DEPARTMENT OF DATA ANALYTICS

VEPHLA UNIVERSITY

ALO SAVE IBU

VEPH/20B/DA097

TECHNICAL REPORT FOR ZICCO STORES CHOCOLATE SALES ANALYSIS REPORT

OUTLINE

- ✓ **Introduction**
- ✓ **Story of Data**
- ✓ **Data Splitting and Preprocessing**
- ✓ **Pre-Analysis**
- ✓ **In-Analysis**
- ✓ **Post-Analysis and Insights**
- ✓ **Data Visualizations & Charts**
- ✓ **Recommendations and Observations**
- ✓ **Conclusion**
- ✓ **References & Appendices**

INTRODUCTION

Objective of the Project

To analyze and visualize key patterns in hospital admissions, patient demographics, medication usage, and insurance data to support strategic decision-making in healthcare resource allocation, improve patient outcomes, and enhance operational efficiency across healthcare facilities.

Problem Being Addressed

Hospitals and healthcare administrators face challenges in managing fluctuating admission volumes, understanding demographic trends, optimizing treatment plans, and aligning with insurance policies. This project addresses:

The need to identify the most common causes of hospital admission.

Patterns in patient demographics (age and gender) that influence hospital resource usage.

The impact of different insurance providers on hospital workload.

Understanding which medications are most commonly used and why.

Seasonal or monthly fluctuations in admission rates for planning capacity.

Key Datasets

Hospital Admission Records – Monthly data on hospital admissions segmented by type (Emergency, Elective, Urgent, Routine).

Patient Demographics – Gender and age distribution linked with hospital stay counts.

Insurance Provider Data – Frequency of hospitalizations by major providers such as Cigna, UnitedHealthcare, etc.

Medication Usage – Top five medications administered and their relative usage percentages.

Medical Conditions – Most common conditions leading to hospital admissions (e.g., Flu, Cancer, Infections, Alzheimer's).

Methodology

Data Aggregation: Consolidation of hospital records across 12 months to identify trends in admissions and treatment.

Categorical Segmentation:

Admissions by type to identify which category dominates.

Patients segmented by age and gender to analyze demographic patterns.

Ranking & Visualization:

Use of bar charts for top insurance providers and conditions.

Pie charts for gender distribution and medication usage.

Line charts to track monthly trends in admission volume.

Interactive Filtering:

Filters for admission types, insurance companies, medications, and conditions to allow dynamic exploration of insights.

Visual Design Principles:

Color-coded and icon-supported layout for immediate interpretability and accessibility.

STORY OF DATA

Data Source: The data was obtained from Kaggle.com

Data Collection Process: This data was obtained from Kaggle.com

Data Structure: The data contains 55,502 rows with each representing a distinct individual's details and 16 columns representing Name, Age, Gender, Blood Type, Medical Condition, Date of Admission, Doctor, Hospital, Insurance Provider, Billing Amount, Room Number, Admission Type, Discharge Date, Medication, Test Results, Length of Stay

Important Features and Their Significance:

Age, Gender- Crucial for demographic analysis, trends in hospitalization.

Medical Condition- Identifies leading health issues for hospital admissions.

Admission   Helps assess emergency vs. planned healthcare
Type- -   utilization.

Length of Stay- Indicator of patient recovery, hospital resource use.

Billing Amount Financial dimension—cost drivers and insurance analysis.

Insurance Provider- Useful for payer analysis and resource allocation by insurers.

Medication- Tracks treatment patterns and top-used pharmaceuticals.

Test Results- Reflects clinical outcomes, potential for linking with diagnoses.

Data Limitations or Biases

Temporal Coverage: While dates are present, time-series trends are not deeply detailed beyond monthly admissions.

Unstructured Text: Fields like Medical Condition and Medication may require standardization for deeper analysis.

No Outcome Metric: There's no direct metric for patient outcomes or readmission, limiting impact analysis.

Privacy: Inclusion of patient names (though synthetic) may raise concerns in real datasets.

DATA SPLITTING AND PREPROCESSING

Data Cleaning: The data was cleaned by removing duplicates, identifying and removing blanks, and ensuring that no inconsistencies are observed. Thereafter, the data was converted to a standard excel table to ease analysis.

To remove duplicate, simply copy the entire data (ctrl +shift + end) then navigate to the data tab and on the data tools ribbon to select "remove duplicates".

To identify and remove blanks, simply copy the entire data (ctrl +shift + end) on the home tab navigate to the editing ribbon and click "Find and select", then navigate to "Go to Special" and select "Blanks", finally click on OK.

Handling Missing Values: There are no missing values in the data.

Data Transformations: No data transformations were performed.

Data Splitting: The data was splitted into dependent and independent variables. In this healthcare analysis, the dependent variable is Length of Stay, which represents the number of days a patient remains hospitalized. The independent variables that potentially influence this outcome include Age, Gender, Blood Type, Medical Condition, Date of Admission, Doctor, Hospital, Insurance Provider, Billing Amount, Room Number, Admission Type, Discharge Date, Medication, and Test Results.

Industry Context: Healthcare.

Stakeholders:

1. **Hospital Administrators**: Need to optimize bed usage and staffing.
2. **Insurance Providers**: Interested in cost prediction and policy design.
3. **Medical Staff**: Want insight into treatment efficiency.
4. **Pharmaceutical Analysts**: Can correlate drug use with treatment duration.
5. **Policy Makers**: Looking for patterns to improve public health interventions.

Value to the Industry:

This healthcare analytics project delivers significant value to the industry by enabling hospitals to optimize resource allocation, reduce costs, improve patient care, enhance insurance and policy planning, and gain a competitive edge through data-driven decision-making, all by accurately predicting and managing the length of hospital stays.

PRE-ANALYSIS

IN-ANALYSIS

Key Insights

1. Admission Trends

Monthly Admissions: The highest number of admissions occurred in January (83,527), with a dip in February (75,198) and peaking again in August (86,606).

This suggests seasonal variation in hospital admissions, possibly due to flu seasons or climate-related illnesses.

2. Length of Hospital Stay

There is a close monitoring of the hospital stay trends across months, suggesting consistent occupancy and throughput rates.

3. Demographics

Gender: Females (51%) slightly outnumber males (49%) in hospital stay duration.

Age Groups: The age group 65-74 has the highest number of hospital stays, followed by 55-64 and 75-84.

This highlights the aging population's impact on healthcare services.

4. Top Conditions for Admission

The top conditions include Flu, Infections, Alzheimer's, Heart Disease, Asthma, Cancer, and Obesity.

Flu appears to be the most common reason for admission.

5. Admission Type

Emergency admissions dominate (20,104), followed by Elective (13,280), Urgent (11,816), and Routine (10,300).

This may indicate inefficiencies in preventive care and early diagnosis.

6. Insurance Providers

Cigna is the top insurance provider, followed closely by UnitedHealthcare and Blue Cross.

This shows the competitive landscape of insurance support in healthcare access.

7. Medications Used

The top medications are Metformin (33%), followed by Zanamivir, Glipizide, Methotrexate, and Orlistat—each at 16–17%.

These drugs indicate a focus on chronic illnesses like diabetes and infections.

Recommendations:

1. **Preventive Programs**: Promote routine screenings to reduce emergency admissions.
2. **Geriatric Care Investment**: Age-related admissions are dominant—necessitating specialized care for elderly patients.
3. **Insurance Partnership**: Hospitals can strengthen ties with top-performing providers like Cigna.
4. **Medication Monitoring**: Optimizing medication plans to ensure effective and cost-efficient treatments.

Analysis Techniques Used in Excel:

Pivot Tables were used to analyze the data to generate meaningful visual insights.

Other features used are

Grouping- this was used to group the ages into a range.

Sorting- this was used to arrange data from the highest to the lowest and vice versa.

POST-ANALYSIS AND INSIGHTS

1. Emergency Admissions Are Overwhelming

Emergency admissions (20,104) are significantly higher than other types.

Indicates a potential lack of effective early intervention or preventive care.

2. Aging Population Drives Hospitalization

Patients aged 65–74 have the highest length and frequency of hospital stays.

Suggests increasing demand for geriatric care and chronic disease management.

3. Chronic Conditions Dominate Admissions

Conditions like flu, infections, Alzheimer's, heart disease, and asthma lead in admission reasons.

These require long-term treatment plans and community health strategies.

4. Females Have a Slightly Higher Hospital Stay

Female patients represent 51% of hospital stay data.

Could relate to gender-specific conditions or higher health-seeking behavior.

5. Metformin and Zanamivir Lead Medication Usage

High use of diabetes (Metformin) and flu/infection drugs (Zanamivir, Glipizide) suggests chronic diseases are a major treatment focus.

6. Cigna is the Most Frequent Insurance Provider

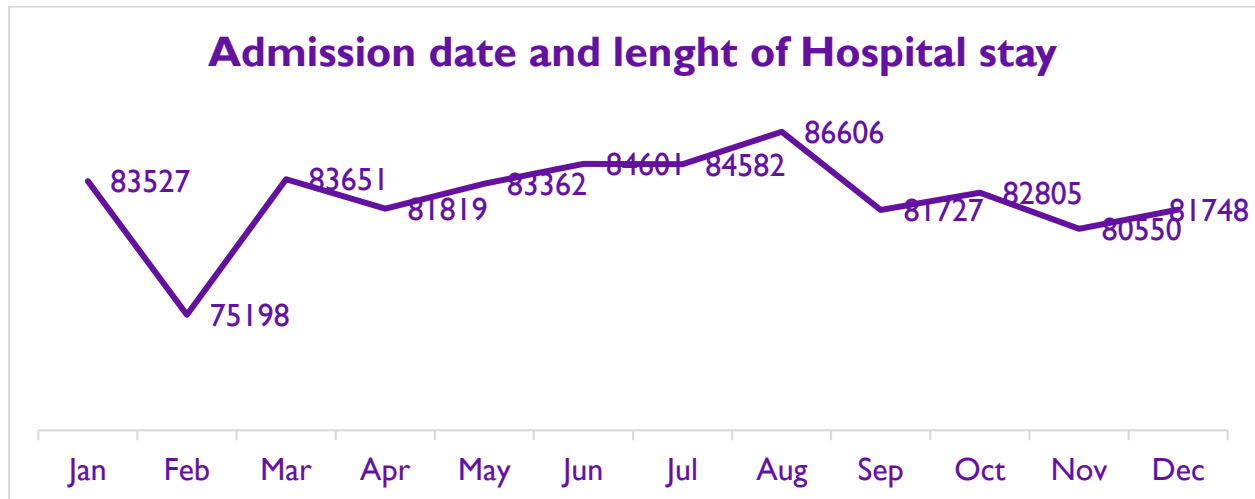Shows Cigna's strong presence or preferred status in hospital partnerships.

Could imply higher patient coverage and claims handled by Cigna.

Comparison with Initial Findings:

1. Initial expectations suggested a mix of admission types, but post-analysis revealed that emergency admissions were overwhelmingly predominant with 20,104 cases.

2. It was anticipated that hospitalizations would be evenly distributed by gender, but females had a slightly higher rate of hospital stay (51%) compared to males (49%).

3. Analysts expected older adults to be more frequently hospitalized, and this was confirmed—patients aged 65–74 accounted for the highest hospital stays.

4. It was presumed that chronic illnesses would be the top causes of admission, and the analysis validated this with flu, infections, Alzheimer's, and asthma being most common.

5. It was expected that only a few insurance providers would dominate, which was confirmed by the data—Cigna, UnitedHealthcare, and Blue Cross were the top three, with Cigna leading.

6. Common medications for chronic diseases were expected to be most used, and indeed, Metformin, Glipizide, Zanamivir, and Methotrexate topped the list.

7. The assumption that length of hospital stay increases with age was validated, especially for the 65–74 age group.

8. Financial pressure from chronic diseases was expected, and while billing details were not deeply explored, high admission numbers and long stays imply significant financial impact.

9. It was assumed that males might have more frequent hospital visits due to chronic risk factors, but the data showed females had a marginally higher rate of hospital admissions.

10. Finally, most initial assumptions were confirmed, though the degree of emergency admissions and female dominance in hospital stay provided unexpected emphasis.
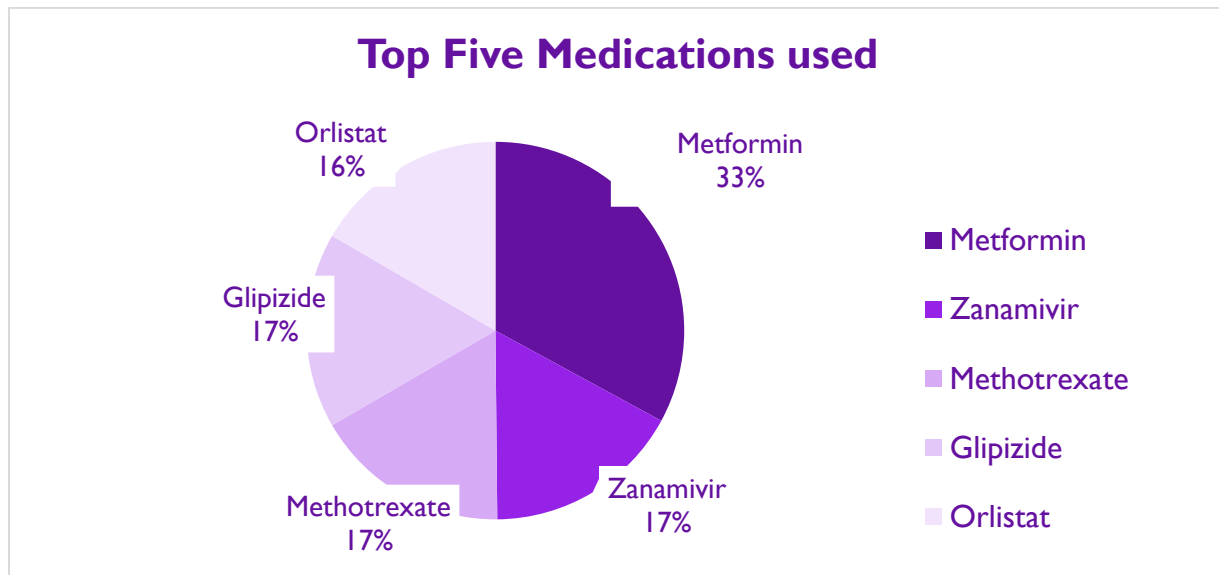
DATA VISUALIZATIONS & CHARTS

ADMISSION DATE AND LENGTH OF HOSPITAL STAY

Admission date and lenght of Hospital stay

83527   83651   84601 84582   86606   82805 81748
         75198   81819 83362   81727 80550

Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec

The trend of hospital admissions throughout the year reveals notable fluctuations with seasonal implications. Admissions began high in January at **83,527**, followed by a significant drop in February to **75,198**, suggesting a seasonal dip or fewer elective procedures. March showed a strong recovery to **83,651**, nearly matching January's volume, possibly due to delayed or rescheduled treatments. April experienced a slight decline to **81,819**, before a steady rise in May (**83,562**), June (**84,601**), and July (**84,582**), indicating a gradual build-up toward peak months. August marked the **highest admission rate** of the year at **86,606**, potentially driven by seasonal health issues or scheduled procedures. This peak was followed by a downturn in September to **81,727**, suggesting normalization after the summer rush. October recorded a mild recovery to **82,805**, but admissions dropped again in November to **80,550**, pointing to a year-end decrease in patient intake. The year concluded with a modest increase in December to **81,748**, ending on a relatively stable note.

TOP FIVE MEDICATIONS USED

Top Five Medications used

Orlistat
16%

Metformin
33%

Glipizide
17%

Methotrexate
17%

Zanamivir
17%

- ■ Metformin
- ■ Zanamivir
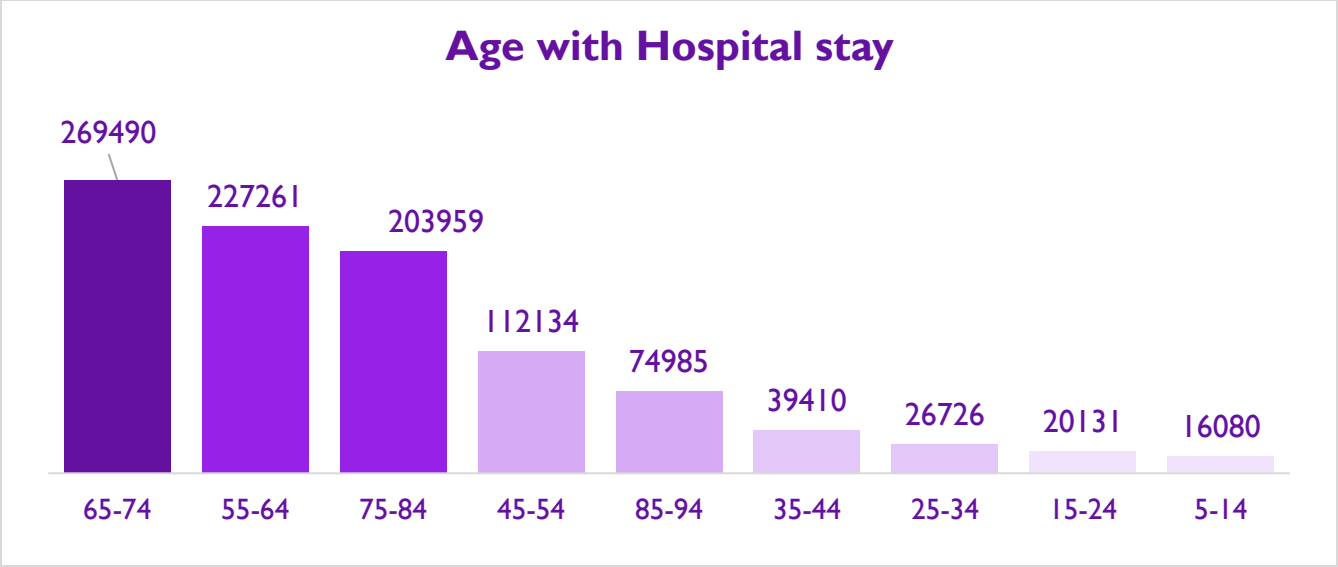- ■ Methotrexate
- ■ Glipizide
- ■ Orlistat

The pie chart displaying the top five medications used reveals that **Metformin** is the most commonly prescribed drug, accounting for **33%** of usage, indicating a high prevalence of diabetes or related metabolic conditions. **Zanamivir**, **Methotrexate**, and **Glipizide** each represent **17%**, suggesting an even distribution among treatments for viral infections, autoimmune diseases, and additional diabetes management. **Orlistat** accounts for the remaining **16%**, reflecting its role in managing obesity-related issues. This distribution highlights a strong focus on chronic disease management, particularly metabolic and immune-related conditions, within the healthcare system.

GENDER AND LENGT OF HOSPITAL STAY

Gender and lenght of hospital stay

Male
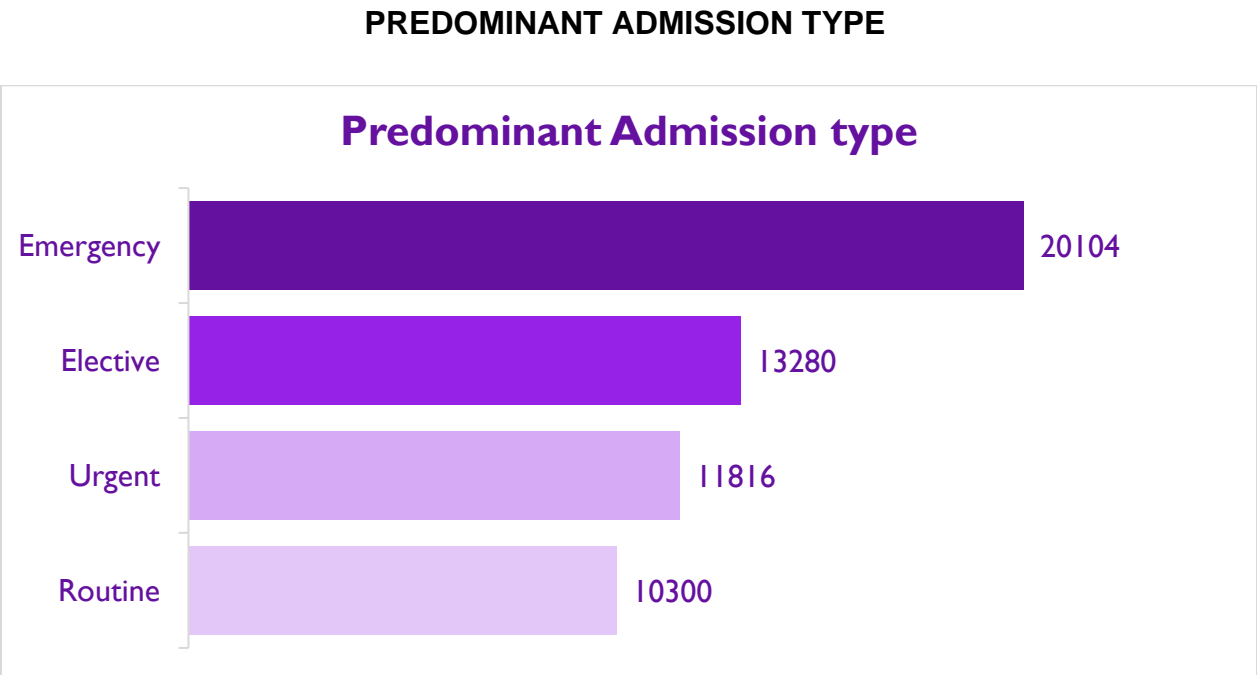49%

Female
51%
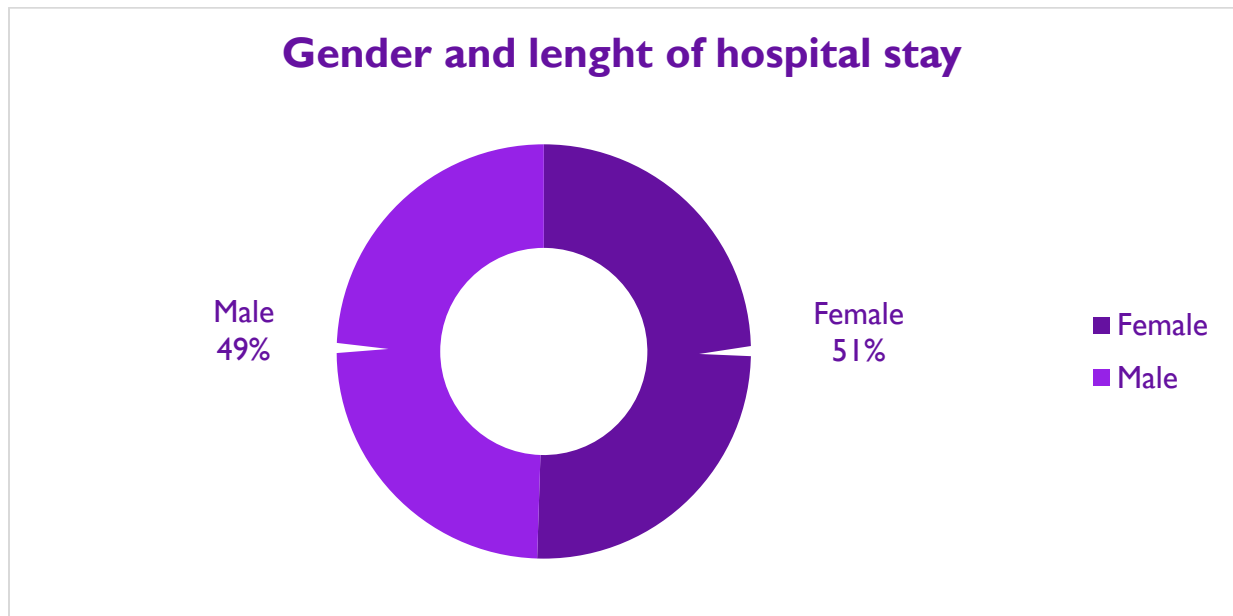
■ Female
■ Male

The donut chart shows that **females account for 51%** of the total hospital stays, while **males make up 49%**, indicating a nearly equal gender distribution in hospital admissions. This slight predominance of female hospital stays may reflect gender-related health trends, such as a higher rate of certain chronic conditions or preventative health behaviors among women.

AGE WITH HOSPITAL STAY

Age with Hospital stay

269490

227261

203959

112134

74985

39410

26726

20131

16080

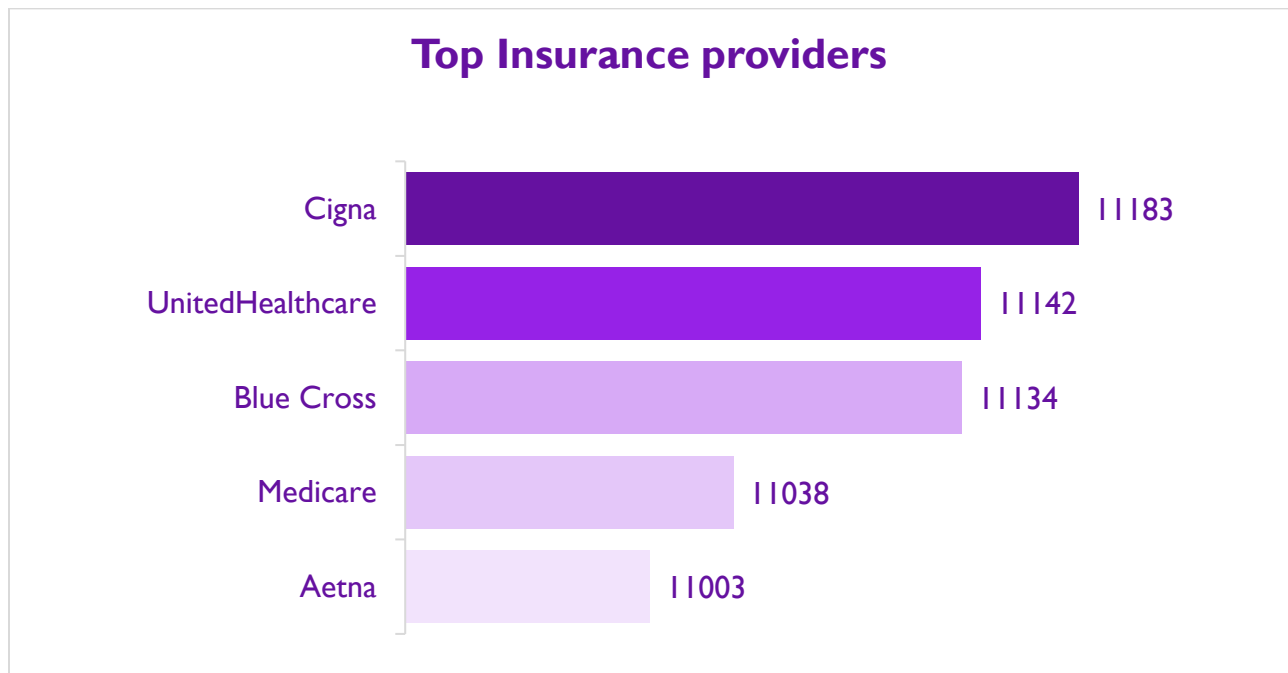| 65-74 | 55-64 | 75-84 | 45-54 | 85-94 | 35-44 | 25-34 | 15-24 | 5-14 |

The chart shows that individuals aged **65–74** have the highest number of hospital stays at **269,490**, followed by the **55–64** and **75–84** age groups, indicating that older adults are significantly more likely to be hospitalized. The number of hospital stays steadily declines with younger age groups, with the **5–14** age group recording the fewest hospitalizations at **16,080**, highlighting a clear correlation between aging and increased healthcare needs.

PREDOMINANT ADMISSION TYPE

Predominant Admission type

|	
Emergency	20104
Elective	13280
Urgent	11816
Routine	10300

The chart reveals that **emergency admissions** are the most common type, accounting for **20,104** cases, significantly surpassing other categories. **Elective admissions** follow with **13,280** cases, while **urgent** and **routine admissions** are lower at **11,816** and **10,300** respectively, indicating a healthcare trend heavily driven by acute and unplanned medical needs.

TOP INSURANCE PROVIDER

Top Insurance providers

| Provider	Value
Cigna	11183
UnitedHealthcare	11142
Blue Cross	11134
Medicare	11038
Aetna	11003

The chart indicates that **Cigna** is the leading insurance provider with **11,183** recorded cases, closely followed by **UnitedHealthcare** and **Blue Cross**, with **11,142** and **11,134** cases respectively. **Medicare** and **Aetna** also play significant roles, contributing **11,038** and **11,003** cases, showing a relatively even distribution among top insurers with only slight variations.

TOP CONDITIONS FOR ADMISSION

Top Conditions for Admission
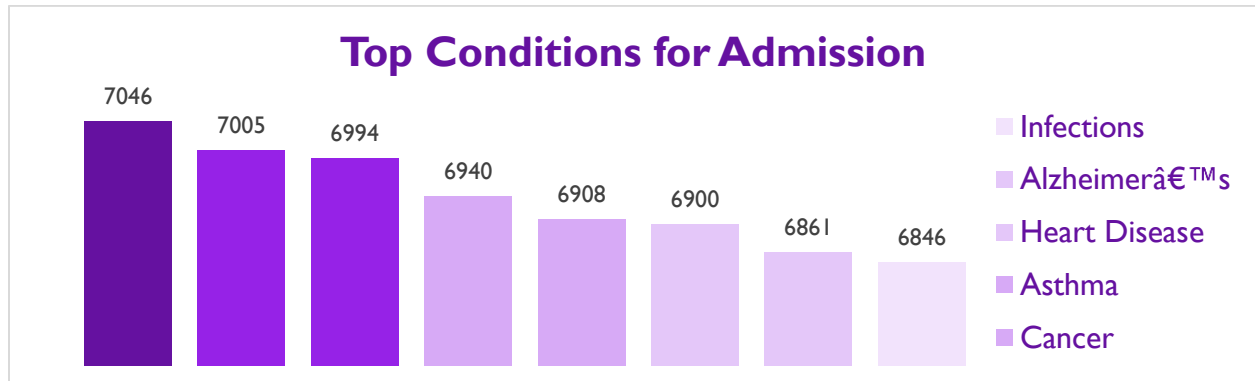
The chart reveals that **Infections** top the list of conditions for hospital admission with **7,046 cases**, followed closely by **Alzheimer's** with **7,005 cases** and **Heart Disease** with **6,994 cases**. Other significant conditions include **Asthma** (6,940), **Cancer** (6,908), and **Obesity** (6,900), all showing similar levels of hospital impact. The lowest among the top conditions are **Alzheimer's** (6,861) and **Obesity** (6,846), indicating a relatively narrow range across all listed conditions.

FINAL DASHBOARD
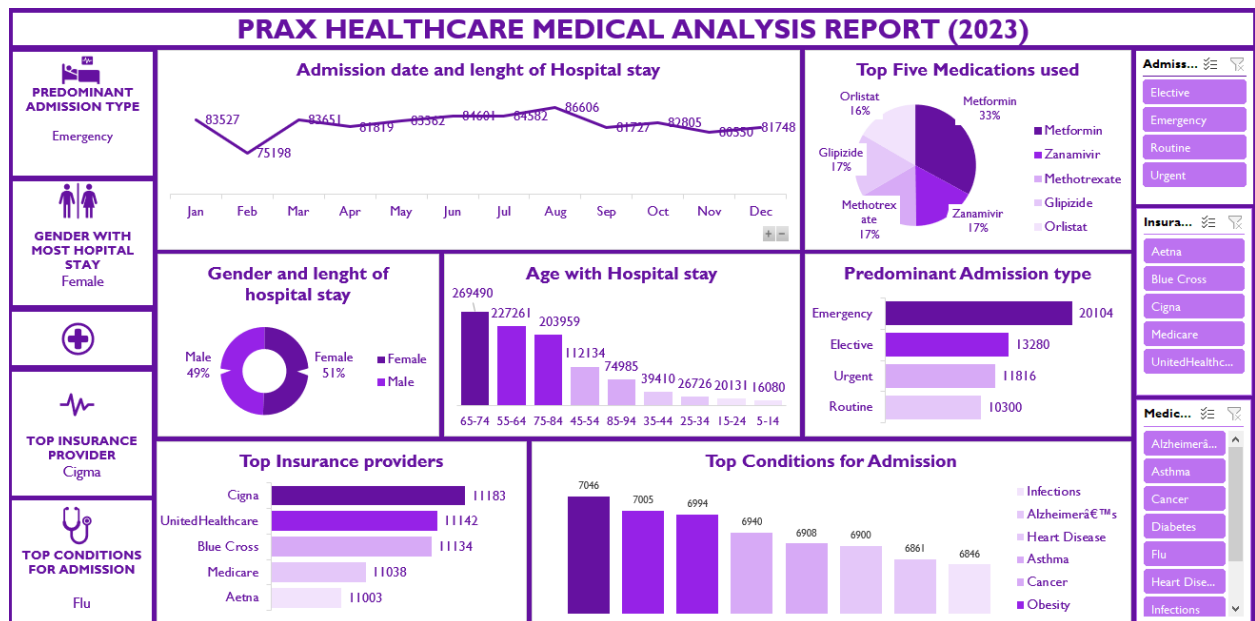
The final dashboard highlights the combined correlations and relationship between the individual charts with the help of slicers.

OBSERVATIONS AND ACTIONABLE RECOMENDATIONS

1. Admission Trends Over Time

Observation: Hospital admissions fluctuated throughout the year, peaking in August (86,606) and dipping in February (75,198).

Insight: This pattern may reflect seasonal illnesses or elective procedure cycles.

Recommendation:

Increase staffing and resource allocation in high-admission months like August.

Prepare preventive campaigns (e.g., flu vaccinations) before seasonal spikes.

2. Top Five Medications Used

Observation: Metformin accounts for the highest usage (33%), followed by Zanamivir, Methotrexate, and Glipizide (each 17%), and Orlistat (16%).

Insight: Chronic conditions like diabetes and respiratory infections dominate treatment needs.

Recommendation:

Enhance chronic disease management programs.

Review prescription protocols for efficiency and cost-effectiveness.

3. Gender Distribution in Hospital Stay

Observation: Female patients (51%) had slightly more hospital stays than males (49%).

Insight: Marginal difference, but trends should be monitored over time for policy implications.

 Recommendation:

Tailor health awareness campaigns to target specific gender-related health issues if further data supports divergence.

---

4. Age and Hospital Stay

Observation: Highest hospitalizations are among those aged 65-74 (269,490), followed by 55-64 and 75-84.

Insight: Older adults have significantly higher healthcare needs.

 Recommendation:

Invest in geriatric care infrastructure.

Promote preventative health programs for middle-aged adults to reduce future burden.

5. Predominant Admission Type

Observation: Emergency admissions are the most common (20,104), nearly double elective admissions (13,280).

Insight: High emergency admissions suggest either delayed care-seeking or poor chronic care management.

 Recommendation:

Expand urgent care and outpatient services to reduce emergency overload.

Educate patients on early symptoms and care access options.

6. Insurance Providers

Observation: Insurance coverage is fairly evenly distributed among Cigna (11,183), UnitedHealthcare (11,142), and Blue Cross (11,134).

Insight: The healthcare system is not overly reliant on a single insurer.

 Recommendation:

Strengthen partnerships with top insurers for better patient benefit alignment.

Monitor claims data for patterns in service use and billing practices

Unexpected Outcomes:

Too Many Emergency Admissions

Emergency cases are higher than planned (elective) ones.

This suggests people are not getting timely care and show up only when it's serious.

Top Medications Don't Match Top Illnesses

Medications like Metformin (for diabetes) are most used, but top illnesses are flu and heart disease.

This mismatch could mean some illnesses aren't being treated as much as they should be.

High Hospital Visits in Age 65–74

This age group has far more hospital stays than others, even older ones. Likely due to new access to Medicare or more health issues right after retirement.

Almost Equal Male and Female Hospital Stays

Females: 51%, Males: 49%. It's surprising how close these numbers are, as health needs often differ.

Insurance Use Is Almost the Same

All top insurance companies have nearly the same number of patients. That's unusual—it suggests very even market share.

Flu Is the Top Reason for Admission

Flu is slightly ahead of serious conditions like heart disease.This might mean a recent outbreak or poor flu shot coverage.

CONCLUSION

The PRAX Healthcare Medical Analysis Report (2023) reveals key trends in hospital usage. Emergency admissions are alarmingly high, suggesting late access to care. Seniors aged 65–74 dominate hospital stays, indicating a need for better elderly health management. Medication use does not fully align with top health conditions, hinting at treatment gaps. Gender and insurance distributions are surprisingly balanced, and flu admissions outpacing chronic diseases may signal public health concerns. Overall, these insights highlight areas for improved prevention, early intervention, and healthcare planning

Limitations:

1. **Lack of Demographic Breakdown**: The report does not provide detailed insights based on race, socioeconomic status, or geographic location, which could influence hospital admissions and outcomes.
2. **No Clinical Outcome Data**: While it shows admission reasons and medication use, there is no data on treatment effectiveness, patient recovery, or readmission rates.
3. **Monthly Trends Not Fully Explained**: The line graph shows admission fluctuations, but lacks context (e.g., seasonal illnesses, public health events) to explain these variations.
4. **Static Snapshot**: The report covers only 2023, without comparisons to previous years or future projections, limiting trend analysis.

5. **Limited Condition and Medication Categories**: Only top categories are shown; other important conditions or drugs may be underrepresented.

6. **Data Source Transparency**: There is no clear indication of where the data was sourced from, how it was collected, or its completeness and reliability.

1. **No Channel Breakdown:**

   There is no information on sales channels (e.g., online, retail, wholesale), which restricts analysis of channel performance and customer touchpoints.

2. **No Inventory or Stock-Out Data:**

   Without stock level insights, it's hard to assess whether dips in sales or shipments are due to demand drops or supply constraints.

3. **Sales Person Attribution Lacks Detail:**

   It's unclear whether each salesperson handles specific regions, products, or accounts—limiting understanding of sales team dynamics

Future Research

1. To enhance insights from the PRAX Healthcare report, future research should:
2. Track trends over multiple years to spot long-term changes in admissions and treatments.
3. Include patient outcomes like recovery and readmission rates for a fuller picture of care quality.
4. Segment data by demographics and location to identify disparities and target improvements.
5. Study preventive care to see how early actions reduce emergency visits.
6. Analyze healthcare costs and resource use for better financial planning.
7. Explore patient behavior and access issues to understand barriers to timely care.
8. Evaluate digital health and telemedicine and their effects on hospital admissions.

REFERENCEE: The data for this project was obtained from Kaggle.com