POWER BI PROJECT (DATA CLEANING)

STEP 1: Data Set Selection And Download.

- Data Selected From the Site “Maven Data Challenges”On The Topic Of “Hospital Dashboard”.(It Focus On To Build a top-level KPI dashboard to help a hospital's executive team quickly understand its recent performance in key areas).
- Downloaded It As a CSV. File Named As Power BI Project.

STEP 2: Load Data into Power BI.

- Imported that CSV Dataset Into Power Query .

STEP 3: Clean the Data.

- Renamed Some Columns For Better Understanding.
- Changed Type Of Some Columns.
- Replaced Some Errors On The Some Columns.
- Capitalized Each Words In Text Columns.
- Removed Erros In Some Columns.
- Removed Duplicates In Some Columns.
- Added Custom Column Named “Patient Paid” By Substracting ‘Insurence Comp Paid’ From ‘Total Cost’.
- Reordered Columns For Better Understanding.
- Changed The Type Of ID Into Text.
- Renamed The ID In Appointment As Appointment ID.
- Split The Start Date And Time Column By Deliminator.
- Changed The Column Type Into Date.
- Renamed The Splited Columns.
- Split The Stop Date And Time Column By Deliminator.
- Changed The Column Type Into Date.
- Renamed The Splited Columns.
- Used 20 Dax measures And 6 Dax Column.

DATA MODELING

Star Schema Characteristics:

- **Central Fact Table:** `Appointments` acts as a central fact table. It contains foreign keys to several dimension tables (`Patients`, `Hospital`, `Payers`) and also includes measures.
- **Direct Relationships:** `Appointments` has direct, single-level relationships with `Patients`, `Hospital`, and `Payers`.

- This is a classic star schema structure where dimensions are directly linked to the fact table.(Appointments (Many) to Patients (One))

Snowflake Schema Characteristics:

- **Normalized Dimensions:** The `Procedures` table can be seen as a sub-dimension or a more granular fact table linked to `Appointments` and `Patients`. If `Procedures` were considered a dimension of `Appointments`, it would be a "snowflaked" dimension because it's not directly linked to the central fact `Appointments` in a flat, denormalized way, but rather through its own set of details.

VISUALIZING

Home Page (Cover Page)

- **Purpose**: Acts as a landing page and navigation guide.
- **Components**:
  - Hospital image, logo, and tagline: *"Healing Hands, Caring Hearts."*
  - Title: **City Hospital Dashboard**
  - Navigation Buttons: "Overview", "Cost&Profits", "Performance Overview", "Patient Flow & Cost Analysis", "Revenue & Patient Demographics"
- **Description Under Each Page**:
  - **Overview**: Key metrics, top procedures, total income & profits.
  - **Cost&Profits**: Cost & profit by payer, class, and year.
  - **Performance Overview**: Patient vs insurance payments, doctor-wise profits, costliest procedures.
  - **Patient Flow & Cost Analysis**: Monthly appointment heatmap, department cost, average patient cost.
  - **Revenue & Patient Demographics**: Insurance revenue, gender trends, and patient age over years.

Page 1: Overview Page

- **Purpose**: Provides high-level hospital KPIs (Key Performance Indicators).

- **Main Metrics**:
  - Total Patients: `974`
  - Total Doctors: `10`
  - Total Costs: `$113K`
  - Total Incomes: `$4,110K`
  - Total Profits: `$3,998K`
  - Total Departments: `6`
- **Top 15 Procedures Table**:
  - Shows procedure name, **average base cost**, and **procedure count**.
  - Example: *Ultrasound scan for fetal viability* done 184 times, avg. cost $9,205
- **Insight**: Quickly spot major revenue-driving procedures and overall hospital performance.

Page 2: Cost&Profits

- **Purpose**: Breaks down financial metrics by **payer** and **appointment class**.
- **Visuals**:
  - **Average Claim Cost by Payer**: Highest - *Medicare ($1,454K)*.
  - **Cost by Appointment Class**: *Ambulatory* class has highest cost ($2M+).
  - **Cost by Year Line Chart**:
    - Shows declining cost trend from 2011 to 2021.
    - Includes **unique patient counts** per year.
  - **Profits by Class**: Ambulatory generates highest profit ($2M).
- **Insight**: Identifies which payers and services contribute most to cost and profit over time.

Page 3: Performance Overview

- **Purpose**: Compares **patient-paid vs insurance-paid**, highlights **costliest procedures** and **doctor profit**.
- **Visuals**:
  - **Top 8 Costliest and Most Frequent Procedures**:
    - Includes procedures like COVID-19, drug overdose, and pregnancy.
    - Bar + Line combo chart shows both cost and frequency.
  - **Patient vs Insurance Comparison**:

        - Breaks down how much patients and insurance companies paid.

    - **Doctor-wise Profit Chart**:
        - Highest: Dr. Naseef Ahmed ($572K), followed by Dr. Beena Mathew ($489K).
- **Insight**: Identifies most expensive treatments and most profitable doctors.

Page 4: Patient Flow & Cost Analysis

- **Purpose**: Shows **seasonal trends, department-wise costs**, and **average patient cost**.
- **Visuals**:
    - **Monthly Appointment Heatmap**:
        - February and January have highest appointments.
        - October and December lowest.
    - **Total Cost by Department (Donut Chart)**:
        - Ambulatory: 50.47% of cost.
    - **Average Cost per Patient**: `$4.22K` (Gauge chart).
- **Insight**: Helps optimize staffing and resources based on seasonal trends and department demand.

Page 5: Revenue & Patient Demographics

- **Purpose**: Deep dive into **insurance revenue, gender trends**, and **patient age analytics**.
- **Visuals**:
    - **Revenue by Insurance Company**:
        - Medicare highest contributor.
    - **Gender Breakdown by Year**:
        - Stacked bar chart shows M/F patient counts over time.
    - **Average Patient Age Over Time**:
        - Trend line shows age rising from 33 to 52 over a decade.
- **Insight**: Reveals demographic patterns, aging patient base, and insurer impact.