STATISTICAL ANALYSIS OF PATIENT RECOVERY PATTERNS IN A MULTI-STATE HOSPITAL CHAIN

1. Client Context

A private healthcare group operating 15 hospitals across 6 U.S. states approached us with a need to analyze inpatient recovery outcomes. Their objective was to understand if recovery time varied significantly by hospital, patient age group, and type of treatment administered. They also wanted the results presented in a format suitable for internal benchmarking and external regulatory reporting.

2. Problem Statement

The management sought statistical evidence to:

- Identify whether specific hospitals had significantly longer or shorter recovery durations.
- Determine if patient age and treatment type influenced average recovery time.
- Present the results in a clear, quantitative format to guide operational adjustments.

3. Tools and Techniques Used

- **Software:** SPSS and Excel
- **Methods:**
    - Descriptive statistics (mean, median, standard deviation)
    - One-way ANOVA and post hoc Tukey tests for comparing recovery times across hospitals
    - Multiple linear regression to identify the effect of age and treatment type
    - Box plots and histogram visualizations to show data distribution and outliers
    - Confidence interval estimation for benchmarking
- **Reporting:** APA-style written report with embedded tables, visuals, and interpretive commentary

4. Data Overview

- **Sample Size:** 8,450 inpatient records from January to December 2023

- **Key Variables:**
    - Recovery_Time_Days (continuous)
    - Hospital_ID (categorical, 15 levels)
    - Treatment_Type (categorical, 6 types)
    - Patient_Age (continuous)
    - State (categorical, 6 levels)

The data was anonymized and cleaned before analysis. Outliers were flagged using interquartile range (IQR) thresholds, and missing values (<2%) were handled using listwise deletion.

5. Key Findings

- **Hospital Comparison:** ANOVA revealed statistically significant differences in average recovery time across hospitals ($p < 0.01$). Post hoc analysis highlighted 3 hospitals with consistently lower recovery durations.
- **Regression Analysis:** A multiple linear regression model showed that both age and treatment type were significant predictors of recovery time ($R^2 = 0.42$). Recovery duration increased by approximately 0.6 days for every 10 years of patient age.
- **Visualization:** A multi-panel boxplot showed treatment-wise variations in recovery time. Heatmaps and summary tables provided clarity on hospital-wise performance.

6. Deliverables

- A 22-page statistical report including:
    - Cleaned dataset summary
    - Step-by-step methodology
    - Hypothesis testing interpretation
    - Graphs and visuals created in Excel and SPSS
    - Executive summary for hospital directors
- Recommendations for:
    - Reducing variability in treatment outcomes
    - Implementing targeted process improvement in underperforming facilities

7. Client Impact

The report was presented to hospital leadership during their quarterly performance review. Based on the insights, the client initiated a clinical audit for the three lowest-performing hospitals and rolled out a standardized protocol for the top-performing treatment regimen. The client also cited the report during a regulatory inspection as evidence of continuous performance monitoring.