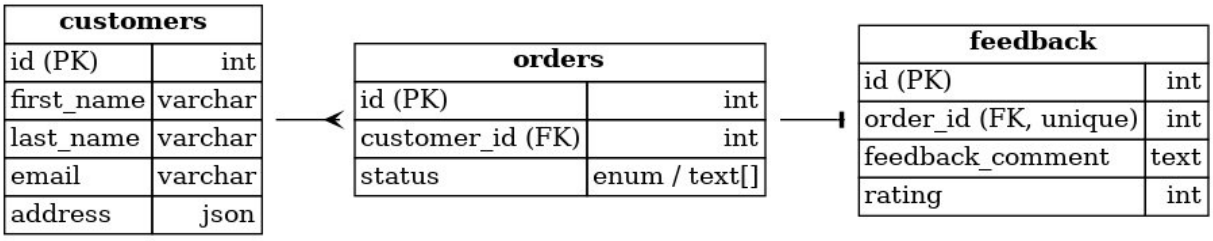
- last_name (varchar)
- email (varchar)
- address (json)

orders

- id (PK, integer)
- customer_id (FK → customers.id, integer)
- status (enum or varchar; candidates may use Postgres ENUM or simple varchar)

feedback

- id (PK, integer)
- order_id (FK → orders.id, integer, **unique if one-to-one**)
- feedback_comment (text)
- rating (integer)



Objectives

1. **Sign up for Snowflake**

   Create a free Snowflake trial account. (You will use this Snowflake account as the pipeline destination.)

2. **Sign up for Hevo via Snowflake Partner Connect**

   Use Snowflake Partner Connect → Hevo to get the Hevo trial account

3. **Set up Postgres**

Install a PostgreSQL database locally (Docker or VM).

4. **Create tables & load data**

   Create the three tables (customers, orders, feedback) as defined in the schema and load them with the provided CSVs.
   CSV link: /hevo-assessment-csv

5. **Set up Hevo pipeline**:

   - Source = PostgreSQL (from Step 3)
   - Destination = Snowflake
   - Ingestion mode must be **Logical Replication**

6. **Apply Transformations**:

   - **Orders table → Event table**
     Based on the status field, create a new table in Snowflake (e.g. order_events) where each row represents an event.
     Example:
     - If an order has status = 'delivered', insert a row into order_events with event_type = 'order_delivered'.
     - Similarly, handle placed, shipped, cancelled.
   - **Customers table → Username field**
     Add a derived field username to the customers data, extracted from the email address (text before the @).
     Example: jane.doe@gmail.com → jane.doe.
   - If required, the following can reference Hevo documentation for Transformations.

7. **Load data to Snowflake**:

   Let Hevo push transformed data into your Snowflake trial account.

8. **Document & Share**:

   Push everything to a **public GitHub repo** including:

   - README.md with:
     - Detailed steps to reproduce
     - Assumptions made (e.g. enum vs varchar for status)
     - How Postgres was connected to Hevo
     - Choices made for transformations

        - Any issues/workarounds during setup
    - SQL DDL files for Postgres tables
    - CSV files (or link to the provided CSV repo)
    - Transformation scripts (code or drag and drop SS) used in Hevo
    - A script or SQL file with validation queries for Snowflake (e.g. count rows, check usernames, confirm event rows)
    - Loom video link (or instructions to add) showing pipeline setup, transformations, transformation test via Hevo UI, and verification in Snowflake

Deliverables (to be attached to the google form)

- GitHub repo link
- Hevo **Account Team Name** and **Pipeline ID**

(All documentation, assumptions, and notes must live inside the README.md file in your repo.)

Important Notes

- Do not hardcode or publish any credentials, database info, or access keys in the repo. Use environment variables, config files outside source control, or Hevo's UI settings.
- Please follow the academic honor code while doing the exercise. Candidates will support this culture of academic honesty by neither giving nor accepting unpermitted academic aid in any work that serves as a component of grading or evaluation, including assignments, examinations, and research.
- Any violations of the above honor code will result in immediate disqualification.