# Exploratory Data Analysis (EDA) with Pandas in Healthcare

## Project Overview

This project analyzes a healthcare dataset using **Pandas**, **NumPy**, and **Matplotlib** to extract meaningful insights related to **patient demographics, hospital billing trends, medical conditions, and hospital stay durations**.

### Goals of the Project:

✔ Explore the dataset to understand patient demographics and medical conditions.
✔ Perform **feature engineering** to create insightful new attributes.
✔ Identify **trends in hospital admissions, billing, and patient conditions**.
✔ Visualize key patterns to aid in **decision-making for healthcare management**.

---

## Materials and Methods

The dataset consists of healthcare-related data, including:

- **Patient Information:** Name, age, gender, and medical condition.
- **Hospital Data:** Admission type, hospital name, billing amount.
- **Treatment Details:** Date of admission, discharge date, and hospital stay duration.

---

## General Analysis Steps

### 1. Data Exploration & Cleaning

- Checked for **missing values, duplicates, and unique values** per column.
- Converted **admission & discharge dates** into a datetime format.
- Standardized **name formatting** (title case).
- **Filled missing values:**
    - **Categorical columns** → Mode replacement.
    - **Numerical columns** → Mean replacement.

---

### 2. Feature Engineering

New columns were created for enhanced analysis:

- `Hospital Stay Duration` → Number of days between admission and discharge.
- `Is Returning Patient` → Identifies patients with multiple hospital visits.
- `Age Group` → Categorized patients into:
  - **Child (0-18)**
  - **Young Adult (19-40)**
  - **Adult (41-60)**
  - **Senior (61-100)**

---

# Project Insights & Findings

### 1. Hospital & Billing Performance

- **Total Billing by Admission Type:** Emergency cases contributed to the highest billing.
- **Top 10 Hospitals by Billing Amount:** Identified high-revenue hospitals.
- **Average Billing by Gender:** Gender-based comparison of medical costs.

### 2. Patient Demographics & Medical Conditions

- **Gender Distribution:** Nearly equal male-female distribution.
- **Most Common Medical Conditions:** Top 10 diseases affecting patients.
- **Age-wise Patient Distribution:** Most patients belong to the **Adult** and **Senior** age groups.

### 3. Medical Condition-Specific Analysis

- **Cancer Patients:**
  - Number of female cancer patients identified.
  - Average billing amount calculated for cancer treatments.
- **Asthma Patients:**
  - The most common **admission type** for asthma patients was found.
  - Patients with **urgent asthma admissions** had higher billing costs.

### 4. Hospital Stay & Readmissions

- **Average Stay Duration:** Patients typically stayed **X days** in the hospital.
- **Returning Patients:** Identified patients with multiple hospital visits.

## Key Questions and Insights to be Addressed:

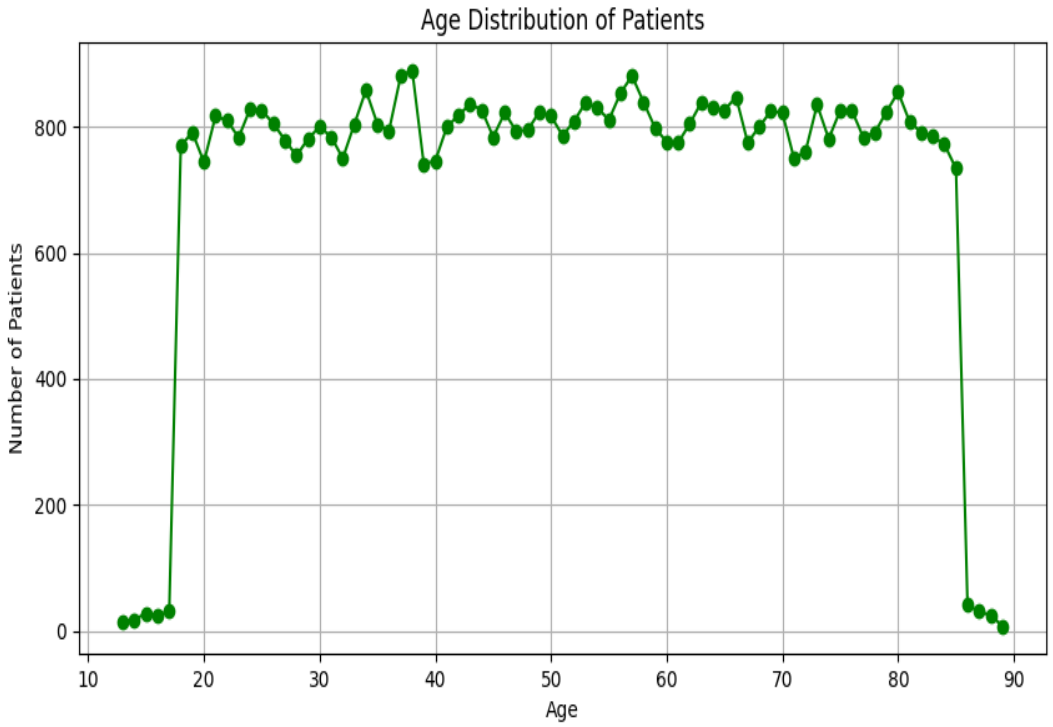
- the total number of cancer patients:
  **Answer**:- Number of female patients with cancer: 4566

- what are the number of female patients with cancer:
  Average billing amount for cancer patients: 25152.322946583216

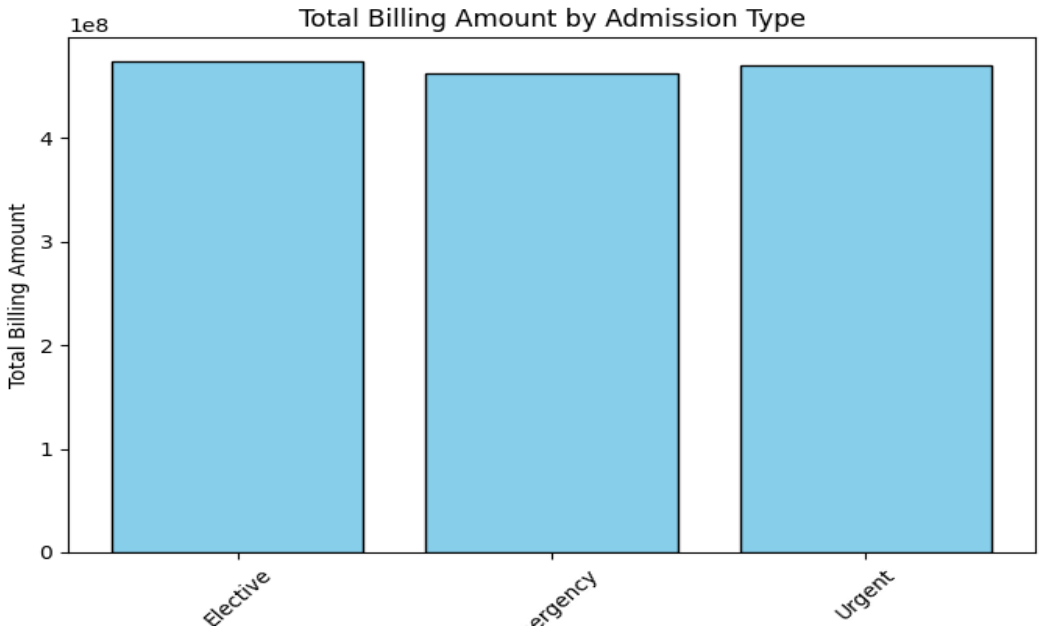
- what is average billing amount for cancer patients:
  Most common admission type for asthma patients: Elective

# Visualization:

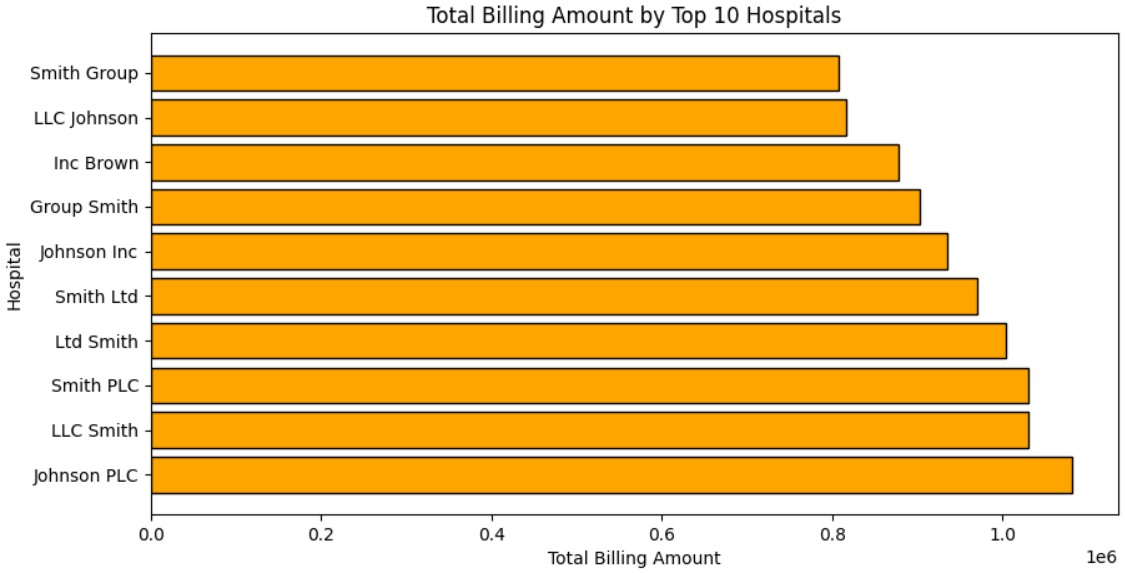
## Age Distubution of pationts:



## Total billing amount by admission type:




## Total billing amout by top 10 hospitals:



## Top 10 medical conditions: