*"I need to generate a large-scale dataset for Swiggy sales analysis using Python. The dataset should include* ***1,500,000 orders*** *and* ***1,500,000 deliveries****, ensuring proper relationships between tables. The dataset consists of five tables:* ***Customers, Restaurants, Riders, Orders, and Deliveries****. The relationships between tables should be maintained through primary and foreign keys. The dataset should follow these specifications:*

**Tables and Relationships:**

**1. Customers Table:**

* customer\_id (Primary Key)
* customer\_name (Indian names using Faker)
* age (18-65)
* gender (Male/Female)
* registration\_date (Date range: **Last 3 years**)

**2. Restaurants Table:**

* restaurant\_id (Primary Key)
* restaurant\_name (Indian names using Faker)
* city (Cities including **Daltonganj, Indore, Ranchi**)
* opening\_hours (Random business hours between **6:00 AM to 11:00 PM**)

**3. Riders Table:**

* rider\_id (Primary Key)
* rider\_name (Indian names using Faker)
* sign\_up (Date range: **Last 2 years**)

**4. Orders Table:**

* order\_id (Primary Key)
* customer\_id (Foreign Key → Customers)
* restaurant\_id (Foreign Key → Restaurants)
* order\_item (50 food items)
* order\_date (**Date range: Last 1 year**)
* order\_time (Random time of the day)
* order\_status (**Completed, Pending, Cancelled**)
* total\_amount (₹100-₹2000)
* rating (Random 1-5)
* **Orders should have a mix of completed, pending, and canceled statuses**

**5. Deliveries Table:**

* delivery\_id (Primary Key)
* order\_id (Foreign Key → Orders)
* delivery\_status (**Delivered, Failed, In Transit**)
* delivery\_time (Random time of the day)
* rider\_id (Foreign Key → Riders)
* **Only Completed orders should have a corresponding delivery**
* **Canceled orders should NOT be in the Deliveries table**

**Additional Requirements:**

✅ **Ensure relationships between tables using primary and foreign keys**  
✅ **Every order should belong to a valid customer and restaurant**  
✅ **Each delivery should be linked to an order and a rider**  
✅ **1,500,000 orders = 1,500,000 deliveries (excluding canceled orders)**  
✅ **Data should be generated using Python (Faker + Pandas) and exported as CSV files**

**Final Deliverables:**

The dataset should be saved as:  
📂 customers.csv  
📂 restaurants.csv  
📂 riders.csv  
📂 orders.csv  
📂 deliveries.csv

This dataset should be structured to reflect real-world food delivery operations. 🚀