'

# DEPI Power Bi Final Project

# Healthcare Dataset Analysis

**Team Members:**

- Joumana Mourad

- Menna Tallah wehdan

- Yousef Shaban

- Mohamed Sherif

## Technical Provider Company:

Alx New Horizons

## Technical Instructor:

Mohammed Abdallah

'

# Index

'

# Project Overview

## Objective:

The Healthcare Data Analysis project utilizes Power BI to analyze and derive insights from healthcare data. The project aims to uncover trends, patterns, and correlations within the data to improve decision-making and operational efficiency in healthcare organizations. By leveraging Power BI's visualization capabilities, we create interactive dashboards that provide valuable insights into patient demographics, treatment effectiveness, hospital utilization, and other key healthcare metrics.

## Deliverables:

- **Patient-centric analysis**: Insights on patient demographics, diagnosis, treatments, and medical expenses.

- **Hospital-centric analysis**: Insights into hospital management, recovery ratings, room costs, and doctor performance.

- **Dashboards**: Interactive visual dashboards for better data-driven decision making.

- **Data Quality Checks**: Handling of missing, inconsistent, or inaccurate data points.

'

# Data Sources

**Healthcare Data Set 1:** Patient-Centric Dataset

Overview of Dataset

- **Patient ID**: A unique identifier for each patient. This can be used as the primary key to connect other datasets or tables.

- **Patient Name**: This is typically considered personal identifiable information (PII) and may not be necessary for analysis.

- **Age**: Numeric value; can be used to create age groups for analysis.

- **Gender**: Categorical data (e.g., Male, Female, Other); helpful for demographic analysis.

- **Blood Type**: Categorical data (e.g., A+, B-, etc.); useful for medical analysis.

- **Diagnosis**: Describes the patient's medical condition.

- **Treatment**: Treatment or procedures provided for the diagnosis.

- **Admission Date**: Date the patient was admitted to the hospital.

- **Discharge Date**: Date the patient was discharged from the hospital.

- **Total Bill**: The total amount billed for the treatment.

- **Full Prescription Details**: Contains all the prescribed medications.

'

**Healthcare Data Set 2:** Hospital-Centric Dataset

Overview of Dataset

- **Patient ID**: A unique identifier for each patient. This can be used as the primary key to connect other datasets or tables.

- **Patient Name**: This is typically considered personal identifiable information (PII) and may not be necessary for analysis.

- **Age**: Numeric value; can be used to create age groups for analysis.

- **Gender**: Categorical data (e.g., Male, Female, Other); helpful for demographic analysis.

- **Blood Type**: Categorical data (e.g., A+, B-, etc.); useful for medical analysis.

- **Diagnosis**: Describes the patient's medical condition.

- **Treatment**: Treatment or procedures provided for the diagnosis.

- **Admission Date**: Date the patient was admitted to the hospital.

- **Discharge Date**: Date the patient was discharged from the hospital.

- **Total Bill**: The total amount billed for the treatment.

- **Full Prescription Details**: Contains all the prescribed medications.

'

# Data Preparation

## Data Cleaning:

1. Inconsistent Date Formats

   • **Issue**: The dates for AdmissionDate and DischargeDate were in different formats (e.g., MM/DD/YYYY, DD/MM/YYYY).

   • **Action**: Standardized all date fields to a consistent format (YYYY-MM-DD) to ensure accurate calculations and comparisons.

   • **Impact**: This improved data accuracy and allowed us to calculate critical metrics like LengthOfStay.

2. Extracting Key Information from Medication Details

   • **Issue**: The column containing information about the medicines prescribed (Full Prescription Details) was too detailed, making it hard to analyze.

   • **Action**: Parsed the prescription data, extracting only relevant medication names and doses for the analysis.

   • **Impact**: Simplified the data, making it easier to track medication usage and perform analysis on prescription patterns.

3.Creation of "LengthOfStay" Column

   • **Action**: A new column LengthOfStay was created by calculating the difference between the DischargeDate and AdmissionDate for each patient.

   • **Purpose**: This metric became essential for understanding hospital stay durations and correlating them with treatment costs and recovery times.

   • **Impact**: Helped in analyzing factors affecting longer or shorter hospital stays.

'

## 4. Query Division Using Power Query

- **Action**: In Power Query, the dataset was divided into multiple queries to handle various aspects of the data:
  - **Date Confusion Query**: Focused on resolving issues related to inconsistent date formats.
  - **Medication Details Query**: Handled the extraction and transformation of the medication information.
  - **Patient & Hospital Merge Query**: Ensured integration between patient-centric and hospital-centric data.

## 5. Merging the Datasets Using Power Query

- **Action**: The two datasets (Patient-Centric and Hospital-Centric) were merged using a full outer join on the PatientID column in Power Query.
- **Purpose**: This ensured that all patient records, even if they appeared only in one dataset, were included in the final analysis.
- **Impact**: Provided a comprehensive view of both patient and hospital data, ensuring no records were omitted.

## 6. Date Confusion Query

- **Action**: A specific query was created to address date formatting issues, ensuring that any inconsistencies were corrected before merging the datasets or performing further analysis.
- **Impact**: Guaranteed reliable time-based metrics, improving the accuracy of reports on admission durations and discharge rates.

## 7. Missing Patient Names

- **Issue:** Some records lacked PatientName values, which could lead to incomplete analysis or challenges in patient identification.
- **Action:** For these missing names, we either filled in the names based on other available records or flagged them as "Unknown" if no relevant data could be found.

'

- **Impact:** Ensured that each record was identifiable, reducing the chance of errors in reporting or patient-based analysis.

8. Handling Null Values in the Total Bill Column

- **Issue**: The TotalBill column had several null values, which could affect financial and cost analysis.

- **Action**: Replaced null values with the average TotalBill for similar cases (based on diagnosis or treatment type) or flagged the records for further review.

- **Impact**: This helped maintain accurate financial analysis, ensuring that calculations like average cost per patient or treatment weren't skewed by missing values.

9. Creating the "AgeGroup" Column

- **Action**: A new column AgeGroup was introduced to classify patients into four distinct categories:

- **Child**: Ages 0-12

- **Teenage**: Ages 13-19

- **Adult**: Ages 20-59

- **Senior**: Ages 60+

- **Purpose**: This allowed for more focused analysis by age group, which could be linked to diagnoses, treatments, and cost structures.

- **Impact**: Enabled the team to analyze trends and patterns across different age groups, improving insights into healthcare services for various demographics.

10. Missing Values in the Hospital Name Column

- **Issue**: Some records had missing values in the HospitalName field, leading to incomplete data on where patients were treated.

'

•        **Action**: For missing hospital names, we either identified the hospital based on other fields such as DoctorName or RoomNumber or flagged them for further review if no relevant data could be found.

•        **Impact**: Ensured that all records had valid hospital information, allowing for accurate hospital-based analysis and reporting.

11. Null Values in the Recovery Rating Column

•        **Issue**: The RecoveryRating column had several null values, potentially skewing analysis of patient outcomes and recovery trends.

•        **Action**:

•        Replaced null values based on similar records (e.g., same diagnosis, treatment type, or patient age group).

•        In cases where no comparable data was available, we left the records as missing but flagged them to exclude from recovery analysis.

•        **Impact**: By minimizing the number of null values in this column, the analysis of patient recovery rates became more accurate, ensuring insights into hospital and treatment performance were reliable.

'

# Data Transformation

### 1. Average Daily Cost

- **Purpose**: This measure calculates the average daily cost for room and therapy across all patients and hospitals.

- **Use Case**: Useful for understanding the general financial burden on patients per day and for comparing daily costs between hospitals or rooms.

### 2. Total Cost per Patient

- **Purpose:** This measure calculates the total cost for each patient by multiplying the daily cost with the length of stay (based on admission and discharge dates).

- **Use Case**: Helps in determining how much each patient was billed based on their stay in the hospital. Can be used for financial analysis and billing trends.

### 3. Average Length of Stay

- **Purpose:** This measure calculates the average number of days patients stay in the hospital.

- **Use Case:** Helps track hospital efficiency and patient turnover rates, allowing hospitals to optimize room usage.

### 4. Recovery Rating by Hospital

- **Purpose:** This measure calculates the average recovery rating for patients in each hospital.

'

- **Use Case:** Useful for assessing hospital performance and effectiveness of treatment. A high recovery rating suggests better patient outcomes.

5. Maximum Rating from the Recovery Rating Column

- **Purpose:** To identify the highest recovery rating achieved by patients, helping to assess overall patient satisfaction and treatment effectiveness.

- **Use Case:** Hospital administrators can use this measure to evaluate the quality of care provided and make informed decisions about improving treatment protocols.

6. Number of Doctors for a Hospital

- **Purpose:** To determine the total number of doctors practicing in a specific hospital, which can help evaluate staffing levels and resource allocation.

- **Use Case:** Hospital management can use this measure to assess whether they have sufficient medical staff to meet patient needs and to identify areas for recruitment.

7. Number of Adult,Child,Teenage,Seniors Patients

- **Purpose:** To calculate the total number of adult patients receiving treatment in the hospital, providing insight into patient demographics.

- **Use Case:** Healthcare analysts can use this measure to analyze trends in adult patient admissions and tailor services accordingly.

8. Number of Patients with Blood Type A- and Other Blood Types

'

- **Purpose:** To quantify the distribution of patients based on their blood types, particularly focusing on A- blood type for targeted resource management.

- **Use Case:** Blood banks and hospital emergency departments can utilize this measure to ensure adequate blood supply and prepare for surgeries or emergencies requiring specific blood types.

## 9. Average Length of Stay for a Hospital

- **Purpose:** To calculate the average duration patients stay in the hospital, aiding in operational efficiency assessments.

- **Use Case:** Hospital administrators can use this measure to evaluate their discharge processes and identify opportunities for reducing length of stay without compromising patient care.

## 10. Number of Rooms in a Hospital

- **Purpose:** To determine the total number of available patient rooms in a hospital, contributing to capacity planning and resource allocation.

- **Use Case:** Hospital management can use this measure to optimize room assignments, manage patient flow, and plan for expansions or renovations.

## 11. Number of Patients in a Hospital

- **Purpose:** To track the total patient count currently admitted to the hospital, providing insight into occupancy rates.

- **Use Case:** Hospital administrators can use this measure to manage resources effectively, ensuring that staffing and supplies meet the needs of current patients.

'

12.  Summary Table of Diagnoses and Associated Medications

•       **Purpose:** To create a comprehensive overview of diagnoses and their corresponding medications, filtered for blank entries, enhancing data clarity and usability.

•       **Use Case:** Healthcare providers can utilize this summary table to analyze medication trends associated with specific diagnoses, improving treatment planning and medication management.

13. Total Number of Non-Blank Medication Entries

•       **Purpose:** To count the total number of medications prescribed across multiple columns, indicating the level of pharmacological intervention in patient care.

•       **Use Case:** This measure can help pharmacy departments track medication usage patterns and support inventory management decisions.

'

Power BI Report Structure

1) Patients Info Dashboard

## 1.1. Diagnosis & Treatment Section

---

## A. Sankey Diagram (Diagnosis → Treatment → Medicine)

- **Purpose**: Visualize the flow of diagnoses to treatments and associated medicines.
- **Structure**:
    1. **Left Side (Diagnosis)**:
        - List of diagnoses: Asthma, Diabetes, Covid-19, Hypertension, etc.
        - These will be your source categories.
    2. **Middle (Treatment)**:
        - Show treatments such as Medication, Therapy, Surgery.
        - These are the intermediate categories.
    3. **Right Side (Medicine)**:
        - Specific medicines prescribed for each treatment (e.g., Furosemide, Amlodipine, Insulin).
        - These are your destination categories.
- **Flow**: The width of the lines connecting each element will represent the number of patients with that diagnosis, undergoing that treatment, and receiving those medicines.

## B. Count of Diagnosis (Bar Chart)

- **Purpose**: Show the number of patients diagnosed with each condition.
- **Structure**:
    1. **X-axis**: Represents the number of diagnoses for each condition.
    2. **Y-axis**: Lists the conditions such as Asthma, Diabetes, Covid-19, etc.
    3. **Bars**: Horizontal bars where the length represents the number of patients diagnosed with each condition.
    4. **Coloring**: Assign distinct colors for each diagnosis (e.g., blue for Asthma, green for Diabetes, red for Covid-19).
- **Top Left Section**:

‘

- o Position the **Sankey Diagram** to show the diagnosis → treatment → medicine flow. This should be large enough for users to follow the flow from left to right.
- **Below Sankey Diagram**:
  - o Place the **Bar Chart** to display the count of diagnoses for each condition. The bar chart should be visually distinct and easy to interpret.

## 1.2. Total Medication Section

---

**Medication Usage by Diagnosis (Stacked Bar Chart)**

- **Purpose**: To display the distribution of various medications used by patients with different diagnoses.
- **Structure**:
  1. **X-axis**: Represents different diagnoses such as Diabetes, Asthma, Hypertension, Covid-19, and Flu.
  2. **Y-axis**: Represents the total number of patients using medications.
  3. **Bars**: Each bar represents a diagnosis. The bar is divided into segments (stacked) where each color represents a specific medication.
     - For example, for **Asthma**, the bar might show portions for **Salbutamol**, **Furosemide**, **Prednisone**, etc.
  4. **Color Scheme**: Each medication gets a unique color. For instance:
     - **Red** for **Amoxicillin**
     - **Green** for **Azithromycin**
     - **Blue** for **Salbutamol**
     - **Purple** for **Insulin**
  5. **Data Insight**: Users can hover over each segment of the stacked bar to see exact values of the number of patients on each medication for each diagnosis.

## 1.3. Total Patients by Age Group and Diagnosis

---

**Bar and Waterfall Chart**

'

- **Purpose**: To visualize the number of patients in different age groups across various diagnoses, and to track the increase or decrease in patient counts for different conditions.

- **Structure**:
    1. **Bar Chart (Total Patients by Age Group)**:
        - **X-axis**: Represents the age groups: **Children**, **Adults**, **Seniors**.
        - **Y-axis**: Represents the total number of patients.
        - For each diagnosis (e.g., **Hypertension**, **Diabetes**, **Flu**, **Covid-19**), a separate bar will represent the number of patients in each age group.
    2. **Waterfall Chart (Increase/Decrease by Diagnosis)**:
        - This portion will show how the patient count has changed for each condition across different age groups.
        - For example, the waterfall chart could show:
            - **Hypertension (Adult)**: +442 patients
            - **Flu (Senior)**: -24 patients
            - **Covid-19 (Senior)**: +12 patients
        - The bars will visually illustrate how the number of patients has increased or decreased for specific conditions in certain age groups.
        - **Design Tips:**
        - The Bar Chart shows the distribution of patients by age group and diagnosis.
        - The Waterfall Chart will provide the increase/decrease in patient count for each diagnosis and age group, making it easy to spot trends.
        - Interactive Elements: Hovering over any bar or segment in both charts should display exact numbers to offer more granular data.

- **Key Insights Findings:**

**1. Diagnosis Patterns:**

- The **Sankey diagram** in the **Diagnosis & Treatment** section reveals the **most common diagnoses**, their corresponding **treatments**, and **medications**. This highlights the most frequent treatment pathways in the healthcare system. For instance:
    - **Asthma** is frequently treated with **therapy** and medications like **Furosemide** or **Amlodipine**.

'

- o **Diabetes** patients often require **medication** for management, while **Covid-19** cases may involve **therapy** and **medication**.

  This chart provides a comprehensive view of the **treatment journey** of patients from diagnosis to medication, highlighting the most common treatment methods.

## 2. Medication Use:

- The **Total Medication section** (Medication Usage by Diagnosis) shows the **distribution of medications** used for different conditions. This helps identify:
  - o Which **medications** are used most often for specific **conditions**, and how the usage might vary across diagnoses.
  - o Potential **over-reliance on specific drugs**, which could indicate areas for clinical improvement, such as exploring alternative treatments or investigating **drug effectiveness**.

## 3. Patient Age Group Trends:

- The **Total Patients by Age Group and Diagnosis** section highlights the **distribution of diagnoses** across various age groups. Some notable trends include:
  - o **Adults** have the highest occurrence of **Hypertension**.
  - o **Children** and **Teenagers** show higher rates of diagnoses for conditions like **Asthma** and **Flu**.
  - o **Seniors** tend to have more diagnoses related to **Chronic Diseases** like **Diabetes** and **Hypertension**.

    This data is crucial for healthcare planning, as it can help providers target prevention and treatment strategies based on age-specific trends.

- 

## 2) Hospital and Financial Costs Info Dashboard

### 2.1. Doctor Daily Cost and Total Bill

### Bar Chart: Doctor Daily Cost and Total Bill

This chart will display the daily costs incurred by each doctor, as well as the total bills generated for their patients. It will allow for an easy comparison of the two key metrics for each doctor, helping users understand the financial impact of individual physicians on the overall healthcare operation.

'

---

**Chart Design Breakdown**

1. **Y-axis**: Represents the **Doctors** (e.g., Barbara Gonzalez, James Rodriguez, etc.). Each doctor will have their own row.

2. **X-axis**: Represents the **cost values** (in dollars). The chart will show two distinct values for each doctor:

    o **Daily Cost** (light blue) – This represents the daily operational costs associated with each doctor.

    o **Total Bill** (dark blue) – This represents the total amount billed for the patients treated by that doctor.

3. **Bars**:

    o Each doctor will have two bars:

        ▪ The **first bar** (light blue) will represent the **daily cost** for that doctor.

        ▪ The **second bar** (dark blue) will represent the **total bill** generated from the patients under that doctor's care.

    o The **bar lengths** will reflect the monetary amounts. For example, if Barbara Gonzalez has a daily cost of $5k and a total bill of $15k, the light blue bar for her daily cost will be shorter than the dark blue bar representing her total bill.

4. **Data Example**:

    o **Barbara Gonzalez**:

        ▪ **Daily Cost**: $5,000 (light blue)

        ▪ **Total Bill**: $15,000 (dark blue)

    o **James Rodriguez**:

        ▪ **Daily Cost**: $7,000 (light blue)

        ▪ **Total Bill**: $20,000 (dark blue)

    o **Olivia Turner**:

        ▪ **Daily Cost**: $6,000 (light blue)

        ▪ **Total Bill**: $18,000 (dark blue)

5. **Color Scheme**:

    o **Light Blue** for **Daily Cost**

    o **Dark Blue** for **Total Bill**

'

---

**Layout Considerations:**

- **Bar Width**: Ensure that the bars are wide enough to clearly differentiate between daily costs and total bills for each doctor. The bars should be next to each other for each doctor but distinct in color.

- **Bar Placement**:
    - Arrange the doctors on the **Y-axis** in alphabetical order or based on some metric (e.g., highest daily cost or highest total bill).

- **Axis Labels**: Include the total cost value at the end of each bar for clarity. For example, the label for **Barbara Gonzalez's Daily Cost** might say **$5k**, and the label for **Total Bill** would say **$15k**.

- **Legends**:
    - Include a **legend** or a **color key** to clearly distinguish between the daily cost (light blue) and total bill (dark blue).

---

**2.2 Treatment Type and Average Total Bill**

This section will display two distinct charts that offer insights into **treatment types** and their associated costs, as well as the **relationship between the length of hospital stay** and the **total bill**.

---

**A. Donut Chart (Treatment Type)**

- **Purpose**: Visualize the distribution of different types of treatments and their associated **average daily cost**. This will provide an easy-to-understand breakdown of treatment types and how much they typically cost per day.

- **Structure**:
    1. **Slices**: Each slice of the donut chart will represent a different **treatment type**, such as:
        - **Counseling**
        - **Medication**
        - **Surgery**
        - **Physical Therapy**

2. **Size of Slices**: The size of each slice will be proportional to the **average daily cost** of each treatment. For example, if **Surgery** has a higher cost than **Medication**, the slice for Surgery will be larger.

3. **Labeling**:
   - Each slice will display the treatment name and the **average daily cost**. For example, **Medication - $1.04k**.
   - You can add percentages to each slice to show the relative cost distribution **Medication**: $1.04k/day

### B. Line Chart (Average Total Bill)

- **Purpose**: Show the relationship between the **length of stay** in the hospital and the **total bill**. This chart will help illustrate how patient bills increase as the length of their stay increases, allowing users to see patterns and trends over time.

- **Structure**:
  1. **X-axis**: Represents the **length of stay** in the hospital (e.g., in days). This could range from 1 day to 30+ days.
  2. **Y-axis**: Represents the **average total bill** in dollars. The values will increase as the length of stay increases.
  3. **Trend Line**: The chart will show a line representing the **average total bill** for each day or range of days that a patient stays. This will likely show an upward trend, with bills increasing as the length of stay extends.
  4. **Peaks and Troughs**: If there are certain time periods or hospital procedures that increase the cost significantly (e.g., ICU stays or complex surgeries), the line will show spikes.

1. **Line Chart Details**:
   - **Line**: A smooth curve showing the increase in total bills with length of stay.
   - **Markers**: Add small markers along the line to show the data points for each range (1-5 days, 6-10 days, etc.).
   - Include a **tooltip** for each marker that shows the exact **average bill** for that range of days.

'

       o **Trend Line**: Ensure that the trend line is clear, perhaps using a slightly thicker line for better visibility.

2. **Color Scheme**:

       o Ensure the color scheme is consistent with the rest of the dashboard.

       o Use subtle shades for the background and brighter, more vibrant colors for the charts to make the data stand out.

**Key Insights Findings:**

**Cost Insights:**

1) The **Doctor Daily Cost and Total Bill** chart reveals the significant variation in **daily doctor costs** across different doctors, as well as how these contribute to the total hospital bill. Some key takeaways include:

       o **High daily costs** do not necessarily correlate with the **total bill**, suggesting the importance of understanding how **doctor interactions** contribute to patient expenses.

  2) The **Treatment Type and Average Total Bill** section further reveals how different treatment types (e.g., **Medication**, **Surgery**, **Physical Therapy**) affect the **average daily cost** and overall **hospital bill**. For example:

       o **Surgical treatments** tend to increase **total costs**, while **medications** might be more affordable but may be associated with high usage rates.

3.

---

**3) Diagnosis and Treatment**

**3.1. Total Patients by Blood Type and Diagnosis**

This section will feature two charts that display the **distribution of patients by blood type** and the **association between blood types and diagnoses**. These charts will provide a quick understanding of how blood type might correlate with various health conditions.

‘

### A. Pie Chart (Blood Type Distribution)

- **Purpose**: To show the **distribution of patients based on blood types**. This chart will help users understand the proportion of patients with each blood type, offering insights into demographic patterns within the patient population.

- **Structure**:

   1. **Slices**: Each slice of the pie chart represents a different **blood type** (e.g., A+, O-, AB-, B+).

   2. **Size of Slices**: The size of each slice will be proportional to the **percentage of patients** with that blood type. For example, if **O+** is the most common blood type among patients, the slice for **O+** will be the largest.

   3. **Labeling**: Each slice will have:

      - The **blood type** name (e.g., A+, O-, etc.)

      - The **percentage** of total patients with that blood type.

   4. **Color Scheme**: Assign each blood type a distinct color:

      - **A+**: Light Blue

      - **O-**: Green

      - **AB-**: Orange

      - **B+**: Purple

      - And so on for other blood types.

### B. Heatmap (Blood Type & Diagnosis)

- **Purpose**: To show the **relationship between blood type and diagnosis**. This heatmap will visually represent how different blood types are associated with various diagnoses (e.g., Asthma, Covid-19, Diabetes, etc.).

- **Structure**:

   1. **Rows**: Represent different **blood types** (e.g., A+, O-, AB-, etc.).

2. **Columns**: Represent different **diagnoses** (e.g., Asthma, Diabetes, Covid-19, Hypertension, etc.).

3. **Cell Color Intensity**: The intensity of color in each cell will represent the **number of patients** with that specific combination of blood type and diagnosis.

    ▪ **Darker colors** (e.g., red) will indicate **higher numbers** of patients with that blood type and diagnosis.

    ▪ **Lighter colors** (e.g., pale blue) will indicate **fewer patients** with that combination.

4. **Color Gradient**: Use a **gradient scale** to represent patient counts. For example:

    ▪ **Dark Red**: High number of patients

    ▪ **Light Yellow**: Low number of patients

---

### 3.2. Average Stay and Length of Stay

This section focuses on visualizing the **average hospital stay** across different patient demographics, including **blood type** and **age group**. These bar charts will provide insights into how long patients typically stay in the hospital based on these two factors.

---

### A. Bar Chart (Average Stay by Blood Type)

- **Purpose**: To visualize the **average hospital stay** for patients with different blood types. This chart will allow users to compare the lengths of stay for each blood type, providing insights into whether certain blood types are associated with longer or shorter hospital stays.

- **Structure**:

    1. **X-axis**: Represents different **blood types** (e.g., O+, A-, B+, AB-, etc.).

    2. **Y-axis**: Represents the **average length of stay** in days.

    3. **Bars**: Each bar will represent the **average stay** for a specific blood type. The height of each bar corresponds to the **average number of days** patients with that blood type stay in the hospital.

'

4. **Color Coding**: Use different colors for each blood type to differentiate between them:

    - **O+**: Light Blue

    - **A+**: Green

    - **B+**: Orange

    - **AB+**: Red

    - And so on.

5. **Labeling**:

    - Add the **average number of days** at the top of each bar to give users precise information.

    - 

## B. Bar Chart (Average Length of Stay by Age Group)

- **Purpose**: To visualize the **average length of stay** for patients across different **age groups**. This chart will show how the age of a patient correlates with their hospital stay length, helping to highlight which age groups tend to have longer or shorter stays.

- **Structure**:

    1. **X-axis**: Represents different **age groups** (e.g., Senior, Child, Teenager, Adult).

    2. **Y-axis**: Represents the **average length of stay** in days.

    3. **Bars**: Each bar will represent the **average stay** for a specific age group. The height of each bar will correspond to the **average number of days** spent in the hospital for that age group.

    4. **Color Coding**: Use distinct colors for each age group to make the chart more visually appealing:

        - **Seniors**: Dark Blue

        - **Children**: Light Blue

        - **Teenagers**: Yellow

        - **Adults**: Green

'

5. **Labeling**:
   - Add the **average number of days** at the top of each bar for quick reference.

### 3.3. Number of Diagnoses by Age Group

This section will feature a **clustered bar chart** that helps visualize the **number of diagnoses** for each age group, broken down by specific conditions. This will allow users to easily compare how different age groups are affected by common diagnoses.

---

### Clustered Bar Chart Design

- **Purpose**: To display the **number of diagnoses** across different **age groups** for multiple health conditions (e.g., Asthma, Covid-19, Diabetes, Flu, Hypertension). This chart will allow users to compare the prevalence of various conditions within different age groups.
- **Structure**:
  1. **X-axis**: Represents the different **age groups** (e.g., Adult, Senior, Child, Teenager). Each age group will be shown as a category on the X-axis.
  2. **Y-axis**: Represents the **number of diagnoses** (e.g., the total number of patients diagnosed with each condition).
  3. **Bars**: For each age group, there will be multiple **clustered bars** representing different diagnoses (e.g., Asthma, Covid-19, Diabetes, Flu, Hypertension).
     - Each bar within a cluster corresponds to a **specific diagnosis** for that age group.
     - The **height** of each bar will represent the **number of patients** diagnosed with that condition within the age group.
  4. **Color Coding**: Use distinct colors for each diagnosis to help differentiate between the conditions:
     - **Asthma**: Light Blue
     - **Covid-19**: Green
     - **Diabetes**: Orange
     - **Flu**: Red
     - **Hypertension**: Purple
  5. **Labeling**:
     - Each bar will have a **label at the top** to show the **exact number of diagnoses** for that condition and age group.

'

---

**Key Insights Findings:**

These insights summarize the key patterns and trends identified from the various sections of the **Patients Info Dashboard**. The data visualizations in the dashboard provide valuable information for improving patient care, hospital efficiency, and cost management.

---

**5. Blood Type Correlations:**

- The **Blood Type & Diagnosis section** uses a **heatmap** to show how certain **blood types** are more prone to specific conditions. Key insights include:
  - Certain blood types, like **O+** or **A-**, may have a higher prevalence of conditions like **Asthma** or **Diabetes**.
  - This information can help **personalize treatment plans** by identifying specific **blood type correlations** with diseases, improving diagnostic accuracy.

'

**Future Enhancements for Healthcare Dashboard**

### 1. Predictive Analytics Integration

- **Objective**: Implement machine learning models to predict patient outcomes, such as the likelihood of a diagnosis based on age, blood type, and medical history.
- **Impact**: This will allow healthcare providers to identify potential high-risk patients early, optimize treatments, and reduce hospital readmission rates.

### 2. Real-Time Data Integration

- **Objective**: Integrate real-time data feeds from hospitals and health centers to update the dashboard dynamically.
- **Impact**: Real-time monitoring will enable decision-makers to act on the most current data, improving patient care and resource allocation.

### 3. Patient Outcome Tracking

- **Objective**: Add a section dedicated to tracking long-term patient outcomes after treatment (recovery, relapse rates, etc.).
- **Impact**: This will help healthcare providers assess the effectiveness of treatments over time and refine their approaches based on real-world outcomes.

### 4. Enhanced User Interaction

- **Objective**: Implement more interactive features like drill-downs, clickable charts, and patient-specific detail pop-ups to make the dashboard more intuitive for healthcare professionals.
- **Impact**: Enhanced interactivity will provide deeper insights with minimal navigation, improving user experience and decision-making.

'

### 5. Comparison Across Hospitals

- **Objective**: Add comparative analysis features that allow hospitals to benchmark their performance against others in terms of diagnosis rates, treatment costs, and patient outcomes.
- **Impact**: Benchmarking will help healthcare administrators identify areas for improvement and adopt best practices from top-performing hospitals.

### 6. Mobile-Friendly Design

- **Objective**: Optimize the dashboard for mobile access, enabling healthcare professionals to view reports and insights from any device.
- **Impact**: This will increase accessibility, especially for healthcare professionals who need insights on the go.

### 7. Patient Satisfaction Analysis

- **Objective**: Incorporate patient feedback and satisfaction scores into the dashboard to correlate outcomes with patient experiences.
- **Impact**: Understanding patient satisfaction will help healthcare providers improve service quality and patient care.

### 8. Integration with Wearable Health Data

- **Objective**: Allow data from wearables (e.g., smartwatches, fitness trackers) to be integrated into the dashboard for a more comprehensive view of patient health.
- **Impact**: Wearable data could provide early indicators of health issues, allowing for preemptive medical intervention.

# Conclusion

The current Power BI dashboard provides significant insights into patient demographics, diagnosis trends, treatment costs, and healthcare outcomes. It enables healthcare professionals to make data-driven decisions, optimizing patient care, resource allocation, and financial performance. The inclusion of age, diagnosis, treatment types, and hospital-specific information gives a clear view of healthcare dynamics.

However, the potential for future enhancements is vast. By integrating predictive analytics, real-time data, and more interactive features, the dashboard can become an even more powerful tool for healthcare providers. With the ability to predict patient outcomes, track performance across multiple hospitals, and improve user accessibility, this dashboard can evolve into a comprehensive decision-support system that enhances patient outcomes and operational efficiency.

The focus on future enhancements like mobile access, real-time integration, and predictive capabilities will ensure that healthcare organizations remain agile, proactive, and patient-centered in a rapidly evolving field.