Executive Summary: Hospital Management Data Analysis

Astrid Villalobos

July, 2025

Introduction

This report summarizes the results of a data analysis project conducted using a fictional hospital dataset. The objective was to uncover actionable insights to improve healthcare delivery, resource planning, and operational efficiency.

Objectives

- Identify trends in patient demographics, appointment scheduling, and billing.
- Support healthcare decision-making through data-driven insights.
- Visualize patterns that can inform staffing and cost management strategies.

Methodology

- Tools Used: Python (Pandas, Seaborn) for data cleaning and exploratory data analysis (EDA), and Power BI for interactive dashboard development.
- Dataset Overview: Analysis based on 200 patient appointments and 50 patient profiles across the year 2023.

Key Findings

Demographics

- The average patient age is 44.6 years.
- The gender distribution is 58% male and 42% female.
- Female patients tend to be older than male patients on average.

Operational Insights

- Tuesdays and Wednesdays are the busiest days for appointments.
- No-shows and cancellations each represent 26% of scheduled appointments.

Financial Overview

- The month with the highest billing was April 2023, totaling $64,271.54.
- The most expensive treatment types were Chemotherapy and MRI.

Recommendations

- Increase staffing levels on peak days (Tuesdays and Wednesdays).
- Implement strategies to reduce no-show and cancellation rates (e.g., reminders, overbooking models).
- Monitor and evaluate high-cost treatments for cost-effectiveness and prioritization.

Limitations

- Limited Sample Size: The dataset includes only 50 patients and 200 appointments, which may not fully represent complex hospital dynamics.
- Synthetic Data: Being fictional, the data lacks real-world variability, such as seasonal trends, insurance coverage, or comorbidities.
- Inconsistencies in Gender Labels: Some mismatches between names and gender categories were detected and only partially resolved.
- Temporal Constraints: The dataset covers only the year 2023, limiting long-term trend analysis.
- Simplified Billing: Billing values are aggregated and lack detailed cost breakdowns (e.g., procedure-specific charges, insurance adjustments).

Conclusion

While based on fictional data, this analysis provides a structured framework for using data-driven methods to improve hospital operations. The combined use of Python and Power BI highlights the value of integrating analytical tools with interactive reporting to support decision-making in healthcare management.