### **Documentation for the Healthcare Datasets**

This document provides a detailed description of the datasets generated for the healthcare use case. Each dataset is designed to simulate realistic data for testing and analysis purposes.

### **Directory Structure**

output/

└── healthcare\_MM-DD/

├── healthcare\_medical\_staff\_MM-DD.csv

├── healthcare\_patients\_MM-DD.csv

├── healthcare\_appointments\_MM-DD.csv

├── healthcare\_hospital\_performance\_MM-DD.csv

├── healthcare\_pharmaceutical\_inventory\_MM-DD.csv

├── healthcare\_billing\_and\_payments\_MM-DD.csv

├── healthcare\_regulatory\_compliance\_MM-DD.csv

├── healthcare\_patient\_outcomes\_MM-DD.csv

├── healthcare\_supply\_chain\_MM-DD.csv

### **Example Prompts**

#### **1. DBT Transformations**

* "Write me dbt code to calculate the average wait time for appointments by department."
* "Write me dbt code to identify patients with multiple missed appointments in the last 6 months."
* "Write me dbt code to create a table that tracks inventory items nearing their expiration date."

#### **2. Lists** [**Correct outputs: healthcare dataset**](https://docs.google.com/spreadsheets/u/0/d/13kCCNi74eDK0MOjss7KUTPOATGwKry_NbX-UuJfnskA/edit)

* "List all patients who have no medical history of chronic conditions." (15 patients)
* "List all staff members working in the Emergency department with a rotating shift schedule." (10 staff)
* "List all appointments scheduled for the next 7 days, grouped by department." (62 appointments)

#### **3. Visualizations**

* "Visualize the trend of bed occupancy rates over the past 30 days."
* "Visualize the distribution of billing statuses (e.g., Paid, Pending, Overdue)."
* "Visualize the frequency of compliance audit findings by department."

### **Dataset Descriptions**

#### **1. Medical Staff**

* **Description:** Contains information about healthcare staff, including their roles, departments, and contact details.
* **Columns:**
  + staff\_id: Unique identifier for each staff member (e.g., STF001).
  + name: Full name of the staff member.
  + role: Role of the staff member (e.g., Doctor, Nurse, Admin).
  + specialization: Area of expertise or specialization (e.g., Cardiologist, Emergency Nurse).
  + department: Department where the staff member works (e.g., Cardiology, Surgery).
  + phone: Phone number of the staff member.
  + email: Email address of the staff member.
  + hire\_date: Date the staff member was hired.
  + shift\_schedule: Work schedule of the staff member (e.g., Morning, Night, Rotating).

#### **2. Patients**

* **Description:** Contains demographic and medical information about patients.
* **Columns:**
  + patient\_id: Unique identifier for each patient (e.g., PAT001).
  + name: Full name of the patient.
  + age: Age of the patient (18–90).
  + gender: Gender of the patient (M or F).
  + address: Residential address of the patient.
  + phone: Phone number of the patient.
  + email: Email address of the patient.
  + date\_of\_birth: Date of birth of the patient.
  + medical\_history: Medical conditions or history (e.g., Hypertension, None).
  + allergies: Known allergies (e.g., Penicillin, None).
  + primary\_physician: Identifier of the patient's primary physician (e.g., STF001).
  + insurance\_provider: Name of the patient's insurance provider.

#### **3. Appointments**

* **Description:** Tracks patient appointments, including scheduling, status, and reasons for visits.
* **Columns:**
  + appointment\_id: Unique identifier for each appointment (e.g., APT001).
  + patient\_id: Identifier of the patient who scheduled the appointment.
  + doctor\_id: Identifier of the doctor assigned to the appointment.
  + appointment\_date: Date of the appointment.
  + appointment\_time: Time of the appointment.
  + status: Status of the appointment (Scheduled, Completed, Cancelled, No-show).
  + reason\_for\_visit: Reason for the appointment (e.g., Regular Checkup, Emergency).
  + follow\_up\_required: Indicates if a follow-up appointment is needed (True or False).

#### **4. Hospital Performance**

* **Description:** Contains daily operational metrics for the hospital.
* **Columns:**
  + metric\_id: Unique identifier for each performance metric (e.g., MET001).
  + date: Date of the recorded metric.
  + bed\_occupancy\_rate: Percentage of occupied hospital beds (0.60–0.95).
  + average\_wait\_time\_minutes: Average patient wait time in minutes (15–120).
  + staff\_efficiency\_score: Efficiency score of the staff (0.70–0.98).

#### **5. Pharmaceutical Inventory**

* **Description:** Tracks inventory of medications, equipment, and supplies.
* **Columns:**
  + item\_id: Unique identifier for each inventory item (e.g., ITM001).
  + item\_name: Name of the inventory item (e.g., Paracetamol Tablet).
  + category: Category of the item (Medicine, Equipment, Supplies).
  + quantity\_in\_stock: Current stock level of the item.
  + reorder\_level: Minimum stock level before reordering.
  + supplier\_id: Identifier of the supplier (e.g., SUP001).
  + expiration\_date: Expiration date of the item.
  + unit\_price: Price per unit of the item.

#### **6. Billing and Payments**

* **Description:** Tracks financial transactions and insurance claims for healthcare services.
* **Columns:**
  + billing\_id: Unique identifier for each billing record (e.g., BIL001).
  + patient\_id: Identifier of the patient associated with the bill.
  + appointment\_id: Identifier of the appointment related to the bill.
  + total\_amount: Total amount billed.
  + insurance\_covered\_amount: Amount covered by insurance.
  + out\_of\_pocket\_amount: Amount paid by the patient.
  + payment\_status: Status of the payment (Paid, Pending, Overdue).
  + payment\_date: Date of payment.

#### **7. Regulatory Compliance**

* **Description:** Tracks compliance audits and corrective actions for hospital departments.
* **Columns:**
  + compliance\_id: Unique identifier for each compliance record (e.g., COM001).
  + audit\_date: Date of the compliance audit.
  + department: Department audited (e.g., Emergency, Pharmacy).
  + audit\_type: Type of audit conducted (e.g., Safety, Infection Control).
  + findings: Summary of audit findings.
  + corrective\_actions: Actions taken to address findings.
  + status: Status of the compliance issue (Open, Closed, In Progress).

#### **8. Patient Outcomes**

* **Description:** Tracks treatment outcomes and progress for patients.
* **Columns:**
  + outcome\_id: Unique identifier for each outcome record (e.g., OUT001).
  + patient\_id: Identifier of the patient associated with the outcome.
  + treatment\_id: Identifier of the treatment plan (e.g., TRT001).
  + diagnosis: Medical condition diagnosed (e.g., Hypertension).
  + treatment\_start\_date: Start date of the treatment.
  + treatment\_end\_date: End date of the treatment.
  + outcome: Result of the treatment (Recovered, Improved, No Change, Deteriorated).
  + notes: Additional notes or observations.

#### **9. Supply Chain**

* **Description:** Tracks inventory orders and delivery statuses.
* **Columns:**
  + order\_id: Unique identifier for each order (e.g., ORD001).
  + supplier\_id: Identifier of the supplier (e.g., SUP001).
  + item\_id: Identifier of the inventory item (e.g., ITM001).
  + order\_date: Date the order was placed.
  + delivery\_date: Expected delivery date.
  + quantity\_ordered: Quantity of the item ordered.
  + quantity\_received: Quantity of the item received.
  + status: Status of the order (Ordered, In Transit, Delivered, Partially Delivered).