Exploratory Data Analysis (EDA) for Hospital Emergency Room Analytics

1. Data Overview and Structure

- Data Sources: The dataset includes patient demographics, ER visit details, wait times, satisfaction scores, department referrals, and admission statuses.
- Initial Checks: Data was checked for completeness, consistency, missing values, and duplicate records. Data types were validated (e.g., dates, numeric fields, categorical variables).

2. Key Metrics and Trends

a. Patient Volume

- No. of Patients: The total number of ER visits (9,216) was calculated to understand the ER's workload.
- Trend Analysis: Line charts show how patient volume fluctuates over time (daily/monthly), revealing peak and off-peak periods.

b. Wait Time Analysis

- Average Wait Time: The mean wait time (35.3 minutes) was calculated to assess ER efficiency.
- Distribution: Trends in wait times were plotted to identify patterns, spikes, or improvements over time.

c. Patient Satisfaction

- Satisfaction Score: The average satisfaction score (4.99) was analyzed to gauge patient experience.
- Trend: Tracked over time to see if operational changes impacted satisfaction.

d. Referrals

- Patients Referred: The number of patients referred to other departments (3,816) was measured to understand the ER's role in the broader hospital workflow.

3. Categorical Analysis

a. Admission Status

- Admitted vs. Not Admitted: Patients were grouped by admission status, showing a near-even split (50.04% admitted, 49.96% not admitted), which helps in understanding ER outcomes.

b. Age Group Distribution

- Age Segmentation: Patients were grouped into age bands (0-9, 10-19, ..., 70-79) to identify which age groups most frequently visit the ER. The 30-39 age group had the highest count.

c. Gender Distribution

- Gender Breakdown: Analysis showed a near-even split between male and female patients, with a small portion unconfirmed.

d. Race Distribution

- Race/Ethnicity: Patients were categorized by race (White, African American, Asian, etc.), highlighting diversity and any disparities in ER usage.

4. Operational Performance

a. Timeliness of Care

- % Seen Within 30 Minutes: 59.3% of patients were seen within the target time, while 40.68% missed the target. This metric is crucial for service quality monitoring.

b. Peak Periods

- Patients by Day and Hour: Bar charts and heatmaps were used to identify busy days (Saturday and Sunday) and peak hours, aiding in resource planning and shift management.

c. Department Referrals

- Referral Patterns: Most patients were not referred to other departments, but among those referred, General Practice and Orthopedics were the most common.

5. Detailed Patient Records

- Patient Details Table: A tabular view was created for granular inspection, including patient ID, name, gender, age, admission date, race, wait time, department referral, and admission status. This supports deeper investigation and case review.

6. Data Quality and Insights

- Missing Values: Checked and handled missing or unconfirmed data (e.g., gender not confirmed).
- Outlier Detection: Looked for unusually high or low values in wait times and satisfaction scores.
- Distribution Checks: Used charts to ensure data was distributed as expected (e.g., no unexpected spikes in certain age groups or races).

Summary of EDA Outcomes

- Workload: The ER handles a high and steady volume of patients, with identifiable peak times.
- Efficiency: Most patients are seen within 30 minutes, but there is room for improvement.
- Demographics: The patient population is diverse, with balanced gender and age group representation.
- Operational Insights: Referral and admission patterns provide actionable information for hospital management.
- Data Quality: The data is generally clean, with minor gaps that are transparently handled in the dashboard.

Conclusion:
The EDA process provided a comprehensive understanding of the ER's operations, patient demographics, and performance. These insights directly informed the design and focus of the Power BI dashboard, ensuring it meets the needs of hospital administrators and supports data-driven decision-making.

Process:

Step 1: Data Collection and Preparation

Data Sources:

- Patient demographics (age, gender, race)
- ER visit details (check-in, wait time, satisfaction scores, admission status)
- Department referrals

Process:

- Imported raw data CSV files into Power BI.
- Used Power Query Editor to clean data: removed duplicates, handled missing values, and standardized column names.
- Transformed data types (e.g., date fields, numeric fields) and created calculated columns such as Age Group and Wait Time.

Step 2: Data Modeling

Actions Taken:

- Established relationships between tables (e.g., linking patient info to visit records using Patient ID).
- Created calculated columns for Age Group (e.g., 0-9, 10-19, etc.) and Wait Time (difference between check-in and seen time).
- Defined DAX measures for key metrics such as:
    - Total Patients
    - Average Wait Time
    - Patient Satisfaction Score
    - % of Patients Seen Within 30 Minutes
    - Admission Status breakdown
    - Patients by Gender, Age Group, Race, and Department Referral

Step 3: KPI and Metric Selection

KPIs Identified:

- No. of Patients
- Average Wait Time
- Patient Satisfaction Score
- No. of Patients Referred

Other Metrics:

- Admission Status (Admitted/Not Admitted)
- Patients by Age Group, Gender, Race
- % of Patients Seen Within 30 Minutes
- Patients by Department Referral
- Patient volume by Day and Hour

Step 4: Dashboard Layout and Design

Navigation:

- Designed a left-side navigation pane with buttons for Monthly View, Consolidated View, and Patient Details.

Filter:

- Added a Date slicer at the top right for time-based filtering.

Visual Hierarchy:

- Top row: KPI cards with trend lines for headline metrics.
- Middle: Bar and donut charts for admission status, patients seen within 30 min, age group, and gender.
- Right: Bar chart and heatmap/matrix for patients by day and hour.
- Bottom: Department referral and patient race breakdowns.
- Separate page: Patient Details table for granular analysis.

Step 5: Building Visualizations

Visuals Used:

- KPI Cards: For No. of Patients, Avg Wait Time, Satisfaction Score, and Patients Referred, each with a trend chart.
- Bar/Column Charts: For Age Group, Department Referral, Patient Race, and Admission Status.
- Donut Charts: For Gender and % Seen Within 30 Minutes.
- Matrix/Heatmap: For Patients by Day and Hour to highlight peak periods.

- Table Visual: For Patient Details (ID, Name, Gender, Age, Admin Date, Race, Wait Time, Department Referral, Admission Status).

Interactivity:

- Enabled cross-filtering so users can interact with the data.
- Added tooltips for additional context.

Step 6: Formatting and Branding

- Applied a consistent color scheme (blues, grays, and highlights).
- Used clear labels and descriptive titles for all visuals.
- Incorporated hospital/medical icons and the hospital logo for branding.
- Ensured layout consistency and alignment across all pages.

Step 7: Testing and Validation

- Validated all calculations and visuals for accuracy.