





***Project name:*** Cara Metrics.

***Project description:***

The Healthcare Data Insights project aims to extract valuable insights from the comprehensive healthcare dataset. By analyzing patient data, we seek to uncover trends, patterns, and associations that can inform healthcare decision-making and improve patient outcomes.

***Project objectives:***

1. Identify Healthcare Trends: Analyze data to uncover trends in healthcare utilization, disease prevalence, and treatment outcomes.

2. Optimize Resource Allocation: Provide insights for better management of staffing, bed capacity, and equipment.

3. Enhance Patient Care: Identify areas to improve patient care, reduce readmissions, and boost patient satisfaction.

4. Support Evidence-Based Decisions: Offer data-driven insights for healthcare professionals and policymakers to make informed decisions.

5. Improve Hospital Efficiency: Recommend ways to enhance operational efficiency and reduce healthcare costs.

***Methodology:***

**Data Overview:**

The dataset contains 55,501 rows and 15 columns, covering patient demographics, medical conditions, admission details, medications, and financial information.

**Data Collection:**

The data was sourced from Kaggle and contained structured healthcare data involving patients' medical and administrative records.

Data Cleaning:

Identify and handle missing or inconsistent values (e.g., incomplete patient records or missing discharge dates).

Standardize categorical variables such as gender and blood type.

Correct formatting inconsistencies in numerical fields like billing amounts



**Data Exploration:**

Use descriptive statistics to summarize patient demographics (age, gender) and health conditions.

Explore correlations between variables (e.g., length of stay and medical conditions).

Visualize key metrics such as common medical conditions, medication usage, and admission types.

**Advanced Analysis:**

Apply predictive models to analyze factors influencing patient outcomes (e.g., readmission rates, cost of care).

Clustering will be conducted to identify patient groups based on their medical and treatment history.

**Data Visualization:**

Generate charts and graphs to present trends (e.g., distribution of admission types, and test results).

***summary***

*Similar prevalence across genders:*

There is no significant difference between the number of females and males for each medical condition.

This suggests that the prevalence of these medical conditions is relatively balanced between the sexes.

*There is no significant difference by blood type:*

The counts of each medical condition (e.g., arthritis, asthma, cancer, diabetes, hypertension, and obesity) are quite similar across different

blood types. There are no very large differences that are prominent, suggesting that blood type may not play a major role in determining susceptibility to these conditions.

Most common conditions: The conditions with the highest counts among all blood types are diabetes and hypertension, suggesting that these conditions are generally more prevalent in the sample population.

Highest and lowest counts:

Diabetes: Blood type B+ has the highest count (1196), while blood type O- has the lowest count (1122).



High blood pressure: Blood type AB+ has the highest number (1215), while blood type O- has the lowest number (1145). Asthma and obesity: These conditions are relatively evenly distributed among all blood types.

*The analysis identifies the hospitals with the highest and lowest average billing amounts*.

The results show a clear distinction between the most and least expensive hospitals based on their average billing charges.

These hospitals have the highest billing amounts, with Hernandez-Morton being the most expensive. The average charges for these hospitals are all above $52,000.

Least Expensive Hospitals:

Rowe, Stone, and Patterson: $49,450.12

*Summary of Infection Trends by Age and Medical Condition:*

1. Cancer:

Highest infection rates are observed in individuals in their mid-fifties.

A significant concentration of infections occurs from the early to late fifties.

Early fifties infections show a correlation with obesity.

2. Diabetes:

Most infections are concentrated in individuals in their late thirties and late forties.

3. Asthma:

Infection rates peak in the mid-twenties and early fifties.

4. High Blood Pressure:

Infection rates rise in the early twenties and again in the early seventies, with noticeable increases also in the late thirties.

5. Joint Disease:

Infection rates are relatively consistent across most ages, but there is a noticeable increase in the late thirties and late fifties.



*The analysis shows how different medical conditions are distributed across three types of hospital admissions: Elective, Emergency, and Urgent.*

Elective Admissions: For elective admissions, conditions such as cancer and diabetes have relatively high occurrence rates. This indicates that many patients are likely planning their hospital stays for treatments or management of chronic conditions.

Emergency Admissions: Emergency admissions see high occurrences for conditions like asthma and hypertension. This suggests that these medical conditions often require urgent intervention, potentially due to sudden complications or exacerbations.

Urgent Admissions: Urgent admissions also have a notable number of cases for conditions such as obesity and diabetes, indicating that these conditions may lead to health emergencies that necessitate immediate care but are less sudden than typical emergencies.

From this data, we can observe that the billing amounts are quite similar across different insurance providers, with only slight variations. Medicare seems to have the highest average billing amount, while UnitedHealthcare has the lowest.

***conclusion***

In conclusion, this analysis provides valuable insights into the relationships between medical conditions, demographics, hospital billing, and admission types. The findings reveal that gender and blood type have little impact on the prevalence of common medical conditions, while hospital billing varies significantly between institutions. Additionally, the distribution of medical conditions across admission types emphasizes the need for both planned and urgent care in managing chronic diseases. These results can help healthcare providers and policymakers better understand patterns in medical care, ultimately improving patient outcomes and resource allocation.



**Tools used:**

1.**Excel**: analyze data and create graphs and reports, used to analyze patient data and create visual graphs and reports that highlight patient admission patterns, treatment types, and distribution across departments.

(Pivot tables, Charts, Filter, Conditional formatting).

2.**Python**: Python was used in this project to analyze healthcare data through the use of Pandas for data manipulation and various functions to perform calculations and extract insights from the datasets.

