# Sri Lanka Institute of Information Technology

IT3021 - Data warehousing and Business Intelligence

Year 3 Semester 2

DWBI – Assignment 02

**IT22923424 – Jayasekara S.S**

## Step 1: Data source for the assignment 2

**Data Set Title -Hospital Appointment Management System**

For this step, we designed a **Star Schema** data warehouse model tailored to the Hospital Appointment Management System. The central **fact table**, **FactAppointment**, captures measurable business events such as appointments, treatment costs, and appointment durations. It is surrounded by **dimension tables** that provide descriptive context about each appointment, including **patient, doctor, department, payment method, and date.**

**Schema :Star Schema**

| Table Name | Description |
|---|---|
| DimDate | Contains full date breakdown for analytics (ex: day, week, month, quarter, year). |
| DimDepartment | Contains department info (Department ID, DepartmentName). |
| DimPatient (SCD) | Describes patient details (PatientID, FullName, Gender, Address). |
| DimDoctor (SCD) | Describes doctor profile (DoctorID, Name, Specialization, DepartmentID). |
| DimPayment (SCD) | Stores PaymentID, payment method and status. |
| FactAppointment | Stores appointment-specific data including AppointmentID, PatientSK, DoctorSK, DepartmentSK, PaymentSK, DateSK, Status, TreatmentCost, and timing details. Linked to all dimension tables via surrogate keys. |

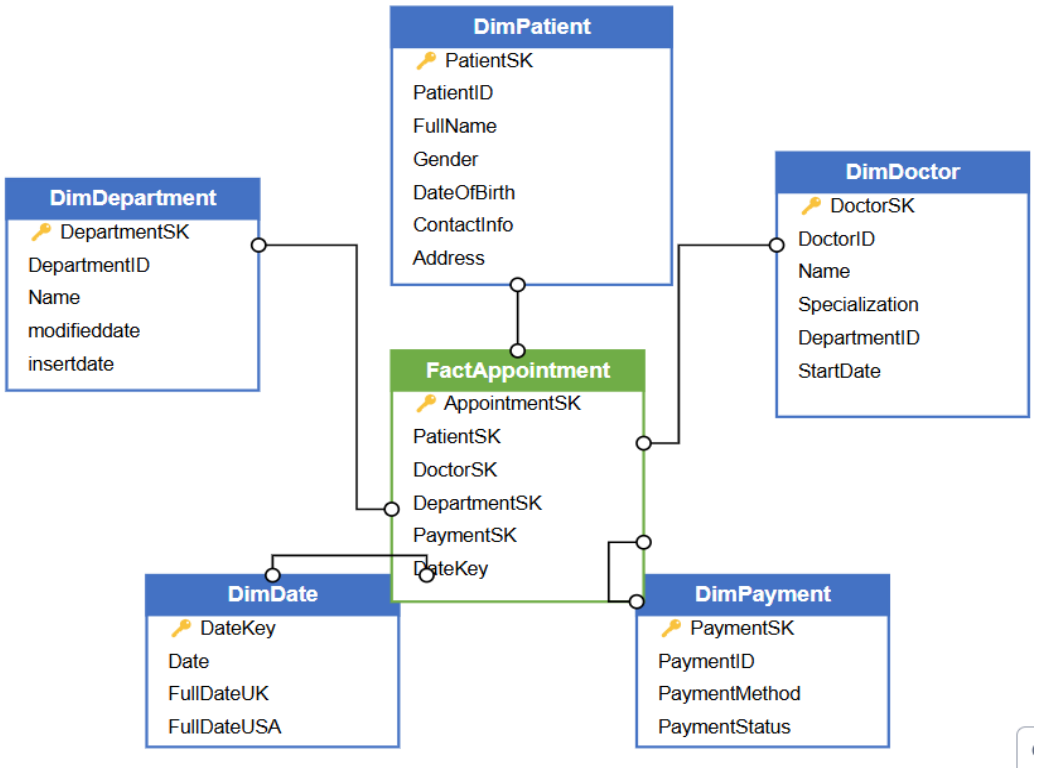
**Assumptions Made**

- A doctor always belongs to **one department**.
- Each appointment has one treatment and one payment.
- Payment method/status may change over time (→ SCD).
- Date values are tracked using a surrogate **date_id**.

**ER Diagram**

| Tables | Attributes |
|---|---|
| Patient | patient_id, full_name, gender, date_of_birth, contact_info, address |
| Doctor | doctor_id, name, specialization, department_id (FK department) |
| Department | department_id, name |
| Appointment | appointment_id, patient_id (FK1 patient), doctor_id (FK2 doctor), appointment_date, status (Scheduled/Completed/Cancelled) |
| Treatment | treatment_id, appointment_id (FK appointment), treatment_type, cost |
| Payment | payment_id, appointment_id (FK appionment), payment_date, payment_method, amount, payment_status |

- A **Patient** can have multiple **Appointments**
- Each **Appointment** is with a **Doctor**
- A **Doctor** belongs to a **Department**
- Each **Appointment** may have **Treatment(s)**
- Each **Treatment** may have one **Payment**



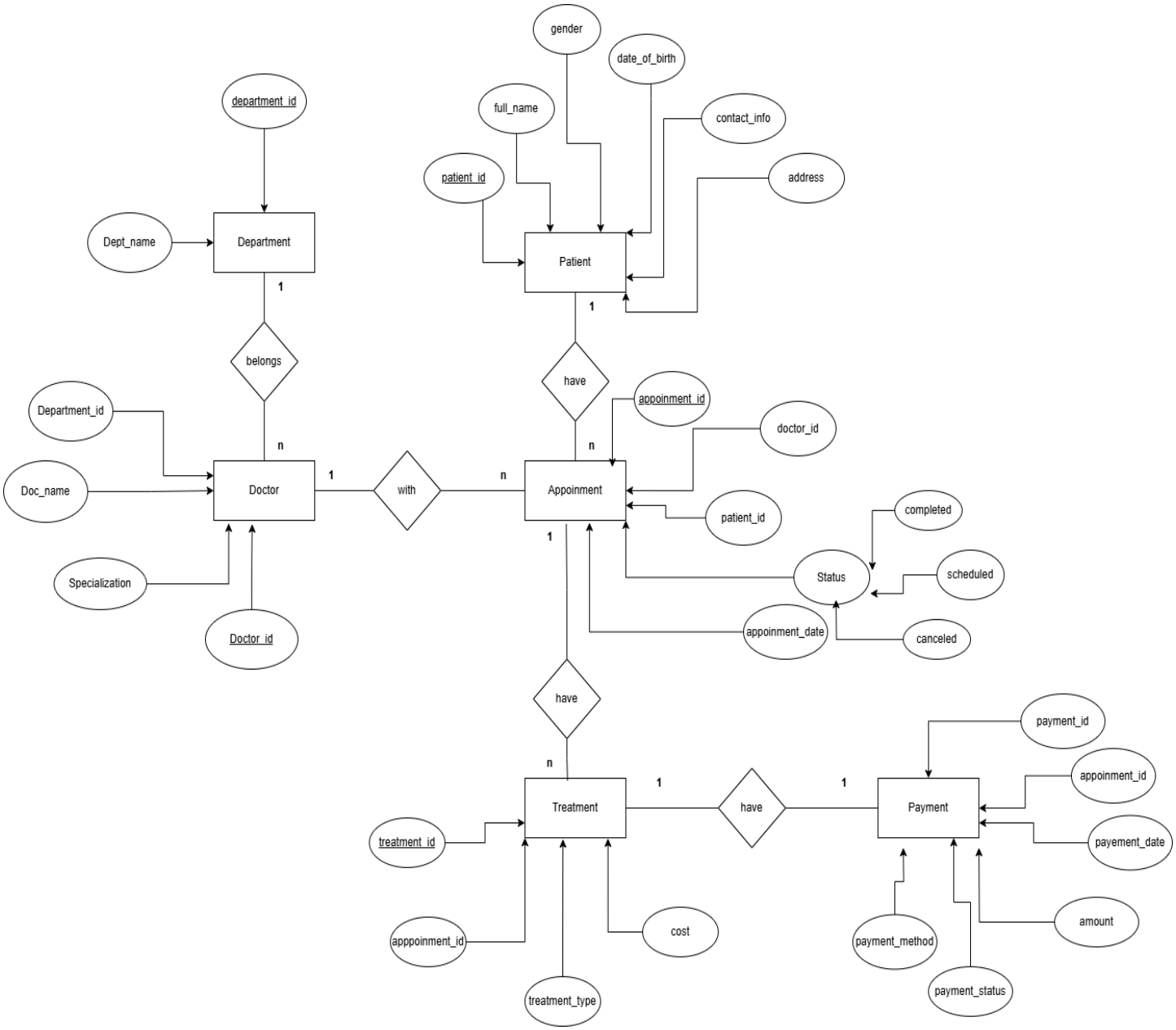



*Figure1:ER Diagram*

## Step 2: SSAS Cube implementation

In this step, a **multidimensional OLAP cube** was designed, developed, and deployed using **SQL Server Analysis Services (SSAS)**. The cube was built on top of the Data Warehouse created in Assignment 01, enabling efficient multidimensional analysis.

### Implementation Process

1. **SSASProjectCreation**

   A new Analysis Services project was created using the Multidimensional and Data Mining template in SSDT.

2. **Data Source and Data Source View (DSV)**

   - A data source was added pointing to the Data Warehouse database.
   - A **Data Source View** was created using the following tables

   (FactAppointment, DimPatient, DimDoctor, DimDepartment, DimPayment, DimDate)

3. **Cube Creation and Configuration**

   - A new cube was created using the **Cube Wizard**.
   - FactAppointment was used as the **Measure Group**.
   - Key measures included:
     - Treatment Cost
     - Appointment Count

4. **Dimension Design**

   - All dimension tables were added as dimensions: DimPatient, DimDoctor, DimPayment, DimDate, and DimDepartment.
   - A regular relationship was defined between each dimension and the fact table using surrogate keys (e.g., PatientSK, DoctorSK).

5. **Hierarchy Setup**

    - A hierarchy was created in DimDate: Year → Quarter → Month → Day

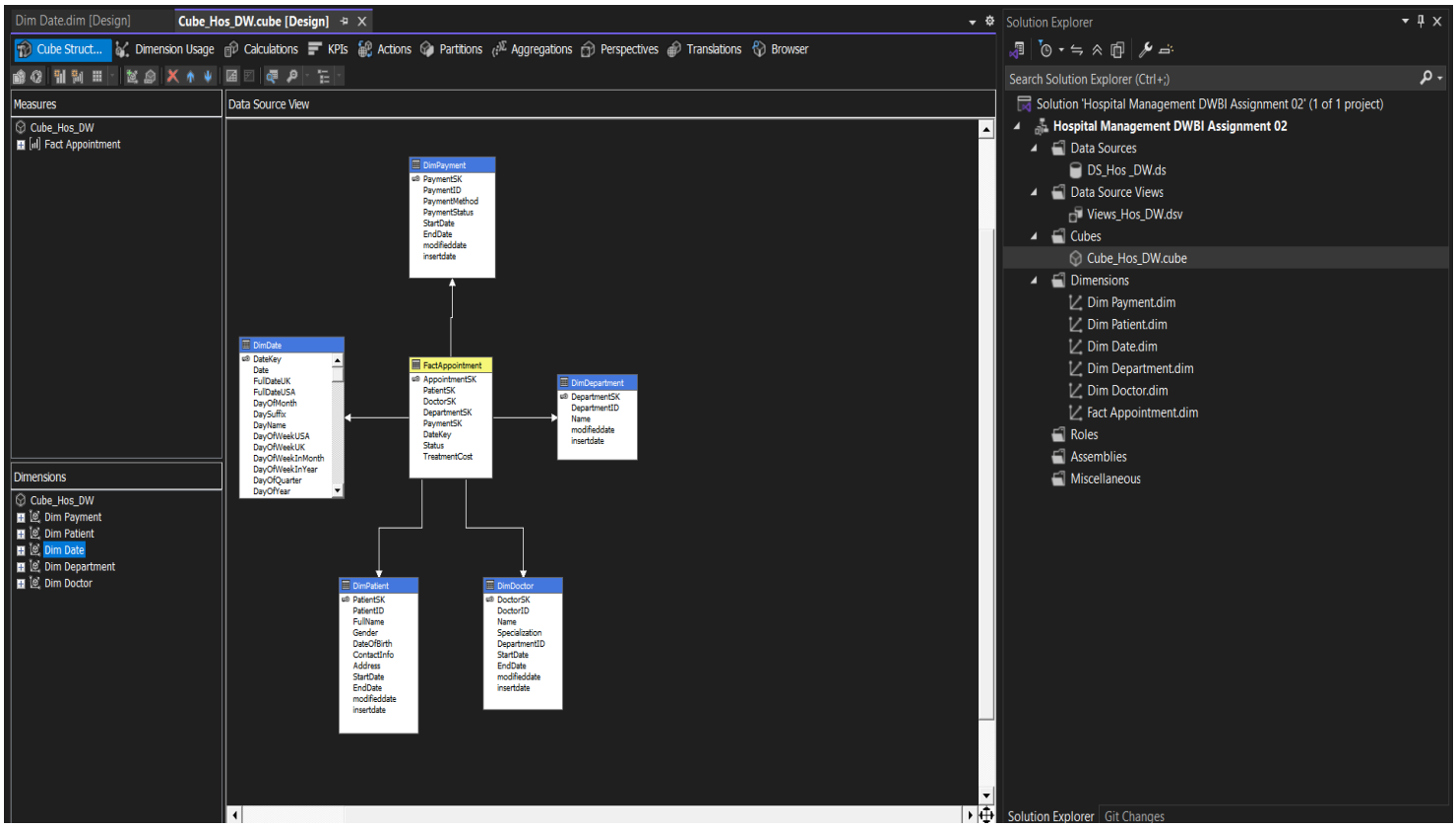
6. **KPI Creation – Treatment Cost Threshold**

    - A Key Performance Indicator (KPI) was added to the cube using the cube_hos_DW design browser in SSDT.
    - KPI Name: High Treatment Cost KPI.
    - Goal: Identify appointments where the Treatment Cost exceeds 50000.
    - Global Expression logic: Treatment Cost >50000.
    - Status and trend indicators were configured to visually highlight high-cost cases in client tools like Excel or Power BI.

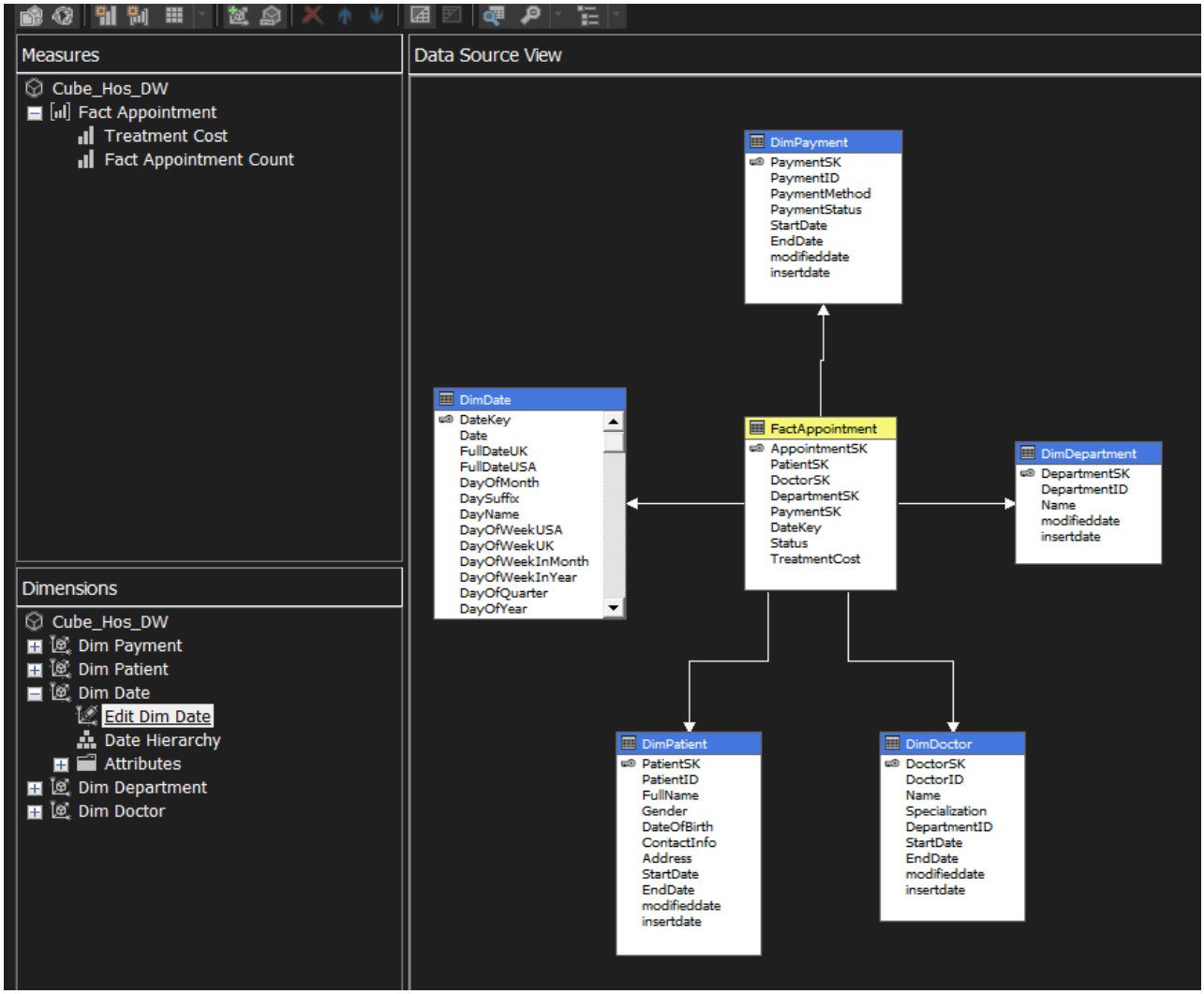
7. **Deployment and Processing**

    - The cube was successfully deployed to the localhost SSAS instance.
    - It was then processed to populate the metadata and data from the Data Warehouse.

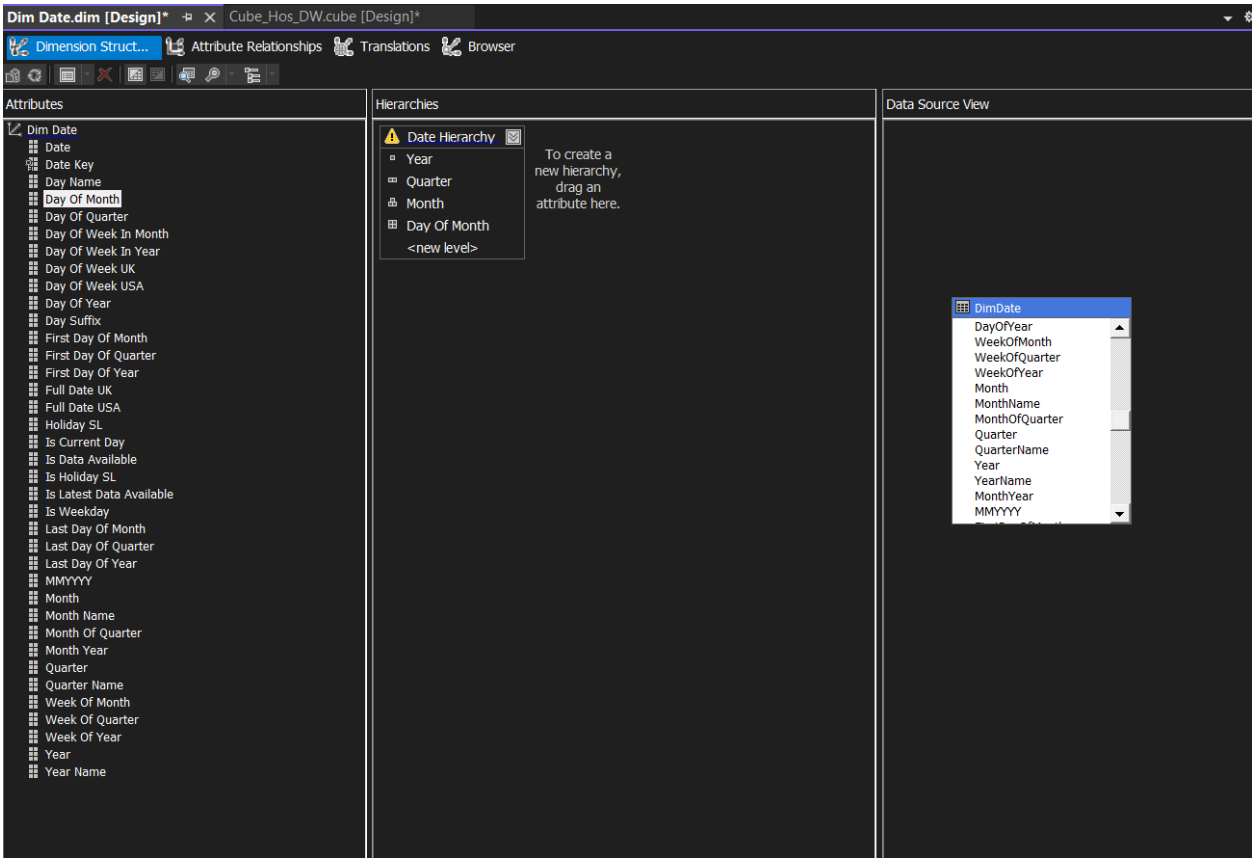
1.Use data warehouse as the data source and create an SSAS cube.

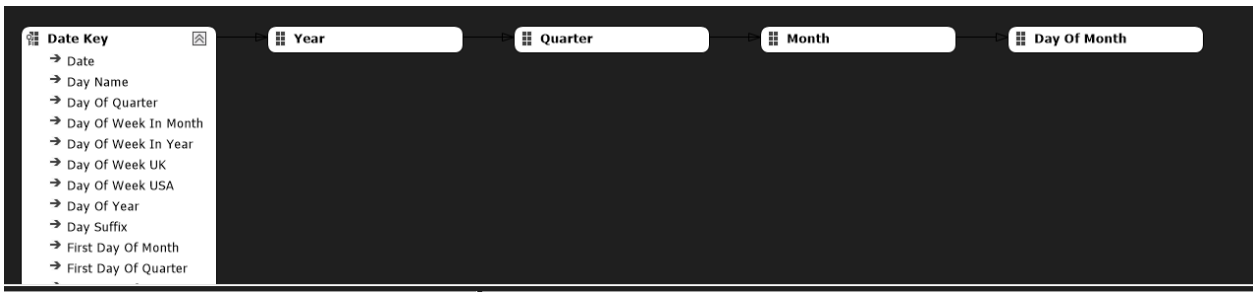


2.Design the cube by including necessary measures in fact table(s) and connecting dimensions.

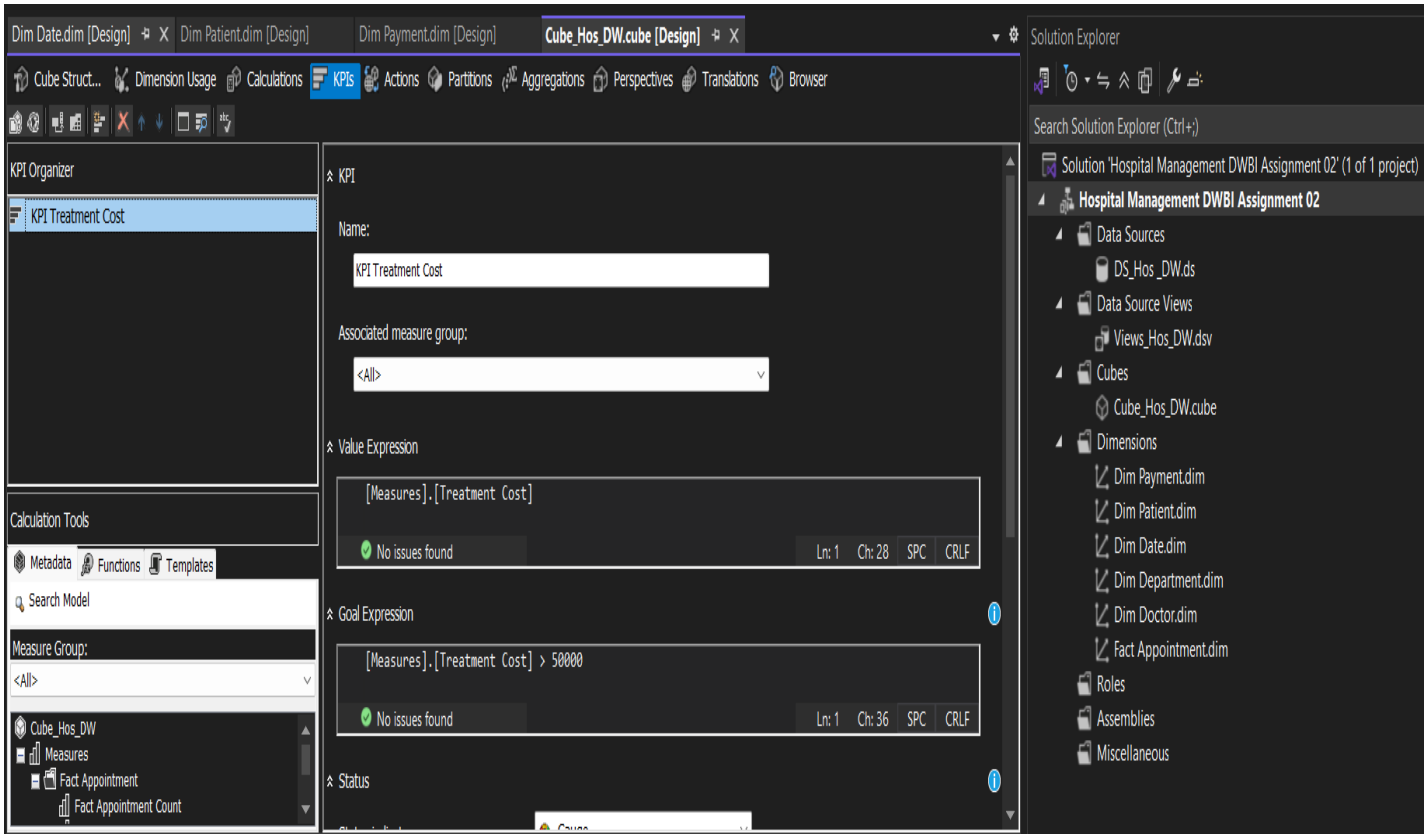


3.Implement at least one hierarchy in the cube. (Date Hierarchy)



4.Implement at least one hierarchy in the cube. (Date Hierarchy)



## Step 3: Demonstration of OLAP operations

In this step, we connected the deployed SSAS cube to **Microsoft Excel** to perform OLAP (Online Analytical Processing) operations. These operations enable interactive and multidimensional analysis of data stored in the data warehouse.

**Step-by-Step Implementation**

1. **Connecting Excel to SSAS Cube**

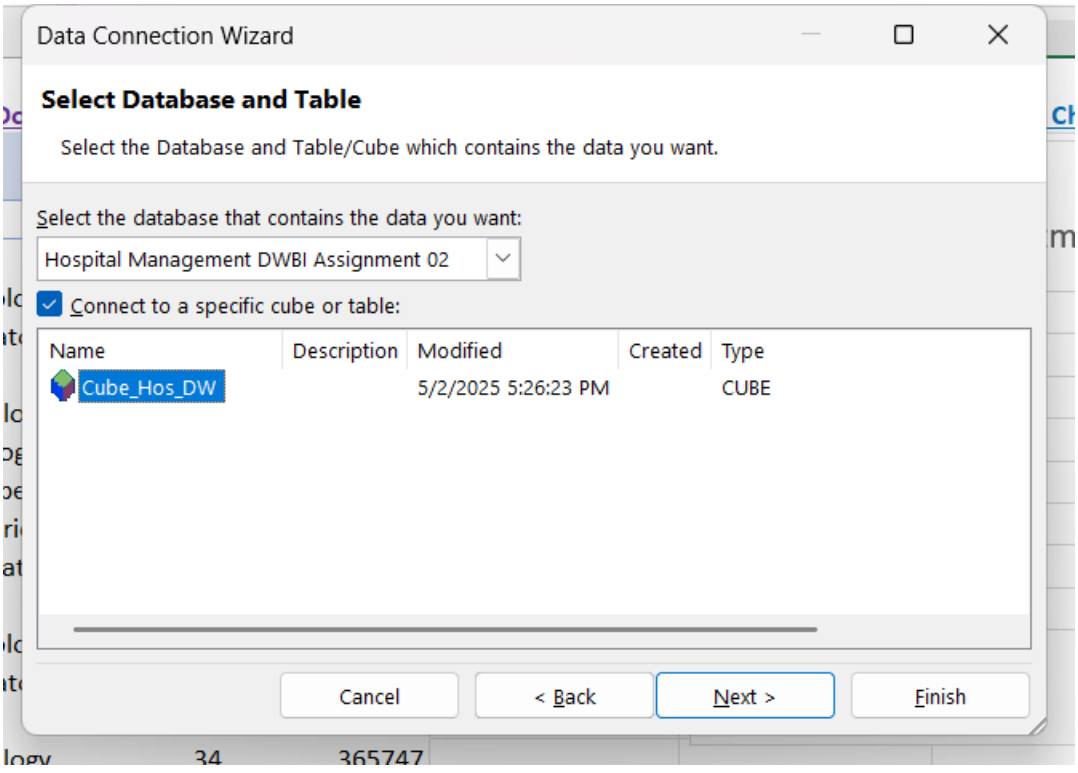
Opened Microsoft Excel, navigated to Data → Get Data → From Database → From Analysis Services, entered the server's name localhost, selected the deployed cube "Hospital DW Cube", and inserted a PivotTable in a new worksheet.

2. **Performing OLAP Operations**

- **Roll-up**

  **Hierarchy Used: DimDate → Year → Quarter → Month → Day**

  Dragged the DimDate hierarchy to Rows. Then added TreatmentCost to Values. Viewed totals at Year level (aggregated). At this point, Excel displayed aggregated treatment costs at the Year level, without expanding into quarters or months. This summarized, higher-level view of the data represents a roll-up, as it aggregates detailed information into broader time periods for easier analysis.



## Rollup &DrillDown using each Department Treatment Cost

| Row Labels | Fact Appointment Count | Treatment Cost |
|---|---|---|
| **2024** | **1380** | **14543159** |
| Cardiology | 78 | 888995 |
| Dermatology | 153 | 1483095 |
| ENT | 401 | 4203099 |
| Neurology | 118 | 1219832 |
| Oncology | 60 | 641622 |
| Orthopedics | 203 | 2250466 |
| Pediatrics | 290 | 3103819 |
| Psychiatry | 77 | 752231 |
| **2025** | **620** | **6358609** |
| Cardiology | 18 | 189153 |
| Dermatology | 58 | 531677 |
| ENT | 193 | 2041769 |
| Neurology | 65 | 706641 |
| Oncology | 28 | 275042 |
| Orthopedics | 96 | 975295 |
| Pediatrics | 136 | 1341647 |
| Psychiatry | 26 | 297385 |
| **Grand Total** | **2000** | **20901768** |

- **Drill Down**

  Expanded the **Year → Quarter → Month → Day** levels using the + signs. Observed progressively more detailed data. This represents a **drill-down** operation.

Rollup &DrillDown using each Department Treatment Cost

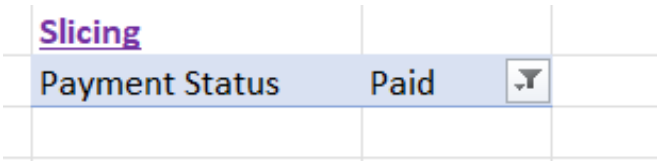
| Row Labels | Fact Appointment Count | Treatment Cost |
|---|---|---|
| **2024** | **1380** | **14543159** |
| **2** | **358** | **3721234** |
| Cardiology | 17 | 150109 |
| Dermatology | 33 | 327183 |
| ENT | 108 | 1091455 |
| Neurology | 31 | 330832 |
| Oncology | 13 | 150332 |
| Orthopedics | 59 | 622004 |
| Pediatrics | 77 | 852970 |
| Psychiatry | 20 | 196349 |
| **3** | **517** | **5415657** |
| Cardiology | 27 | 326630 |
| Dermatology | 56 | 500580 |
| ENT | 156 | 1662604 |
| Neurology | 34 | 365747 |
| Oncology | 28 | 286402 |
| Orthopedics | 82 | 877279 |
| Pediatrics | 108 | 1158635 |
| Psychiatry | 26 | 237780 |
| **4** | **505** | **5406268** |
| Cardiology | 34 | 412256 |
| Dermatology | 64 | 655332 |
| ENT | 137 | 1449040 |
| Neurology | 53 | 523253 |

et1

- **Slice**

  Dragged dimensions like Department, Payment Status, and Year into the **Filters** area. Selected single values **Payment Status =Paid**. This isolated a subset of data to a slice.

  Then we sliced the data by Payment Status = Paid to analyze only successful payments.




- **Dice**

  Applied **multiple filters** across dimensions: Department = ENT (Using Slicing). Dragged **Doctor Name to Rows** and **Payment Method to Columns.** And added to the **Treatment cost** for **values.** This allowed focused analysis of treatment cost handled by doctors in Cardiology department using a specific payment method.

Dice

| Name | ENT |
|---|---|

| Treatment Cost | Column | | | | |
|---|---|---|---|---|---|
| Row Labels | Card | Cash | Insurance | Online Tra | Grand Total |
| Amanda Powell | 300956 | 193084 | 187267 | 295459 | 976766 |
| Jennifer Shelton | 246971 | 318149 | 209320 | 188586 | 963026 |
| Karen Mcintosh PhD | 258305 | 217284 | 227865 | 269263 | 972717 |
| Mark Walker | 345828 | 345428 | 272133 | 142686 | 1106075 |
| Sean Martin | 358152 | 388805 | 318362 | 283687 | 1349006 |
| Timothy Schneider | 219061 | 186011 | 280240 | 191966 | 877278 |
| **Grand Total** | **1729273** | **1648761** | **1495187** | **1371647** | **6244868** |

- **Pivot**

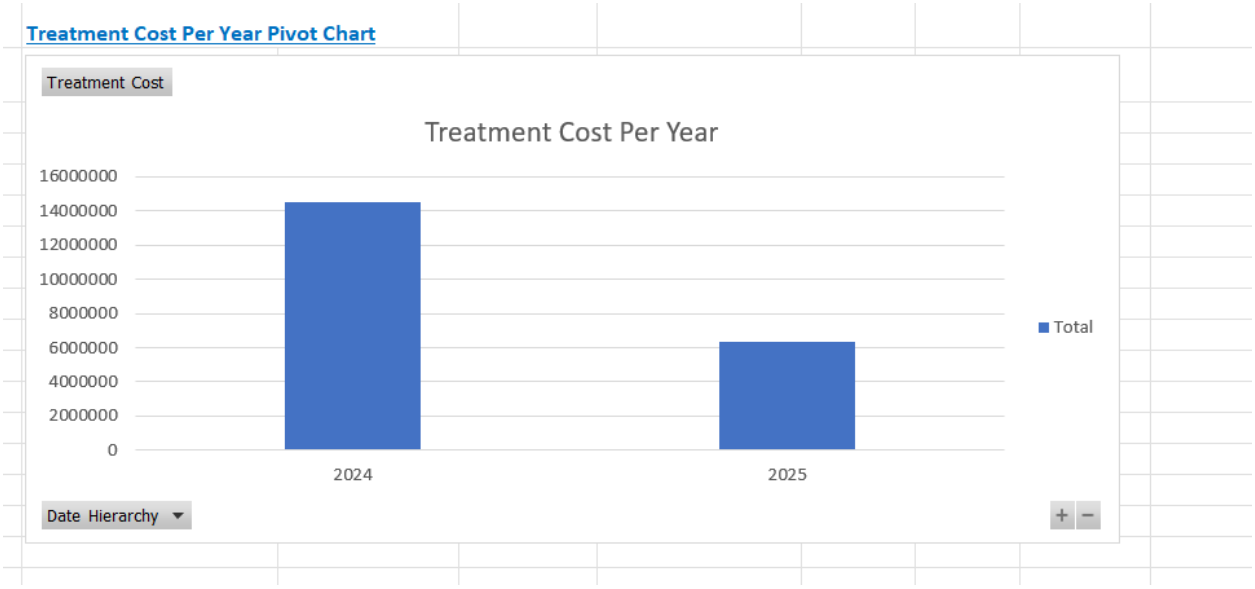
  Initially placed patient Name in **Rows** and Department in **Columns**.

  **Treatment Cost** used for **values** and the **filter** is used **payment method**.

  Then we performed a **pivot** to compare treatment costs across departments and patient from different perspectives.

**Pivot Table**

| Payment Status | Paid | |
|---|---|---|
| | | |
| **Treatment Cost** | Department | |
| Patient Name | ENT | Grand Total |
| Amanda Powell | 296722 | 296722 |
| Jennifer Shelton | 305526 | 305526 |
| Karen Mcintosh PhD | 362641 | 362641 |
| Mark Walker | 329757 | 329757 |
| Sean Martin | 483694 | 483694 |
| Timothy Schneider | 286913 | 286913 |
| **Grand Total** | **2065253** | **2065253** |



## Step 4: Power BI Reports

Power BI Desktop was used to create visual reports connected live to the SSAS cube developed in Assignment 01. This enabled real-time, interactive analysis of hospital appointment and treatment data.

**Connection Process:**

- Connected Power BI to the SSAS cube using localhost
- Chose **live connection** to ensure real-time interaction with cube data

**Visual Reports Created:**

- **Matrix Table** displaying TreatmentCost across time and departments
- **Bar Chart** showing doctor-wise appointment counts
- **Pie Chart** to analyze distribution by Payment Methods
- **KPI Visual** to highlight treatments exceeding the cost threshold (TreatmentCost > 50000)
- **Slicers** to filter by Year, Payment Status, and Department

**OLAP Features Demonstrated:**

- Drill-down and roll-up using DimDate hierarchy
- Real-time slicing and dicing using slicers
- Pivot-like interactivity by rearranging fields in visuals

**Outcome:**

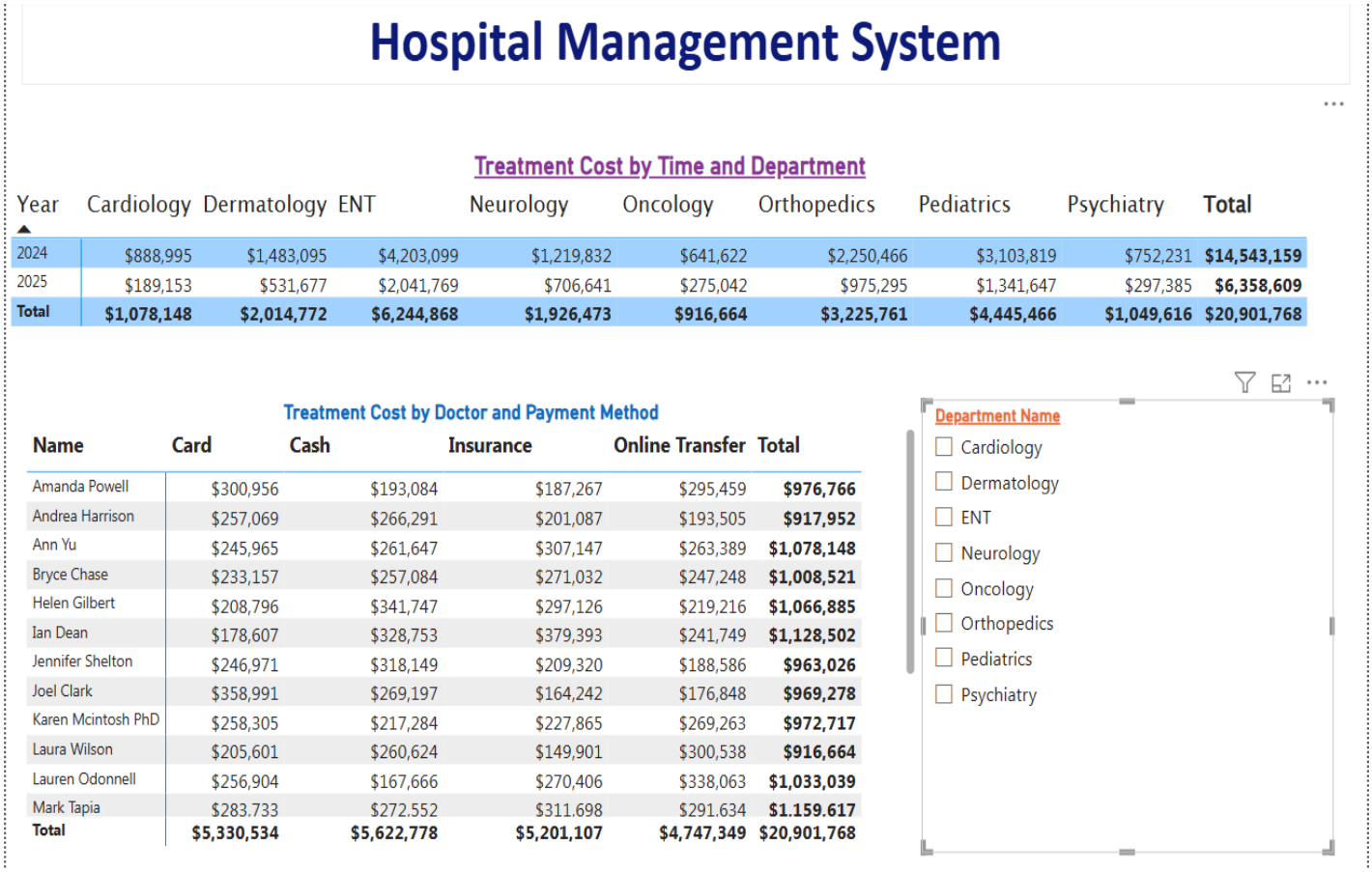
This step provided a user-friendly and analytical front-end, enabling decision-makers to explore data across multiple dimensions with minimal effort and maximum insight.

## Report 01

### Matrix Visual for Tabular Analysis

In the first Power BI report, a **Matrix visual** was used to display detailed tabular data with **row and column groupings**. The matrix used the DimDate hierarchy (Year → Quarter → Month) as **row groups** and Department Name from DimDepartment as **column groups**. The measure TreatmentCost was placed in the **Values** area to show the total treatment cost across departments and time periods.

This report helps stakeholders monitor **department-level performance trends** over time and supports **drill-down** and **roll-up** analysis directly within the matrix. It serves as a foundational view for data inspection and financial comparison across departments and months.



## Report 02

**Slicers and Visual Insights (Live SSAS Connection)**

In this report, we used Power BI connected live to the SSAS cube to build an interactive dashboard. Although custom cascading slicers are not editable due to the live connection, we added multiple slicers that filter the entire report:

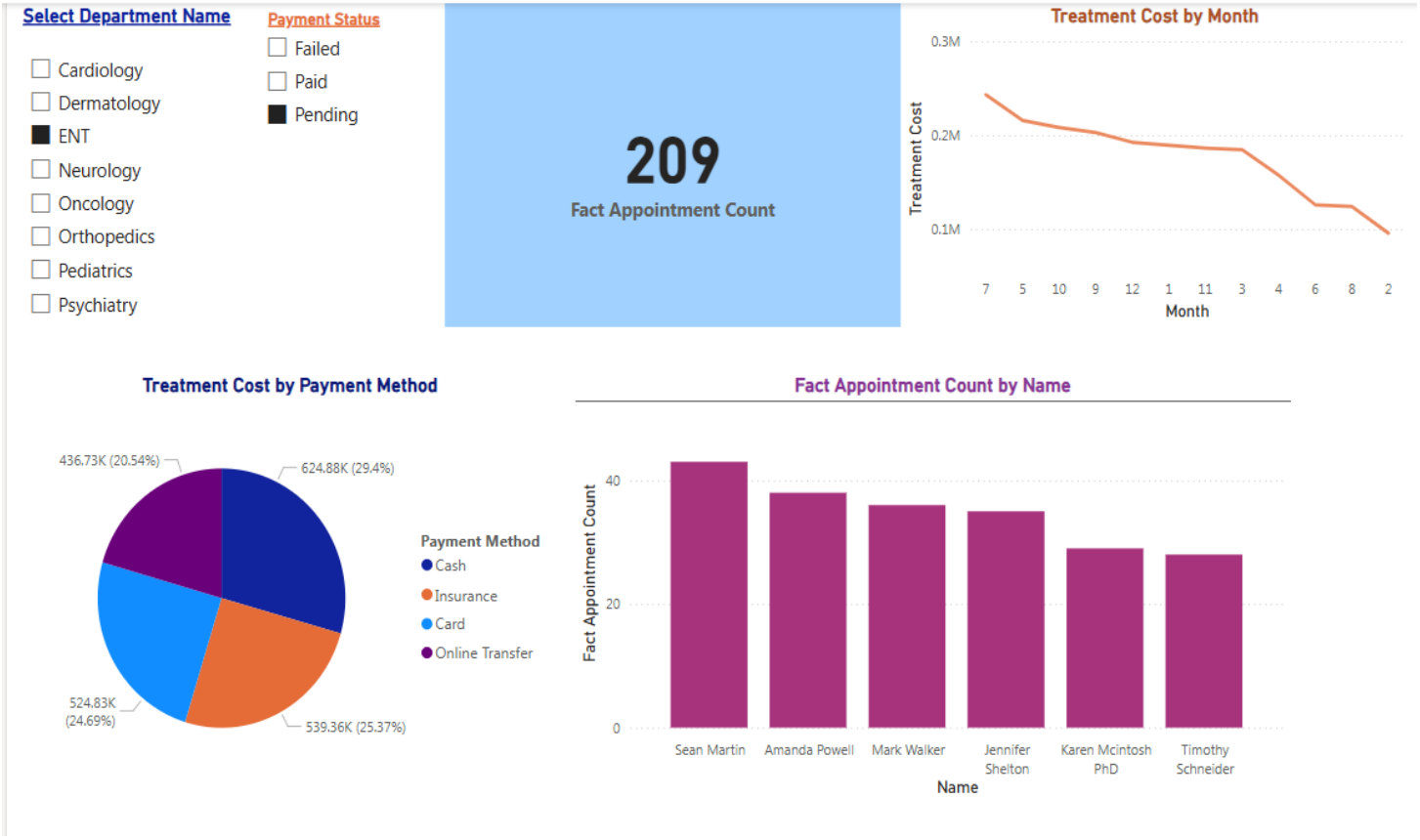
- **Department**
- **Year**
- **Payment Status**

The report includes:

- A **bar chart** showing appointments by doctor
- A **pie chart** showing treatment cost by payment method
- A **line chart** showing monthly treatment trends
- A **card visual** displaying appointment total

This report allows stakeholders to explore insights by slicing across multiple dimensions, even when relationships cannot be manually defined.

## Report 02


Select Department Name
Cardiology
Dermatology
ENT
Neurology
Oncology
Orthopedics
Pediatrics
Psychiatry
Payment Status
Failed
Paid
Pending
209
Fact Appointment Count
Treatment Cost by Month
0.3M
0.2M
0.1M
Treatment Cost
7 5 10 9 12 1 11 3 4 6 8 2
Month
Treatment Cost by Payment Method
436.73K (20.54%)
624.88K (29.4%)
524.83K (24.69%)
539.36K (25.37%)
Payment Method
Cash
Insurance
Card
Online Transfer
Fact Appointment Count by Name
Fact Appointment Count
40
20
0
Sean Martin
Amanda Powell
Mark Walker
Jennifer Shelton
Karen Mcintosh PhD
Timothy Schneider
Name

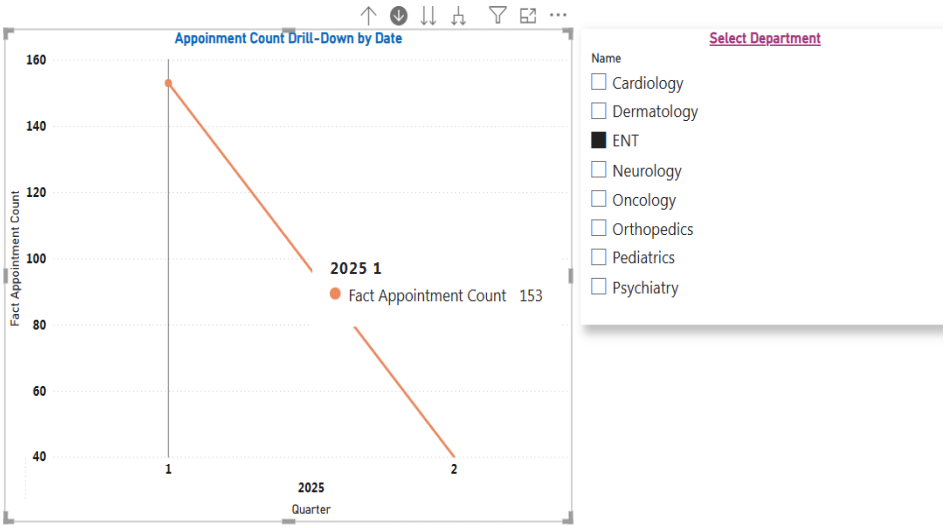
## Report 03

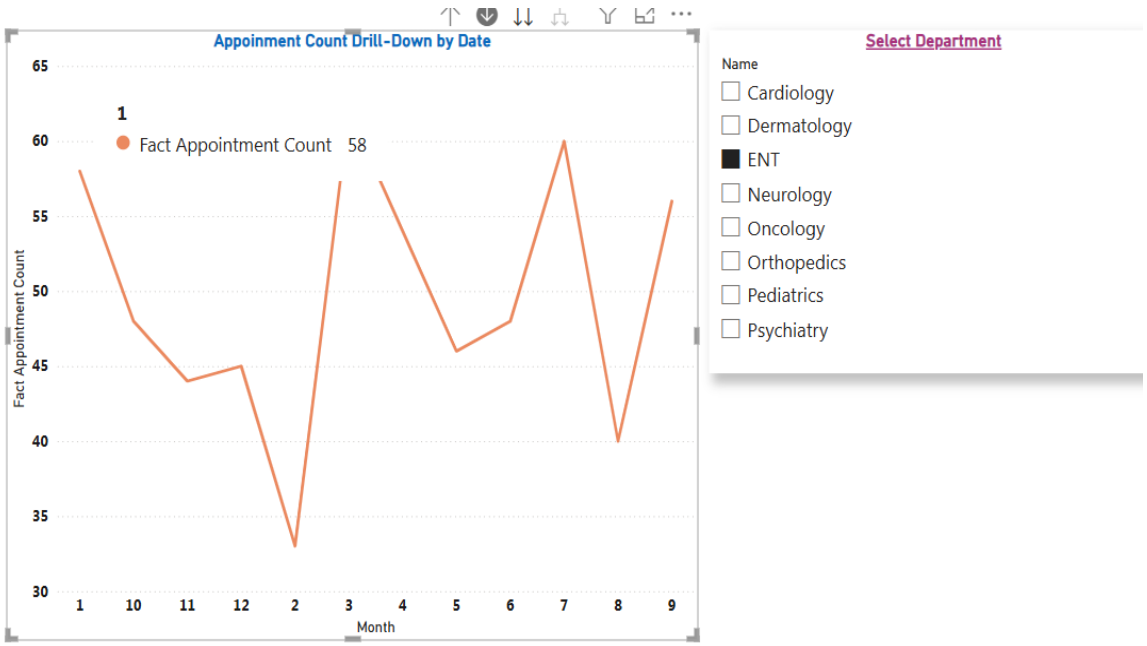
### Drill-down Hierarchy Report

This report uses a **bar chart with a date hierarchy** from the DimDate dimension to support interactive drill-down analysis. The hierarchy includes **Year → Quarter → Month**, allowing users to explore TreatmentCost trends over time.

Users can click on any year to drill into quarters and further into months. This enables intuitive navigation of time-based data and helps stakeholders identify trends at different granularities.

A measure such as TreatmentCost is plotted on the Y-axis, providing insight into how costs evolve over time. The drill-down functionality is enabled through native Power BI features using a hierarchy defined in the SSAS cube.



## Report 04

### Department Drill-Through Report

This report enables users to **right-click on a department in a visual** (e.g., bar chart) and **navigate to a separate page** showing filtered, detailed insights about that department.

The Department Details page was configured using a **drill-through filter** on the Department Name field. Visuals such as pie charts, bar graphs, cards, and tables update dynamically based on the selected department.

This drill-through interaction supports **contextual analysis**, allowing users to explore specific departments in greater detail for more informed decision-making.