**Project Report: SAS-based Analysis of Hospital Length of Stay Across Insurance Groups**

**Author**: Bryan Jacobs
**Date**: 11-03-2025
**Tools**: SAS (Base SAS, PROC GLM, PROC SGPLOT)
**Dataset**: MIMIC-IV Admissions (Electronic Health Records)

**Objective**
This analysis examines whether hospital length of stay varies by insurance type in the MIMIC-IV admissions dataset. The goal is to understand if certain insurance categories are associated with longer or shorter hospital stays, while controlling for patient and admission characteristics.

**Data Preparation**
Access to MIMIC-IV datasets was obtained via a referral from the University of Arizona's Assistant Professor of Biostatistics, Dr. Yiwen Liu. The dataset was imported from a `.csv` file and cleaned in SAS. Key preprocessing steps included:

- **Length of Stay Calculation**: LOS was computed as `(dischtime - admittime) / 86400` to convert seconds into days.
- **Filtering**: Records with non-positive LOS values were removed.
- **Feature Selection**: Retained key demographic and administrative fields such as insurance, admission type, marital status, race, and whether the patient expired during hospitalization.

The cleaned dataset, `admissions_clean`, was saved for analysis.

**Exploratory Analysis**
Descriptive statistics and visualizations were created to summarize the data:

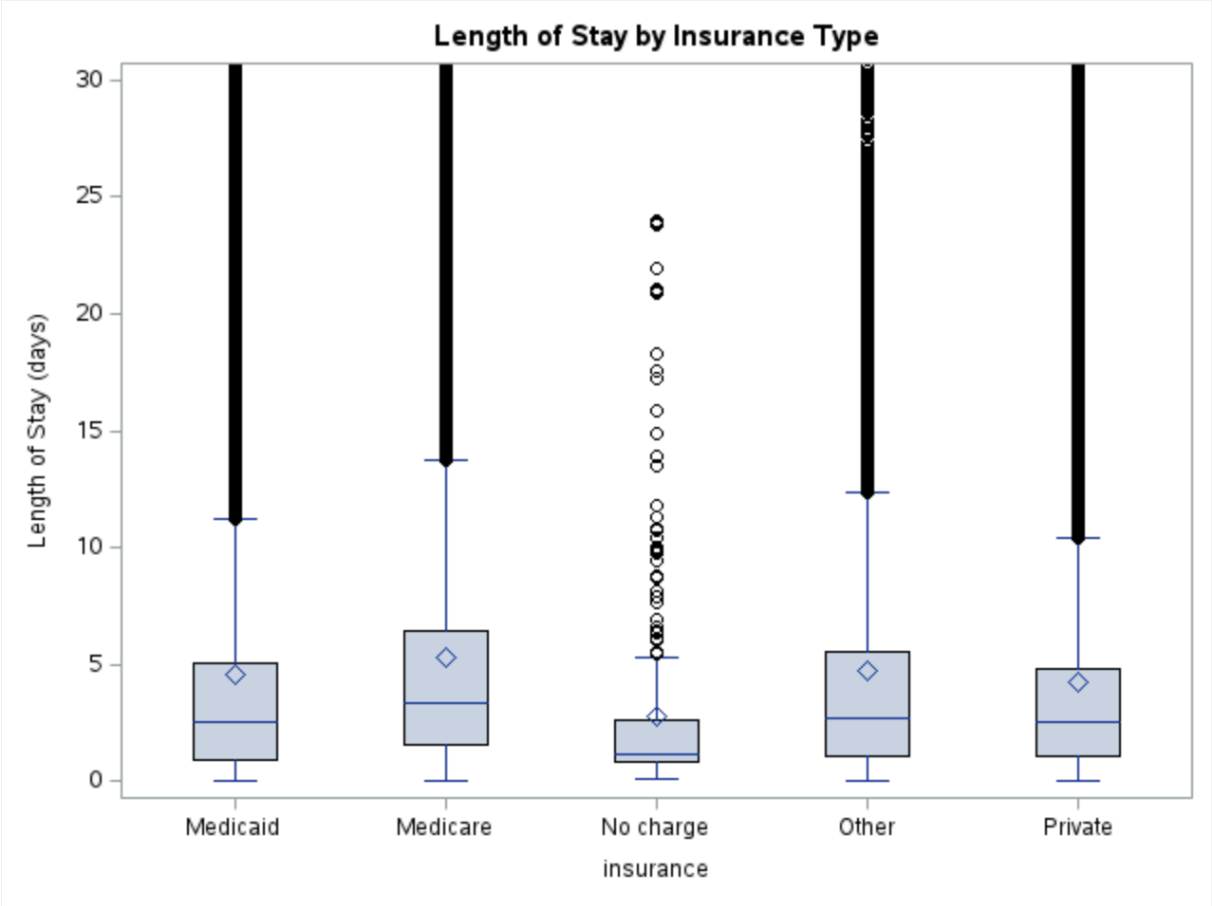
- **Frequency Tables**: Displayed counts for `insurance`, `admission_type`, `marital_status`, and `race`.
- **Descriptive Means**: Mean LOS was calculated by insurance type to provide initial comparisons.

## Mean Length of Stay by Insurance Type

### The MEANS Procedure

| Analysis Variable : length_of_stay | | | | | |
|---|---|---|---|---|---|
| insurance | N Obs | Mean | Std Dev | Minimum | Maximum |
| Medicaid | 104196 | 4.5612330 | 8.0291274 | 0.000694444 | 515.5625000 |
| Medicare | 244500 | 5.3096506 | 7.2399702 | 0.000694444 | 308.9638889 |
| No charge | 463 | 2.7974532 | 4.9151920 | 0.0847222 | 58.0236111 |
| Other | 13995 | 4.7134831 | 7.0124930 | 0.000694444 | 127.8548611 |
| Private | 173360 | 4.1967631 | 6.5786764 | 0.000694444 | 295.9881944 |

- **Boxplot**: A visual comparison (via `PROC SGPLOT`) showed LOS differences among insurance groups.



Preliminary results suggested variation in LOS across insurance types, with Medicaid and Medicare patients tending to stay slightly longer than those with private coverage.

**Modeling Approach**

An analysis of covariance was conducted using linear regression via `PROC GLM`:

```
Length of Stay = Insurance + Admission Type + Marital Status + Race
```

- The reference group for insurance was private insurance
- Admission type, marital status, and race were included as covariates to adjust for potential confounding effects.

**Key Findings**

- The overall model was statistically significant ($p < 0.0001$), explaining about 11% of the variance in length of stay (R-squared = 0.113533).
- Insurance type remained a significant predictor of LOS even after adjustment ($p < 0.0001$).
    - Medicare (+0.78 days) and Medicaid (+0.66 days) were associated with longer stays compared to private insurance.
    - Patients with "No charge" insurance had significantly shorter stays (-2.97 days).
- Admission type had the largest impact, as expected.
- Race and marital status were also statistically significant, though their effect sizes were for the most part negligible.

**Interpretation**

After controlling for admission type, race, and marital status, insurance type still shows a meaningful relationship with hospital length of stay. Patients covered by public insurance programs (Medicare and Medicaid) tend to experience longer hospitalizations, while those with no insurance saw significantly shorter stays. While further research is necessary to draw conclusions, these results could be suggestive of the benefit of government-sponsored insurance programs. It is possible that uninsured patients saw shorter stays due to an inability to pay for necessary treatment, while those covered by Medicare and Medicaid were financially able to stay long enough to receive necessary treatment.

These findings align with existing literature suggesting insurance type influences hospital utilization and discharge timing. https://pmc.ncbi.nlm.nih.gov/articles/PMC5142530/