جامعة الإسكندرية
ALEXANDRIA UNIVERSITY

Healthcare Data analysis Project

Project Members:

Amir Mohamed Abd elmoniem (220590)

Malak Mohamed Abd elmoniem (220591)

Nadeen Ahmed Nabil (220594)

Course name

National and International Healthcare systems

Date of Submission

Spring 2025

Healthcare Dataset Description and Entity Relationships

This dataset captures comprehensive healthcare-related information for a hospital or medical center. It is designed to track detailed data about patients, their visits, associated diagnoses, medical procedures, healthcare providers, hospital departments, insurance details, and the cities where patients reside. The interconnected tables provide a structured way to follow each patient's journey and medical interactions across multiple visits, diagnoses, and procedures.

Dataset Overview

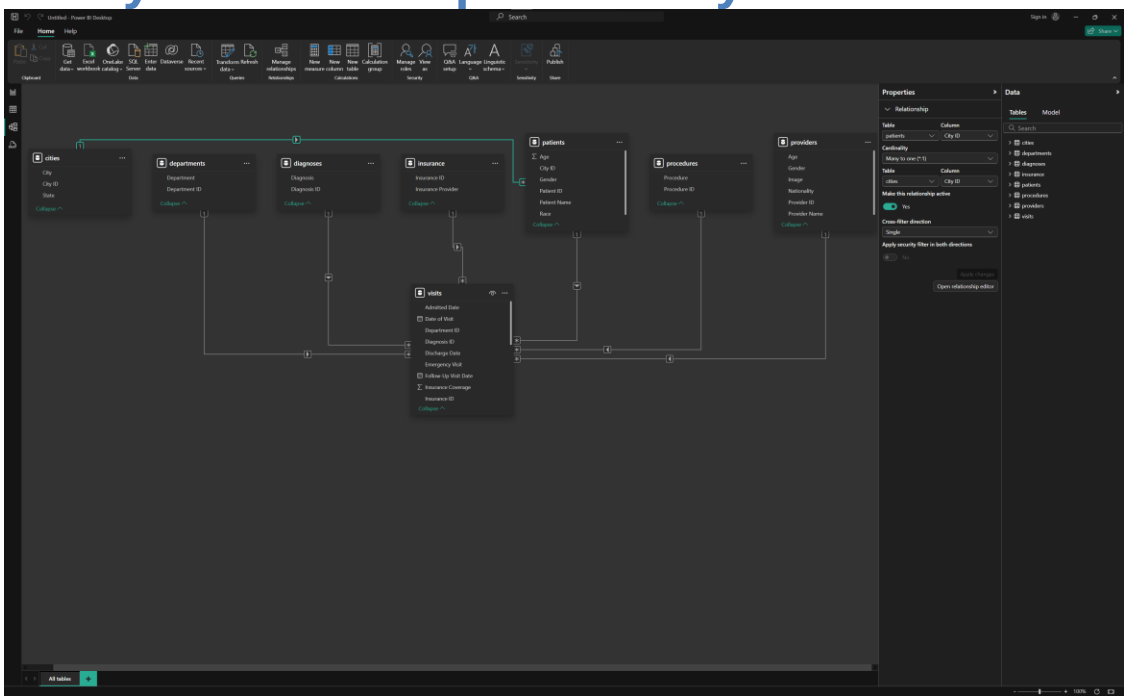
The dataset is organized into several key tables, each representing a distinct entity within the hospital ecosystem. The primary focus is on **patients** and their **visits**, where various health-related interactions and administrative processes take place. Other supporting entities such as **diagnoses**, **procedures**, **providers**, **departments**, **insurance**, and **cities** are linked to build a detailed and holistic picture of healthcare delivery.

Data Dictionary and Entity Descriptions

Table	Primary Key	Description	Key Relationships
cities	City ID	List of cities where patients reside.	Referenced by: *patients* (City ID)
departments	Department ID	Hospital departments such as cardiology, emergency room (ER), etc.	Referenced by: *visits* (Department ID)
diagnoses	Diagnosis ID	Diagnostic codes or descriptions of medical conditions.	Referenced by: *visits* (Diagnosis ID)
insurance	Insurance ID	Information about insurance providers and their policies.	Referenced by: *visits* (Insurance ID)
patients	Patient ID	Demographic and metadata information for patients.	References: *cities* (City ID) Referenced by: *visits* (Patient ID)
procedures	Procedure ID	Descriptions of medical procedures.	Typically related to visits or diagnoses.
providers	Provider ID	Medical professionals including doctors,	Usually linked to visits or procedures.

Table	Primary Key	Description	Key Relationships
		nurses, and specialists.	
visits	Visit ID	Records of patient visits to the hospital or clinic.	References: *patients* (Patient ID), *departments* (Department ID), *diagnoses* (Diagnosis ID), *insurance* (Insurance ID) May reference: *providers* and *procedures* in extended models

Entity Relationship Summary



- **Patients** is the central entity in the dataset. Each patient record includes demographic data and links to the city where they reside via the City ID.
- **Visits** act as a key junction table, connecting patients to their hospital interactions. Each visit references the patient, the department attended, the diagnosis made, and the insurance provider for that encounter.
- **Departments, Diagnoses, and Insurance** are referenced directly by visits through their respective IDs to provide context for each hospital encounter.
- **Procedures and Providers** are included in the dataset as well, though explicit relationships to visits are not shown in the preview. Typically, these entities are linked to visits or diagnoses to capture which procedures were performed and which providers were involved.

Use Case and Importance

This healthcare dataset is valuable for analyzing patient pathways, understanding departmental workload, evaluating diagnosis trends, and assessing insurance impacts on care delivery. By structurally linking these entities, hospitals and researchers can efficiently track the full spectrum of patient care, from admission to treatment outcomes, supporting improved clinical decision-making and resource management.

DAX Measures Explanation

Patient Count:
Description: This measure tallies the total number of unique patients in the dataset.

Procedure Count:
Description: This measure computes the number of unique procedures performed, which could be used to understand procedure diversity or frequency.

High Cost Visits:
Description: This measure filters and counts the number of visits where the cost exceeds a certain threshold, indicating high-cost care episodes.

Note: The threshold (e.g., 1000) can be adjusted based on your specific definition of 'high cost.'

Average Age:
Description: This measure calculates the average age of all patients, providing insight into the general age distribution of your patient base.

Age Group:
Description: This measure categorizes patients into age groups for better demographic analysis.

Male Patients

Description: This measure calculates the number of male patients in the dataset, offering demographic insights into the patient gender distribution.

Total Visits

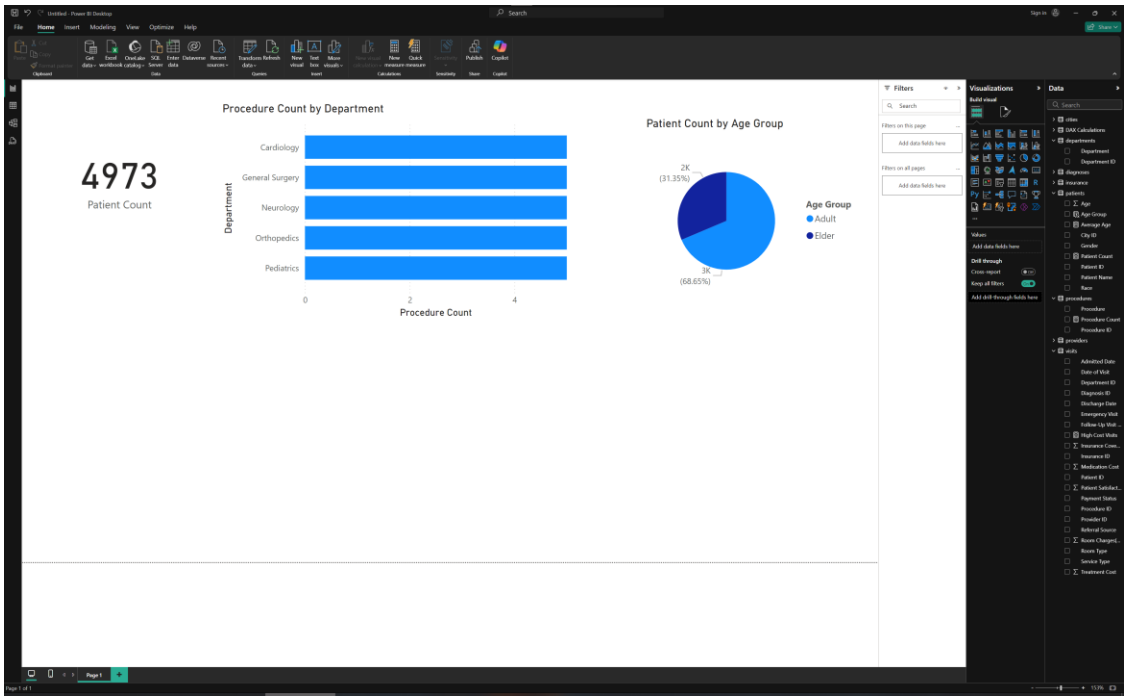
Description: This measure calculates the total number of patient visits recorded in the dataset. It provides insights into service utilization and visit frequency, which are critical for operational and capacity planning.

Purpose

- **Operational Insights**: Helps in understanding the volume of patient interactions, assisting in resource planning and staff scheduling.
- **Trend Analysis**: Utilized for identifying patterns over time, such as peak visit times or seasonal variations.
- **Performance Metrics**: Enables the drafting of key performance indicators related to patient turnover and clinic efficiency.

```
Age Group =

SWITCH(

    TRUE(),

    'Patients'[Age] < 18, "Child",

    'Patients'[Age] >= 18 && 'Patients'[Age] <= 60, "Adult",

    'Patients'[Age] > 60, "Elder",

    "Unknown"

)

Average Age = AVERAGE(patients[age])


High Cost Visits =
CALCULATE(
   COUNTROWS(visits),
   visits[billing_amount] > 10000
)


Procedure Count = COUNTROWS(procedures)


 Patient Count = COUNTROWS(patients)

Male Patients =
CALCULATE(
   COUNTROWS('Patients'),
```

```
  'Patients'[Gender] = "Male"
)

Total Visits = COUNTROWS(visits)
```



Visuals Review

Patient Count:

- **Visual Type**: Card
- **Content**: Displays the total patient count as 4973. This provides a quick snapshot of the total unique patient population.

Procedure Count:

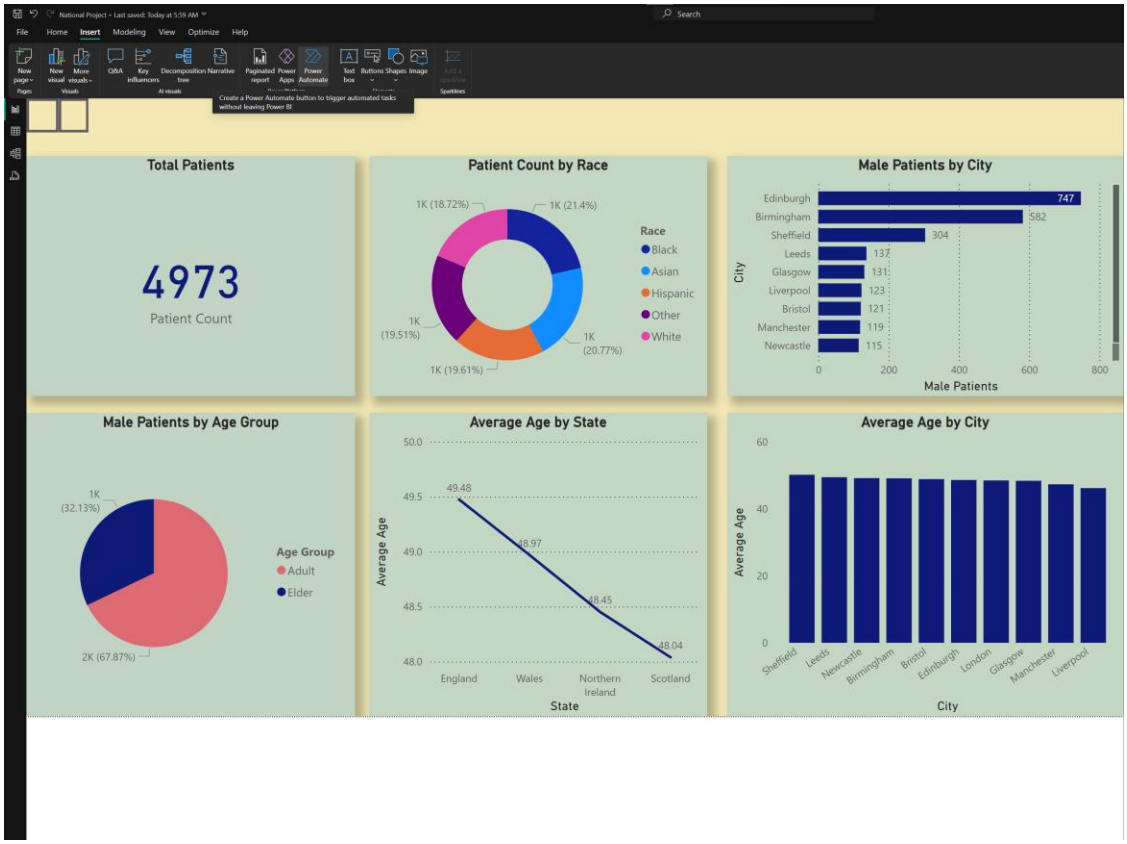
- **Visual Type**: Bar chart
- **Content**: This chart presents proportions of different procedure categories or counts, which helps in visualizing the distribution of procedures conducted.

Age Group Distribution:

- **Visual Type**: Pie/Donut Chart
- **Content**: Represents the distribution of patients across different age groups like Adult and Elder. It shows that approximately 31.35% are

Elders, while 68.65% are Adults. Possibly designed to quickly understand the age demographic of patients.

Dashboard 1 : Demographics



Total Patients (Card Visualization)

- **Purpose**: Quickly communicates the total number of patients (4973), an essential metric for assessing the healthcare service's reach and growth.
- **Insights**: Provides a benchmark for understanding patient base size over time.
- **Business Decisions**: Decide on resource allocation, expansion plans, or evaluate if patient acquisition strategies need adjustment.
- **Limitations**: Lacks context without historical or benchmark comparison data; currently isolated from other temporal indicators.

Male patients by Age Group (Pie Chart)

- **Purpose**: Illustrates the age distribution across patient demographics.
- **Insights**: With 32.13% as Elders and 67.87% as Adults, the chart highlights potential age-related service demands.
- **Business Decisions**: Plan services and staffing better for age-prevalent health issues. Consider developing age-specific programs or services.

- **Limitations**: Categories might be overly broad. Age brackets can be more detailed to provide nuanced insights.

Patient Count by Race (Donut Chart)

- **Purpose**: Displays patient distribution across racial groups.
- **Insights**: Understanding racial diversity aids in culturally competent care and targeting community outreach.
- **Business Decisions**: Tailor healthcare services and community engagement initiatives to cater to dominant or underserved racial groups.
- **Limitations**: Could benefit from normalized data comparing these counts to local demographics or trends.

Male Patients by City (Bar Chart)

- **Purpose**: Breakdown of male patients distributed by city.
- **Insights**: Helps identify urban areas with higher male patient counts for targeted campaigns.
- **Business Decisions**: Allocate resources geographically, optimize clinic locations, or enhance services in regions with higher male populations.
- **Limitations**: Focused on only one gender. A comparative chart with female patient data would provide more value.

Average Age by State (Line Chart)

- **Purpose**: Displays the average patient age per state.
- **Insights**: Provides a comparative view of age-related demographic trends across regions.
- **Business Decisions**: Adjust healthcare offerings based on regional demographic insights; states with higher average ages might need more geriatric services.
- **Limitations**: Limited view if regions outside the states listed are significant.

Average Age by City (Column Chart)

- **Purpose**: Highlights average age trends across cities.
- **Insights**: Analyzes specific urban centers' demographic makeup and trends.
- **Business Decisions**: Develop city-specific strategies for age-relevant medical needs and resources.
- **Limitations**: Disparate scales between charts; results might be misleading without data validation on sample representation.

General Dashboard Insights

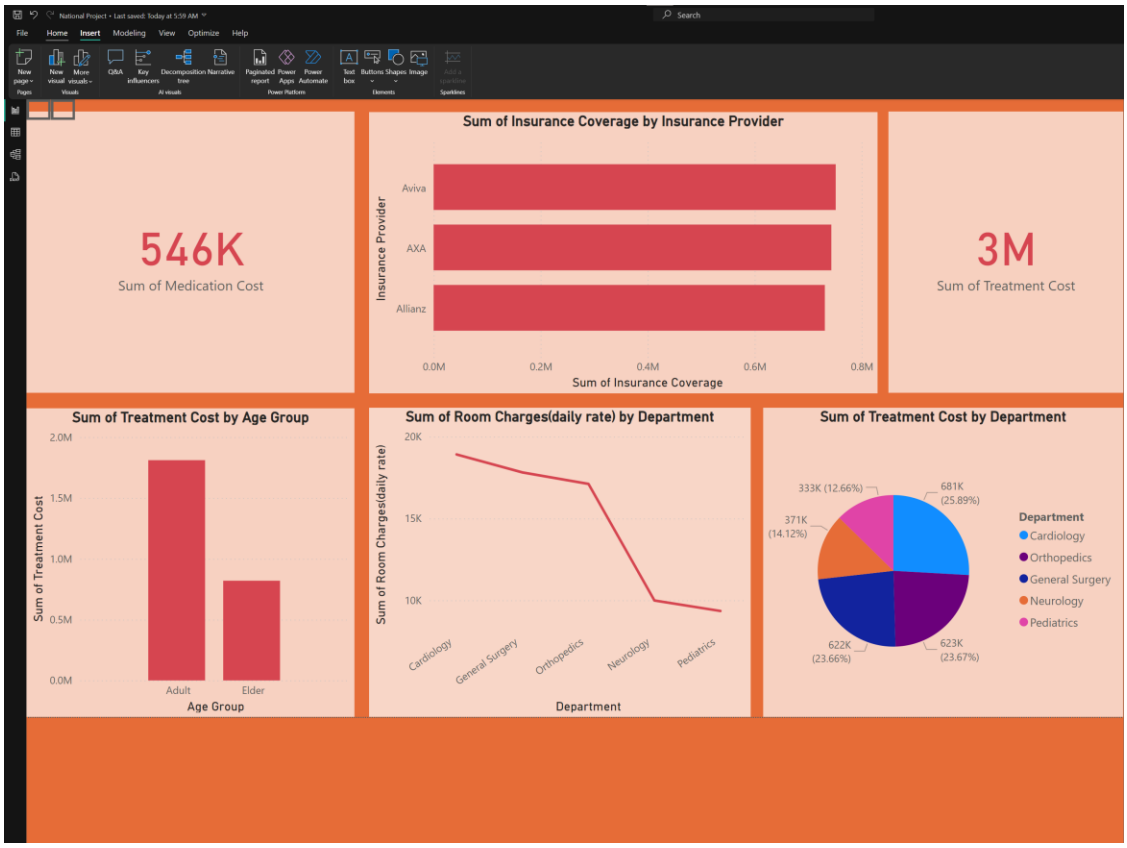
- **Highlighting Demographics**: The dashboard prioritizes demographic insights, aiding strategic planning and personalized care initiatives.

- **Enabling Resource Optimization**: With the geographic and age breakdowns, organizations can tailor resources where they're most needed.

Limitations

- **Lack of Temporal Context**: The absence of trends over time (e.g., historical comparison data) limits predictive analytics.
- **Data Detail**: Adding breakdowns (e.g., age bands within Adult/Elder) could enrich understanding.
- **Diverse Representation**: Including comparatives for female patients or expanding racial and geographic data can enhance inclusivity.
Financial Dashboard Visualizations and Analysis

Dashboard 2 : Financial analysis



Sum of Treatment Cost (Card Visualization)

- **Purpose**: Displays the total cost incurred for treatments, providing a quick overview of financial expenditure.
- **Insights**: This metric helps assess the financial burden of treatment services.

- **Business Decisions**: Evaluate budgeting, detect areas for cost reduction, and plan financial resources accordingly.
- **Limitations**: Lacks breakdown for deeper insights into specific cost drivers or comparison against previous periods.

Sum of Treatment Cost by Age Group (Bar Chart)

- **Purpose**: Analyzes the treatment cost distribution across different age groups (Adult vs. Elder).
- **Insights**: Indicates which age group incurs higher treatment costs, highlighting demographic cost differences.
- **Business Decisions**: Tailor financial planning and service pricing for demographics with higher medical expenses.
- **Limitations**: Limited to broad age groups which may obscure insights into more nuanced age cost trends.

Sum of Room Charges by department (Line Chart)

- **Purpose**: Highlights total room charges by department.
- **Insights**: Identifies which departments generate the most room charges, correlating with service utilization and length of stay.
- **Business Decisions**: Optimize room allocation, improve department efficiency, and evaluate room pricing strategy.
- **Limitations**: Needs integration with occupancy rates for comprehensive analysis.

Sum of Insurance Coverage by Insurance Provider (Bar Chart)

- **Purpose**: Illustrates the distribution of insurance coverage among different providers.
- **Insights**: Reveals reliance on specific insurers and their contribution to coverage.
- **Business Decisions**: Develop strategic partnerships with top insurers and negotiate terms for mutual benefit.
- **Limitations**: Might not account for claim success rates or payment timelines.

Sum of Treatment Cost by Department (pie Chart)

- **Purpose**: Compares treatment costs among departments.
- **Insights**: Demonstrates cost intensity of services provided by each department, key for cost management.
- **Business Decisions**: Assess and realign departmental budgets based on cost contributions; evaluate departmental profitability.
- **Limitations**: Lacks insightful breakdown on treatment types or patient outcomes.

General Dashboard Insights

- **Cost Allocation**: The dashboard focuses on consolidating cost data, aiding in financial oversight and strategic budgeting.
- **Insurance Strategy**: Insurance data supports understanding institutional reliance on insurance funding.
- **Departmental Performance**: Data provides perspective on departmental financial health and resource allocation strategies.

Limitations

- **Temporal Analysis**: Introducing time trends would provide deeper insights into whether costs are rising or in control year-over-year.
- **Outcome-Linked Costs**: Linking treatment costs to health outcomes could refine financial analysis by assessing cost-effectiveness.
- **Coverage and Claim Analysis**: Delve into actual insurance payouts and claim efficiency for comprehensive insurance evaluations.