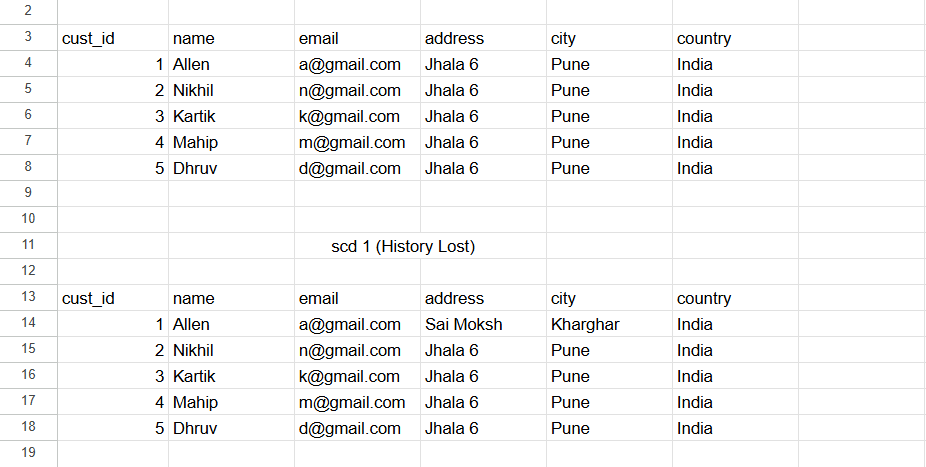
### **Day 3 – [ETL Concepts And Architecture]**

#### **1. What I Learned**

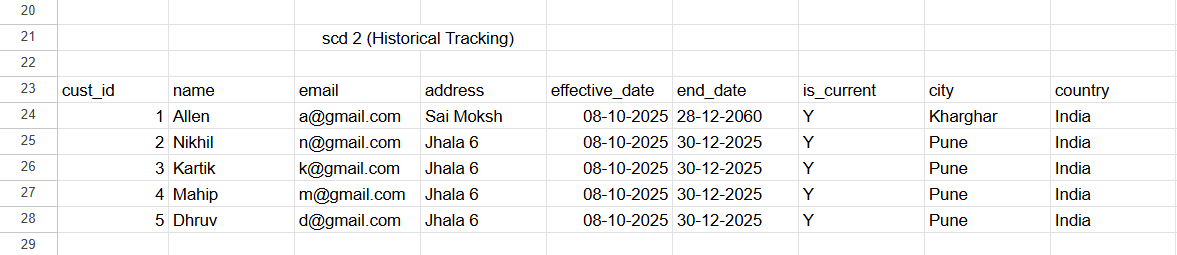
* Revised concepts of ELT vs ETL, Data Lakes vs Data Warehouse, and Lakehouse architecture.
* Performed manual ETL exercises implementing SCD Type 1 and Type 2 for data updates and history tracking.
* Created a detailed flowchart illustrating all ETL process steps from data extraction to loading.

#### **2. Code Snippets / Screenshots**

**Exercise 1**: Manual ETL with SCD Type 1 (Update in Place)



**Exercise 2**: Manual ETL with SCD Type 2 (Historical Tracking)



**Exercise 3**: Identify SCD Types (Theory Exercise)

For each case, decide which SCD type is appropriate:

**1. Employee salary (needs audit history)**  
**Ans.** **SCD Type 2** : Salary changes require complete audit trail for legal purposes.  
 Organizations need to track salary changes and what the  
 previous values were for auditing.

**2. Product description (only current matters)**  
**Ans. SCD Type 1**: Product descriptions historical value doesn’t benefit in any business  
 intelligence.

**3. Customer phone number (keep current + one previous)**  
**Ans.** **SCD Type 3:** Limited history is required for business continuity

**4. Vendor tax ID (should never change)  
Ans. SCD Type 0:** Tax ID is a legal identifier which should remain constant

**5. Store operating hours (track all changes)  
Ans. SCD Type 2:** Operating hours affects business analytics. Full history enables  
 analysis of how hours changes impact on sales.

**Exercise 4**: Manual ETL Simulation

