# 🏥 Hospital Patient Record Analysis

## 📌 Objective

The objective of this project is to explore and analyze synthetic hospital patient data to extract meaningful insights that can inform hospital operations, patient care strategies, and resource allocation.

## 📊 Dataset Overview

- **Total records**: 1,000,000
- **Columns**: 25 attributes including demographics, medical history, diagnosis, treatment, billing, and hospital details.

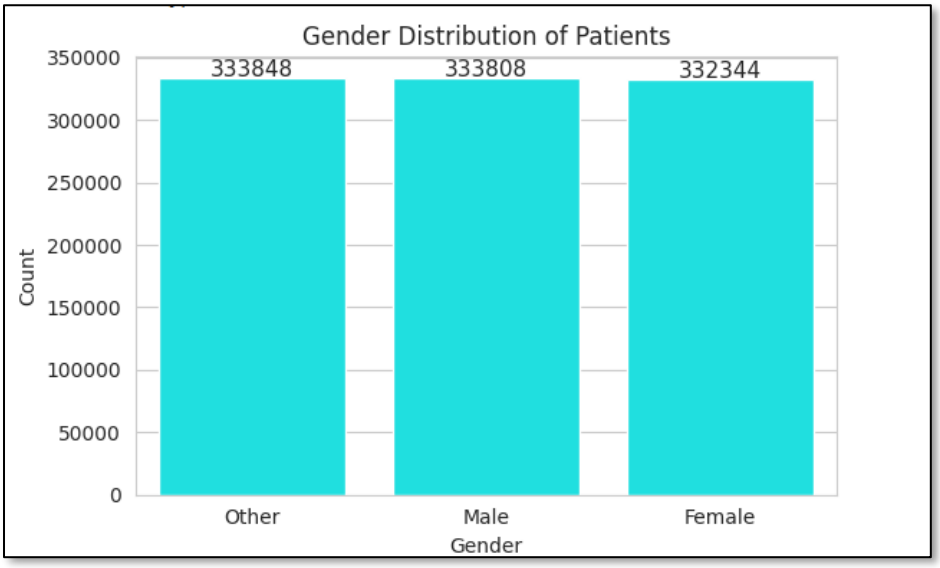
## 🔍 Exploratory Data Analysis (EDA)

### 1. Average Age of Patients

- **Average Age**: ~50.49 years
- Majority of patients fall in the middle-aged category, indicating a high demand for chronic illness management.

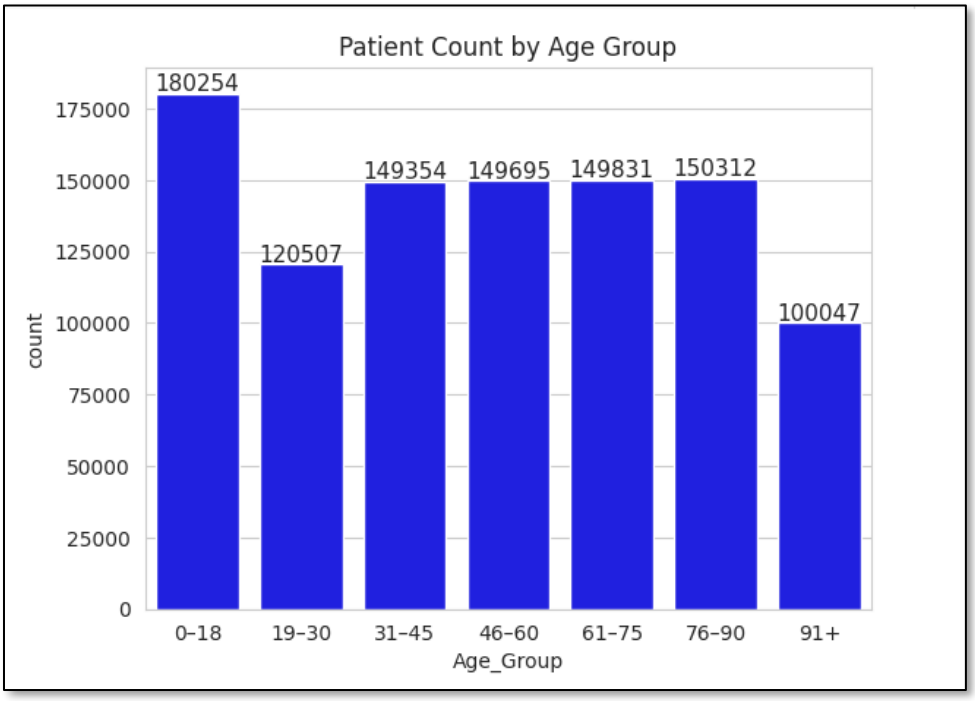
### 2. Gender Distribution

- **Male**: ~33.38%
- **Female**: ~33.23%
- **Other**: ~33.38%



### 3. Most Common Age Group

- **0-18 years** has the highest hospital admissions.
- **Older adults (60+)** account for nearly 45% of all hospitalizations.

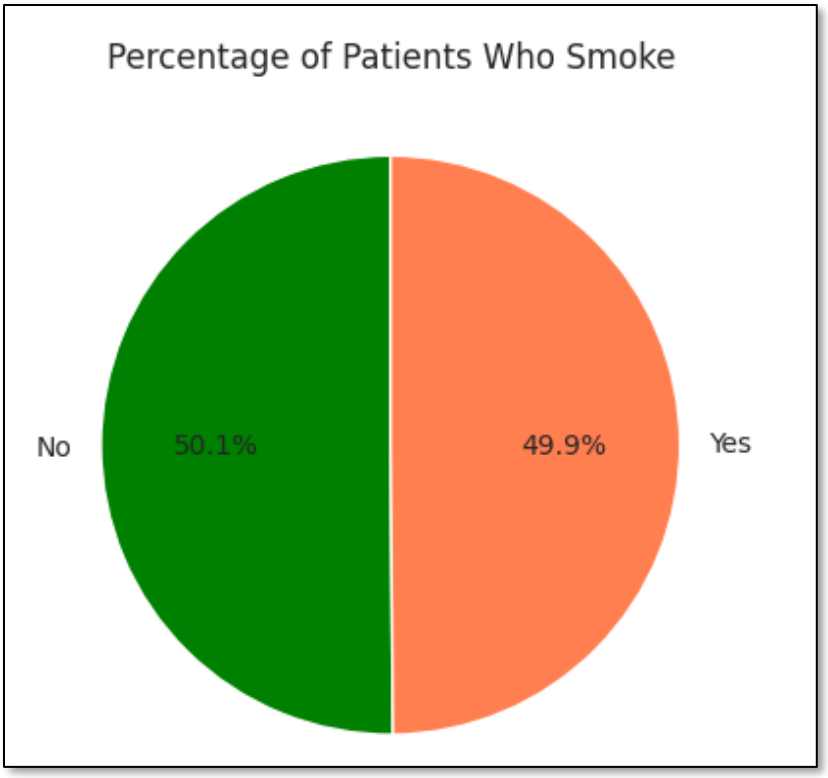



### 4. Medical History

- **Yes**: 50.03%
- **No**: 49.97%
- Half of the patients have a medical history, reinforcing the need for continuous care and monitoring.

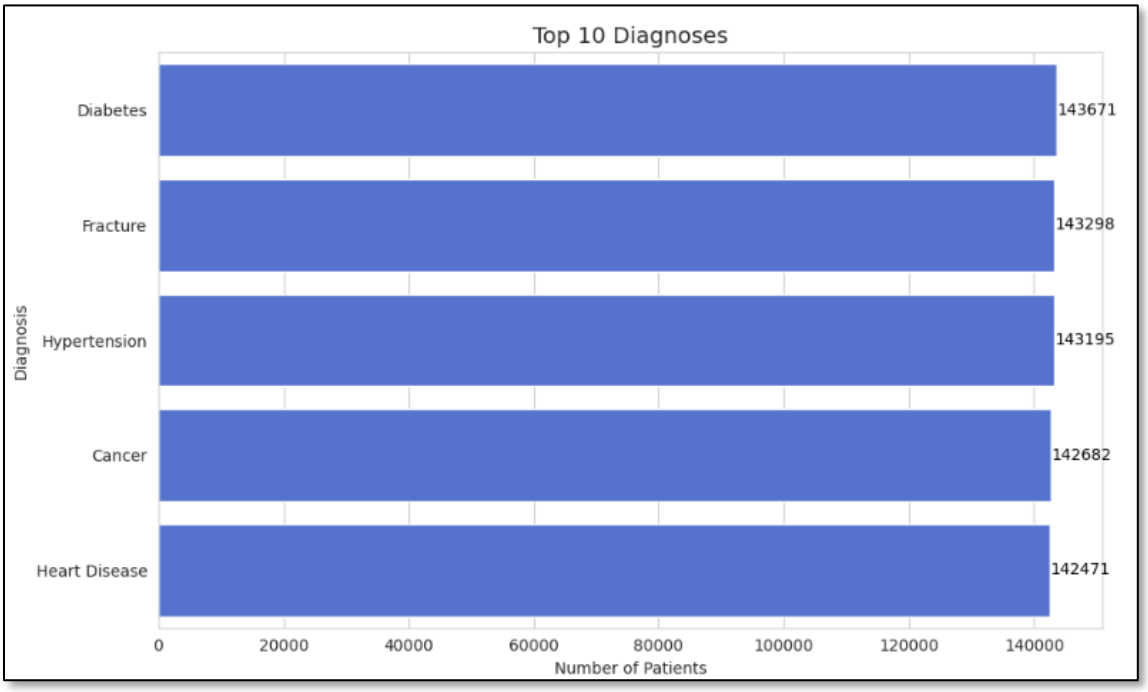
### 5. Smoking Status

- **Smokers**: 49.9%
- **Non-Smokers**: 50.1%
- Balanced distribution allows investigation of lifestyle-related health impacts.



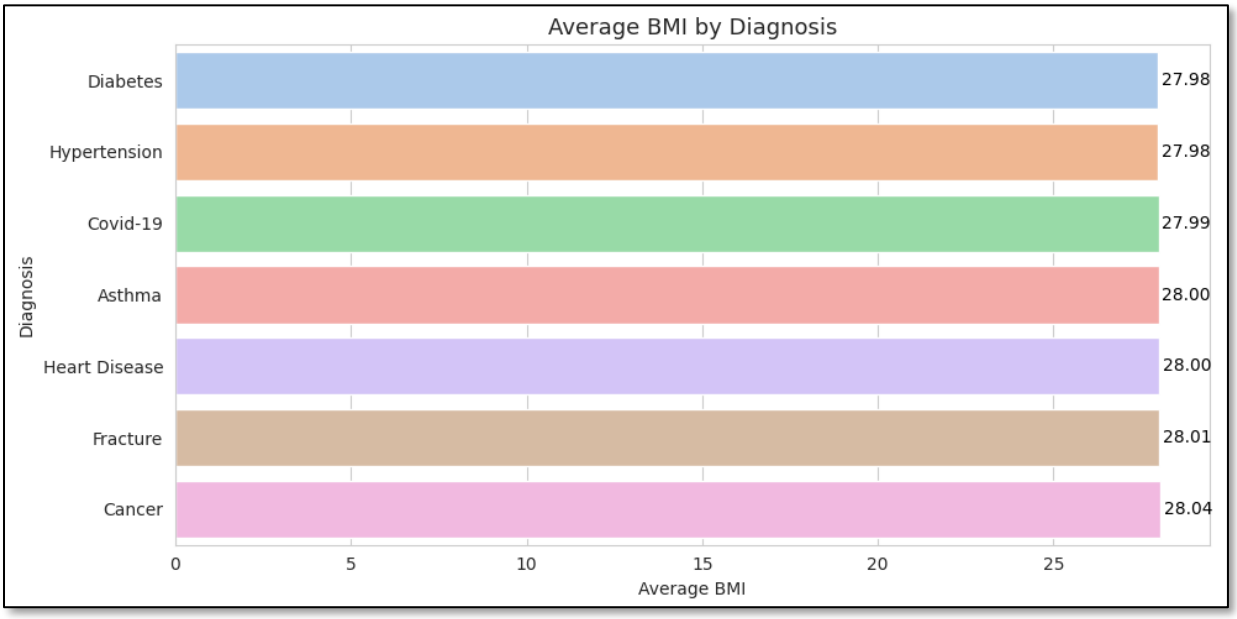

**6. Top 5 Diagnoses**

- Diabetes, Fracture, Hypertension, Cancer, Heart Disease.
- Reflects a mix of chronic conditions and emergency care cases.



### 7. BMI vs Diagnosis

- Average BMI varies slightly across diagnoses (27.97 to 28.04).
- No strong correlation found between BMI and specific conditions.

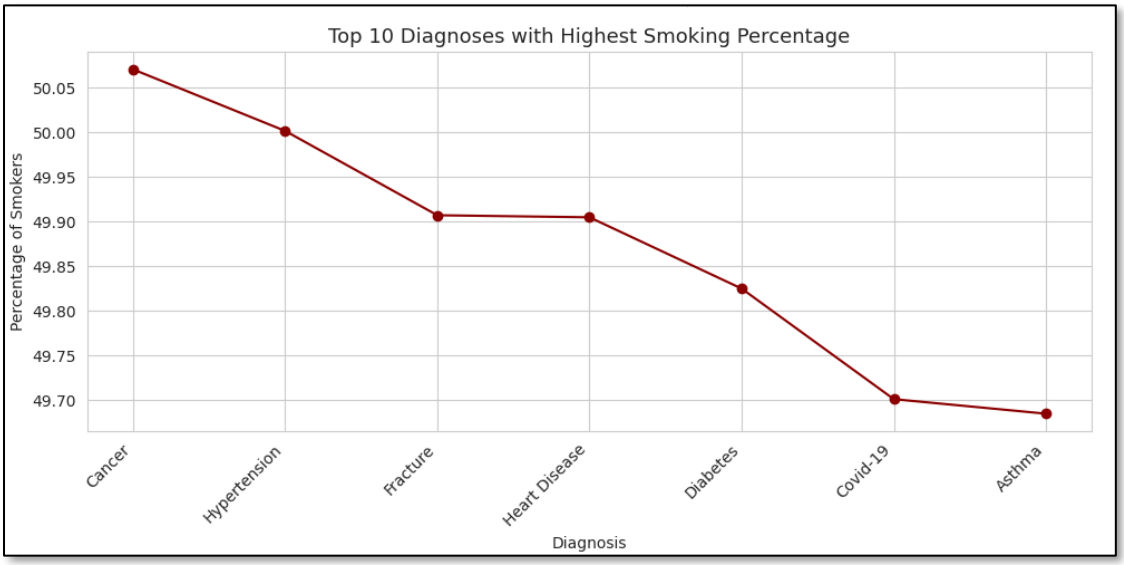



### 8. Hospitalization Duration by Gender

- All genders have an average stay around 182 days.
- No significant variation by gender.

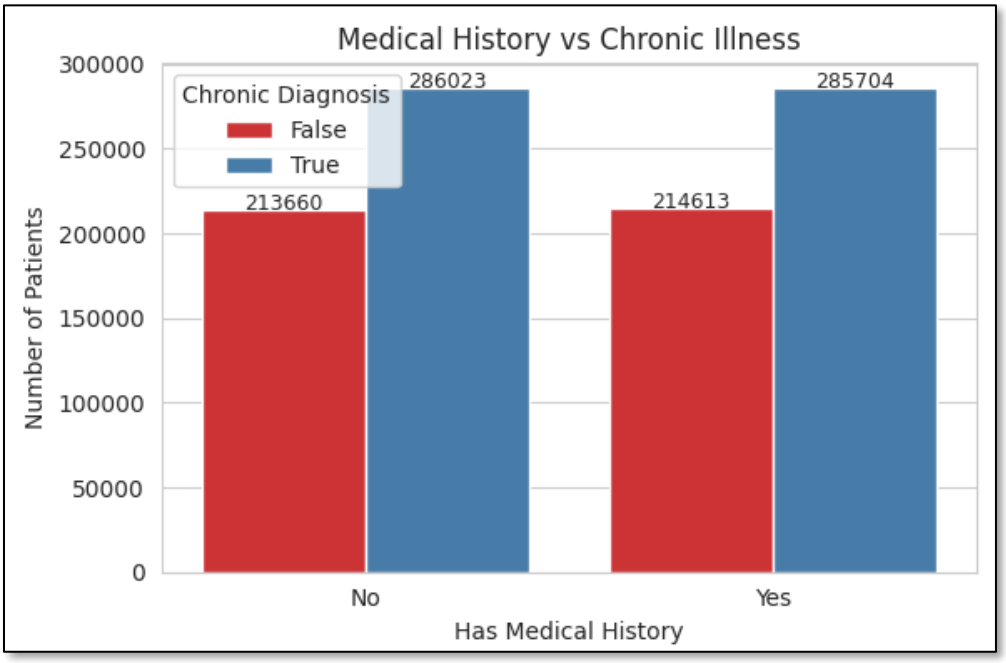
### 9. Smoking Status vs Diagnosis

- Smokers are more prone to heart disease, cancer, and respiratory conditions.



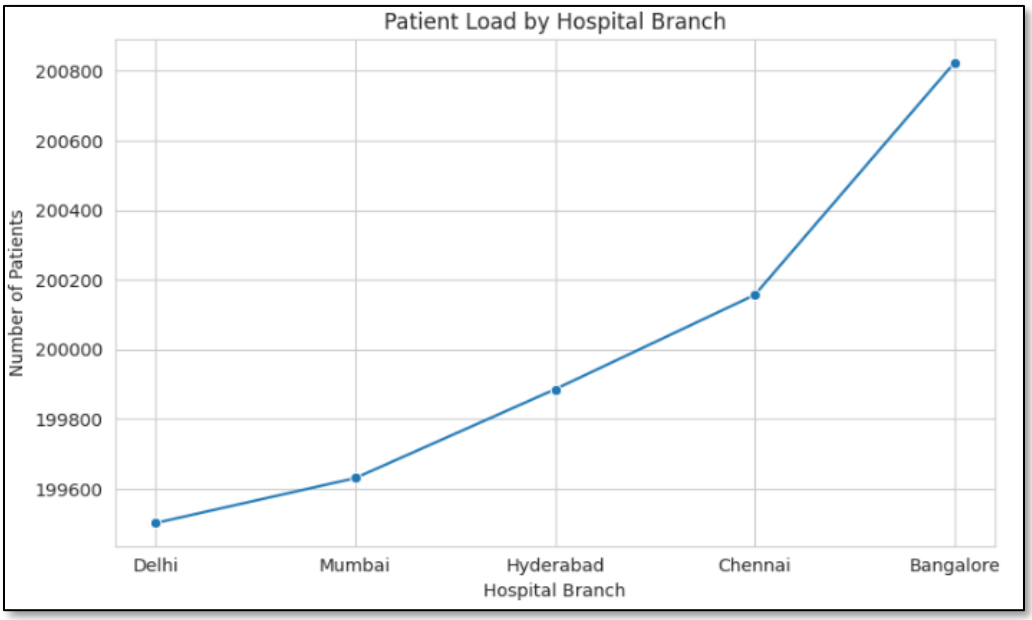
### 10. Medical History vs Diagnosis

- Both patients with or without a medical history can be diagnosed with chronic illnesses like diabetes and hypertension.




### 11. Hospital Branch Load

- **Bangalore** has the highest patient load.
- All branches handle close to 200,000 patients, indicating balanced operations.

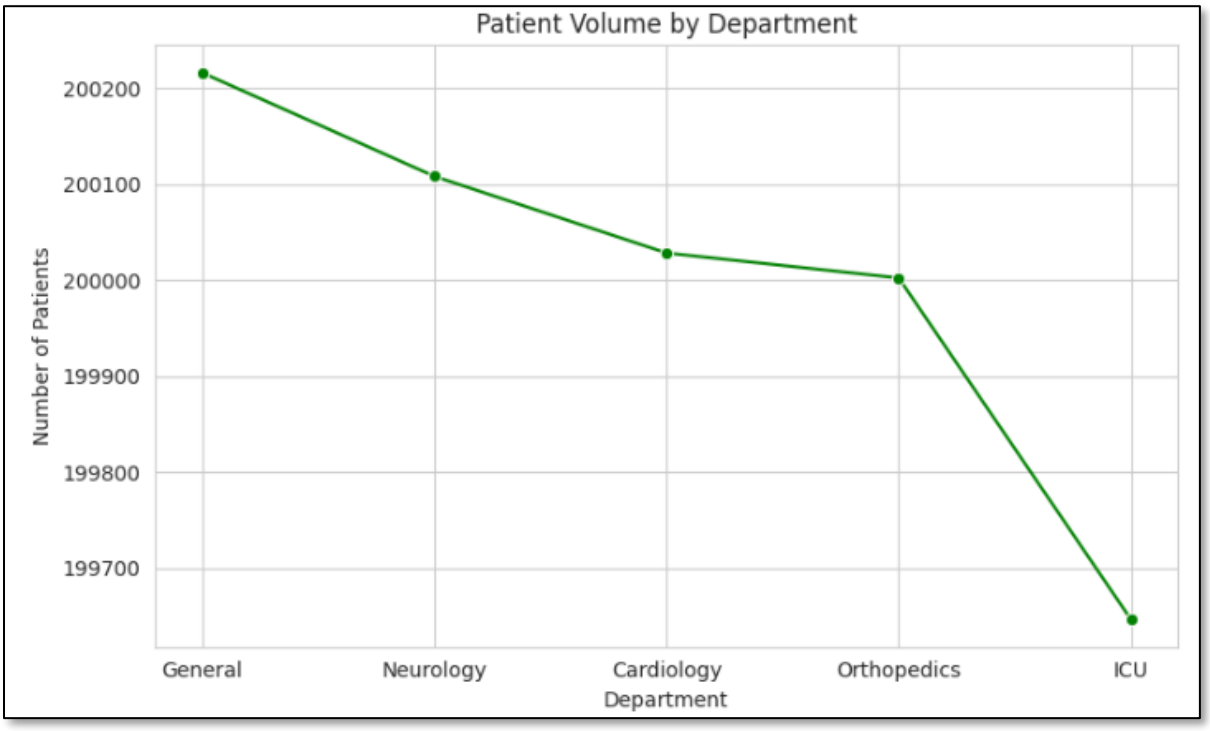


**12. Average Length of Stay**

- **182.02 days** — unusually high and may reflect long-term care cases or synthetic data design.

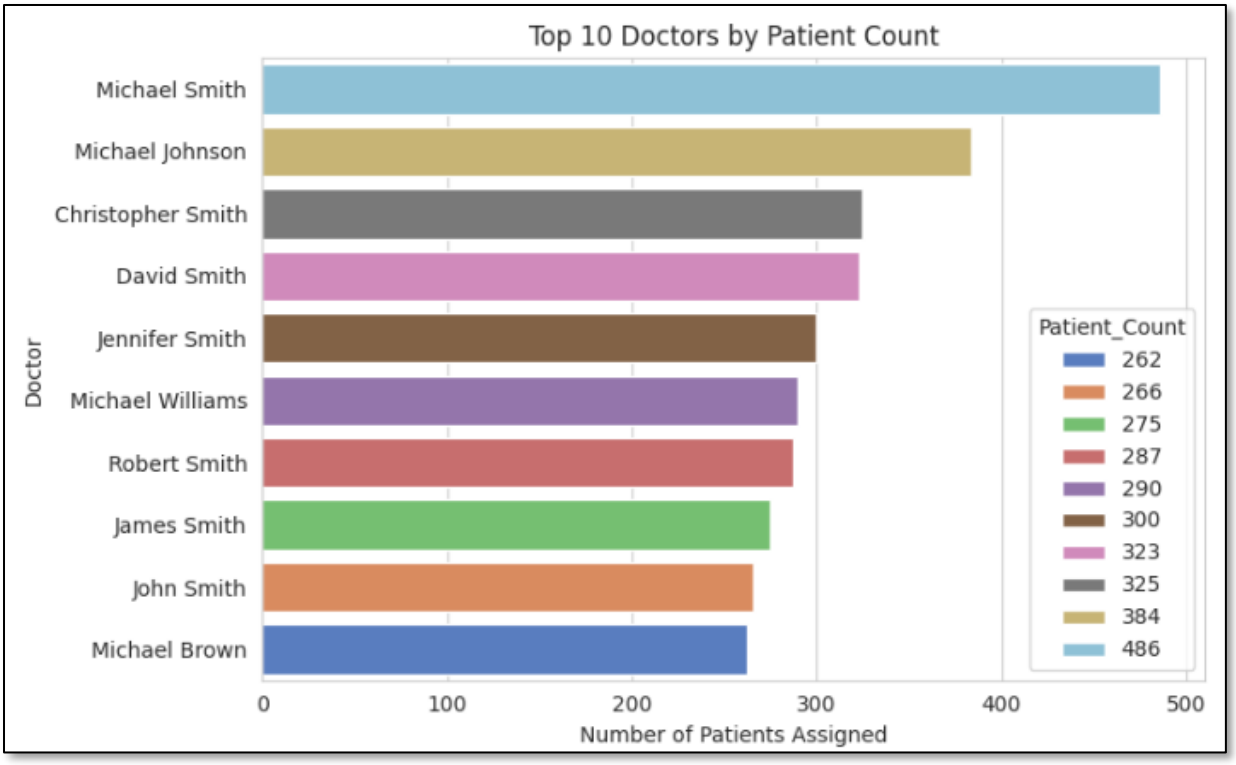
**13. Department Load**

- **General Department** sees the highest patient volume.
- Neurology, Cardiology, and Orthopedics are also heavily utilized.




**14. Doctor Assigned**

- **Michael smith** has received most patients.




## 📣 Conclusion

This project demonstrates how data analytics can uncover meaningful patterns in hospital operations and patient health trends.

Key insights include:

- The **General department** handles the highest patient volume.
- **Diabetes** and **Hypertension** are among the most common diagnosis.

Data-driven healthcare can significantly improve **efficiency, preparedness, and patient outcomes**.