**Faculty of Computers and Artificial Intelligence**

**Data Warehousing – May 2025**

**Hospital Data Warehouse Project**

**1. Team Members**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Name** | **ID** | **Group Number** |
| **1-** | **Abdelrahman Mostafa Sayed** | **20220197** | **3IS-S7** |
| **2-** | **Mariam Assem** | **20221142** | **3IS-S2** |
| **3-** | **Hager Omar** | **20221190** | **3IS-All** |
| **4-** | **Reem Ahmed Abdallah** | **20221063** | **3IS-S2** |
| **5-** | **Samar Mahmoud** | **20221076** | **3IS-S2** |

**2. Physical model of the source system:**

A computer screen shot of a computer

AI-generated content may be incorrect.

**3a. Define the business processes that you will model:**

**1. Fact\_Pharmacy – Medication Dispensing Process:**

* **Tracks the distribution and usage of prescribed medicine.**
* **KPIs:**
  + - **Total quantity dispensed by medicine type and year**
    - **Most prescribed medicines per year**
    - **Unique patient count per medicine**

**2. Fact\_Appointments – Patient Visit Process**

* **Tracks patient visits and interactions with doctors.**
* **KPIs:**
  + - **Appointment count per patient per year**
    - **Average appointments per doctor per year**
    - **Top doctors by number of appointments**

**3. Fact\_Billing – Revenue Collection Process**

* **Captures all billing transactions and financial flow.**
* **KPIs:**
  + - **Average billing per patient per year**
    - **Total revenue by year**
    - **Monthly revenue trends**
    - **Top 5 patients by billing**

**4. Fact\_Cleaning\_Service – Facility Hygiene Management Process**

* **Tracks room cleaning activities and workload per staff.**
* **KPIs:**
  + - **Avg cleaning services per staff per room per year**
    - **Monthly cleaning services per room**
    - **Avg daily services per staff**
    - **Total services by room type**

**3b. Declare the grain of each fact table:**

**The grain of each fact table is defined at the most detailed (minimum) level of the respective business process. This minimum grain allows flexible aggregation and supports detailed analysis across multiple dimensions.**

|  |  |
| --- | --- |
| Fact Table | Grain (One row represents...) |
| Fact\_Pharmacy | **One medication dispensed to a patient on a specific date.** |
| Fact\_Appointments | **One appointment made by one patient with one doctor on a specific date .** |
| Fact\_Billing | **One bill issued to one patient on a specific date** |
| Fact\_Cleaning\_Service | **Number of cleaning services performed by one staff member in one room on a specific date** |

**3c. Define the type of each fact table:**

|  |  |  |
| --- | --- | --- |
| Fact Table | Fact Table Type | Reasoning |
| Fact\_Pharmacy | **Transaction Fact Table** | **Each row represents a medication dispense event.** |
| Fact\_Appointments | **Factless Fact Table** | **Represents appointment events with no measures. Counting only.( Each row = a** **distinct doctor-patient appointment)** |
| Fact\_Billing | **Transaction Fact Table** | **Each record is a single billing transaction for a specific patient.** |
| Fact\_Cleaning\_Service | **Periodic Snapshot** | **Each row = a summary of cleanings**, **not individual cleanings ( count of cleaning)** |

**3d. Define the dimensions and the type of each one:**

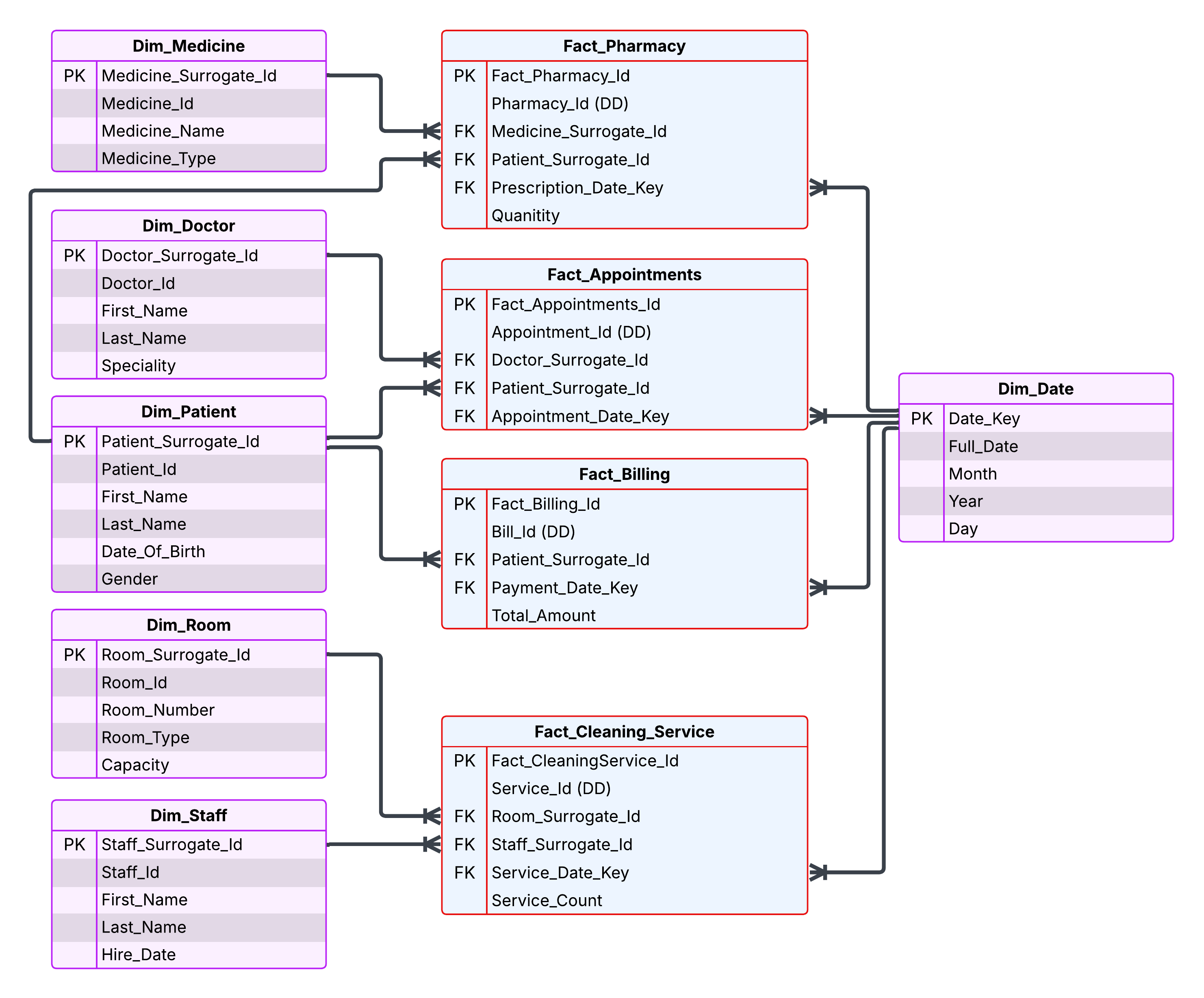
|  |  |  |
| --- | --- | --- |
| Dimension Table | Type | Reasoning |
| Dim\_Patient | **Conformed** | **Used across multiple fact tables like Fact\_Appointments, Fact\_Billing, Fact\_Pharmacy.** |
| Dim\_Date | **Conformed, Role-Playing .** | **Used across all facts for date-related analysis , and reused with different date roles (appointment date, billing date, cleaning service date)** |
| Dim\_Room | **Static Dimension** | **Room number, type, and capacity don’t typically change often** |
| Dim\_Staff | **Slowly Changing Dimension (SCD)** | **staff attributes like First\_Name, Last\_Name may change over time** |
| Dim\_Medicine | **Static Dimension** | **Medicine name and type are usually fixed; changes are rare and often treated by introducing a new medicineID.** |
| Dim\_Doctor | **Slowly Changing Dimension (SCD)** | **‘Speciality’, ‘First\_Name’, ‘Last\_Name ‘ attribute may be change over time** |

|  |  |  |
| --- | --- | --- |
| Dim\_Pharmacy | Degenerate Dimension | It has ONLY primary key attribute “Pharmacy\_Id” and this attribute is stored in the fact table “Fact\_Pharmacy” and has no  associated dimension table |
| Dim\_Appointment | **Degenerate Dimension** | **It has ONLY primary key attribute “Appointment\_Id” and this attribute is stored in the fact table “Fact\_Appointments” and has no**  **associated dimension table** |
| Dim\_Bill | **Degenerate Dimension** | **It has ONLY primary key attribute “Bill\_Id” and this attribute is stored in the fact table “Fact\_Billing” and has no associated dimension table** |
| Dim\_Service | **Degenerate Dimension** | **It has ONLY primary key attribute “Service\_Id” and this attribute is stored in the fact table “Fact\_Cleaning\_Service” and has no**  **associated dimension table** |

**3e. Define the measures that will appear in the fact tables and the type of each one:**

|  |  |  |
| --- | --- | --- |
| Fact Table | Measure | Measure Type |
| Fact\_Pharmacy | **Quantity** | **Fully Additive - Can be summed across all dimensions including time.** |
| Fact\_Billing | **Total\_Amount** | **Fully Additive - Can be summed over time,patients.** |
| Fact\_Cleaning\_Service | **Service\_Count** | **Fully Additive - Number of cleanings can be summed across all dimensions.** |

**3f. Physical model of DWH (the final star/galaxy schema):**

****

**4. Screenshots of the data flow tasks, and control flow tasks used for building the DWH:**

**Control Flow for the Whole Package:** **A diagram of a flowchart

AI-generated content may be incorrect.**

**Data Flow Task for Load Staging Tables:** A screenshot of a computer flowchart

AI-generated content may be incorrect.

**Data Flow Task for Load Dim\_Patients:** A screenshot of a computer screen

AI-generated content may be incorrect.

**Data Flow Task for Load Dim\_Doctor:**  A screenshot of a computer screen

AI-generated content may be incorrect.

**Data Flow Task for Load Dim\_Medicine:**  A screenshot of a computer screen

AI-generated content may be incorrect.

**Data Flow Task of Load Dim\_Room:A diagram of a room

AI-generated content may be incorrect.**

**Data Flow Task for Load Dim\_Staff: A screenshot of a computer

AI-generated content may be incorrect.**

**Data Flow Task for Load Dim\_Date: A diagram of a company

AI-generated content may be incorrect.**

**Data Flow Task for Load Fact\_Pharmacy: A screenshot of a computer screen

AI-generated content may be incorrect.**

**Data Flow Task for Load Fact\_Appointments: A screenshot of a diagram

AI-generated content may be incorrect.**

**Data Flow Task for Load Fact\_Billing: A screenshot of a diagram

AI-generated content may be incorrect.**

**Data Flow Task for Load Fact\_Cleaning\_Service: A screenshot of a diagram

AI-generated content may be incorrect.**

**5. Queries on each fact table to let me understand what this fact table represents and what insights we can get from it, and a screenshot of the result set of each query:**

**6. Screenshots of the deployed packages in SSIS with their schedule.**

**7. [Bonus] Build an interactive dashboard for the DWH using any data visualization tool (Ex: Microsoft Power BI).**