Hospital Patient Analysis

Understanding Charges and Patient Characteristics

Student Name: Vo, Minh Thien An

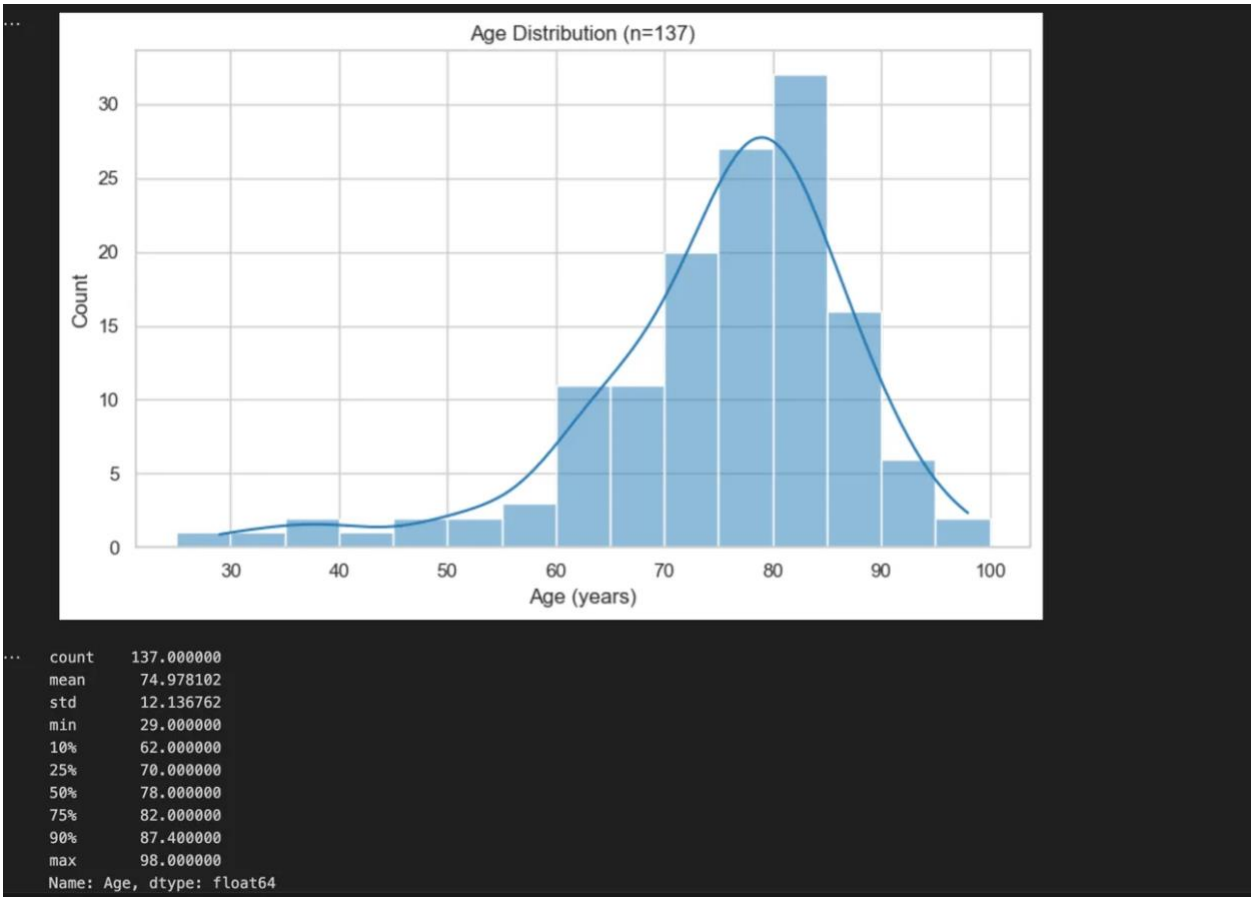
Student ID: m0543138

Introduction

In order to determine which characteristics and circumstances are most helpful in forecasting the total expenses at the conclusion of a hospital stay, this report analyzes 137 patients who were admitted to our facility. To create a trustworthy framework for forecasting patient expenses, we look at charge trends, patient profiles, and the connections between different clinical and demographic factors using descriptive statistics and graphical techniques.

Understanding Our Patient Population

With ages ranging from 29 to 98, the average patient admitted to our hospital is roughly 75 years old (mean = 74.98 years). With half of all patients being 78 years of age or older, the age distribution shows a concentration in the senior population. Our institution primarily serves a senior population, as evidenced by the distribution, which has a distinct bell-shaped pattern focused around the 70–85 age range.




```
count    137.000000
mean      74.978102
std       12.136762
min       29.000000
10%       62.000000
25%       70.000000
50%       78.000000
75%       82.000000
90%       87.400000
max       98.000000
Name: Age, dtype: float64
```

Figure 1: Age distribution

With female patients accounting for 51.1% of admissions (70 patients) and male patients for 48.9% (67 patients), the patient sample exhibits nearly ideal gender balance. This equitable distribution offers a strong basis for comparing results between the sexes.

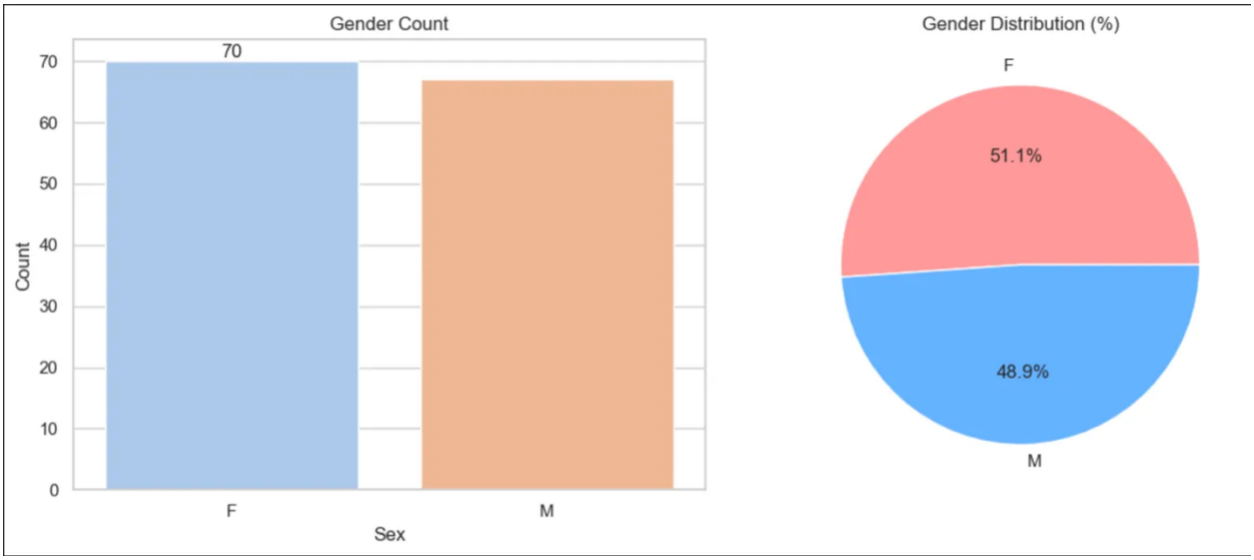


Figure 2: Gender distribution

The average length of stay in the hospital is 4.4 days, although it can vary significantly from 1 to 16 days. 75% of patients are discharged after 5 days, meaning that most spend only a short time. Nonetheless, the existence of longer stays suggests that the hospital treats both common and complicated situations that call for lengthy, extended care.

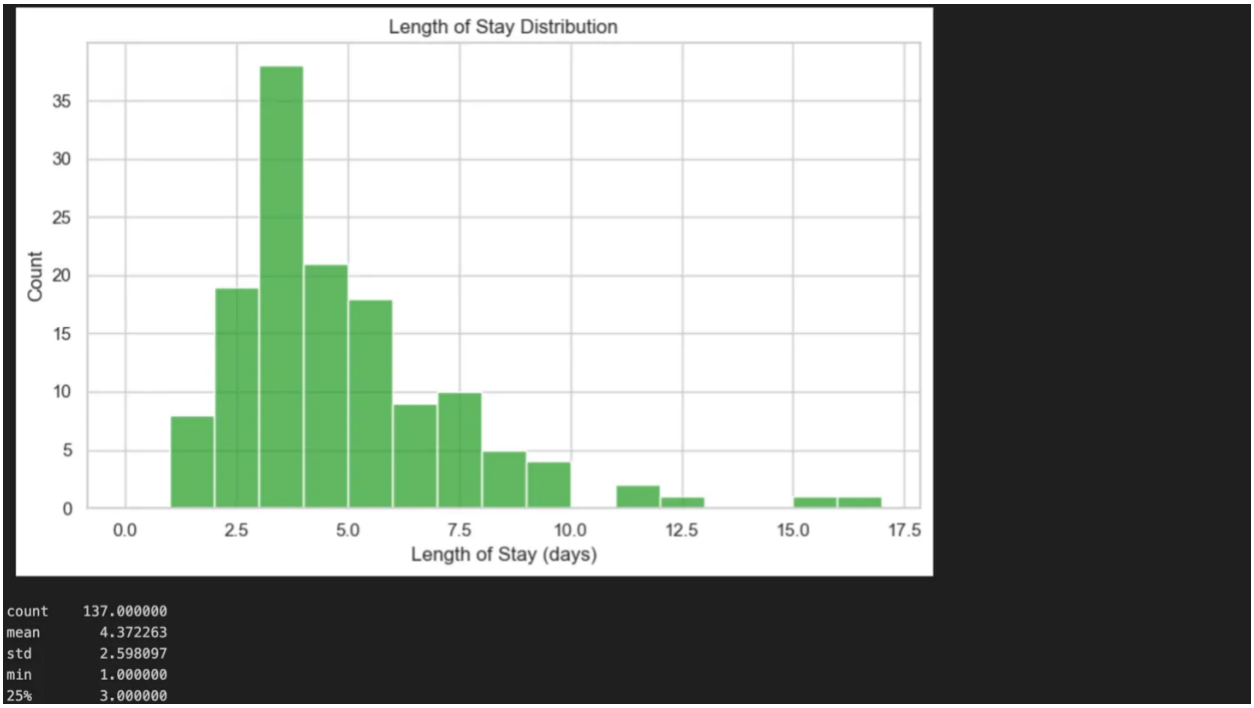


Figure 3: Length of stay distribution

Overall Charge Patterns

The average total charges for all 137 patients come to $6,446. This number, however, covers a wide range of expenses, which range from about $1,500 to more than $40,000. From simple procedures to expensive interventions involving prolonged hospitalization and several treatment modalities, this broad range reflects the varied complexity of medical conditions.

Gender Differences in Charges

When prices are broken down by gender, the average cost for male patients is $6,859, which is $826 more than the average cost for female patients, which is $6,033. Nevertheless, statistical analysis shows that this difference is not significant (t=1.05, p=0.2946), suggesting that gender is not a significant predictor of overall hospital costs. The median values and ranges of the charge distributions for male and female patients are similar. It implies that cost difference is driven by clinical characteristics and treatment intensity rather than patient gender.

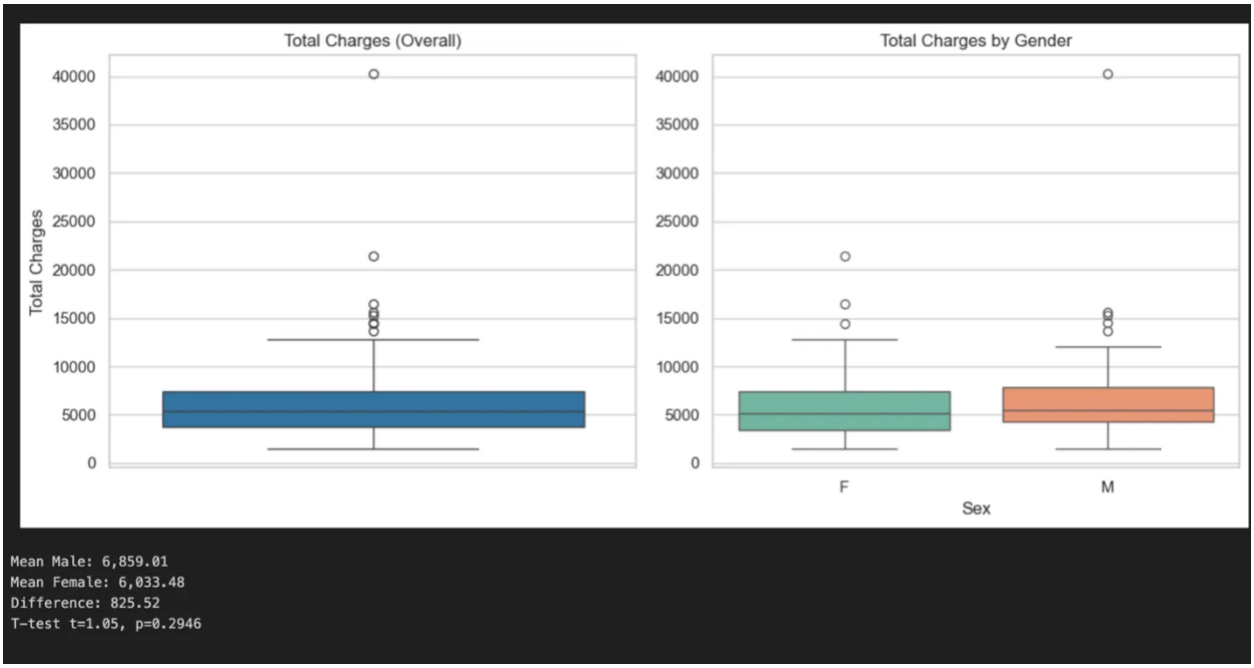


Figure 4: Overall total charge vs. gender-based total charge

The Relationship Between Length of Stay and Total Charges

Whether longer hospital stays translate into higher costs is one of the most obvious problems in healthcare economics. According to our analysis, the length of stay and total charges have a high positive link ($r = 0.77$). This correlation is very strong ($p < 0.001$), suggesting that patients who stay in the hospital longer do, in fact, have much greater expenses.

The scatterplot shows a definite upward trend: patients who remain for three days or less usually pay less than $5,000, while those who stay for ten or more days frequently have total expenditures of more than $15,000. Given that every extra day involves staying fees, nursing care, prescription drugs, meals, and continuing medical supervision, this link makes clinical sense. Even though each instance is unique, the general trend shows that the length of stay is a reliable indicator of the total expenditures.

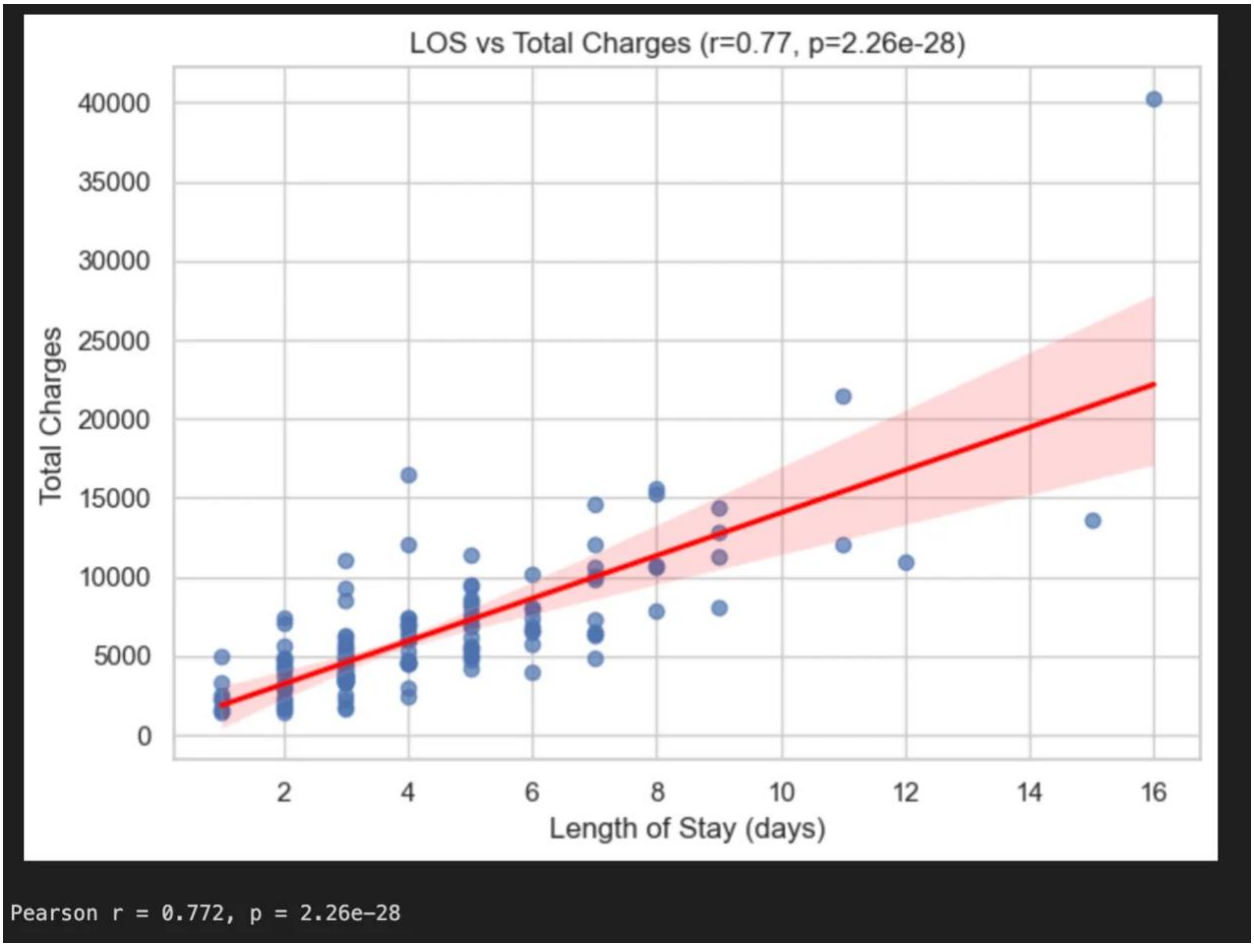


Figure 5: Scatter plot length of stay vs total charge

Identifying the Strongest Predictors of Total Charges

We looked at correlations between all quantitative variables and total charges in order to identify which variables most accurately predict total charges. Several crucial insights for prediction are shown by this analysis:

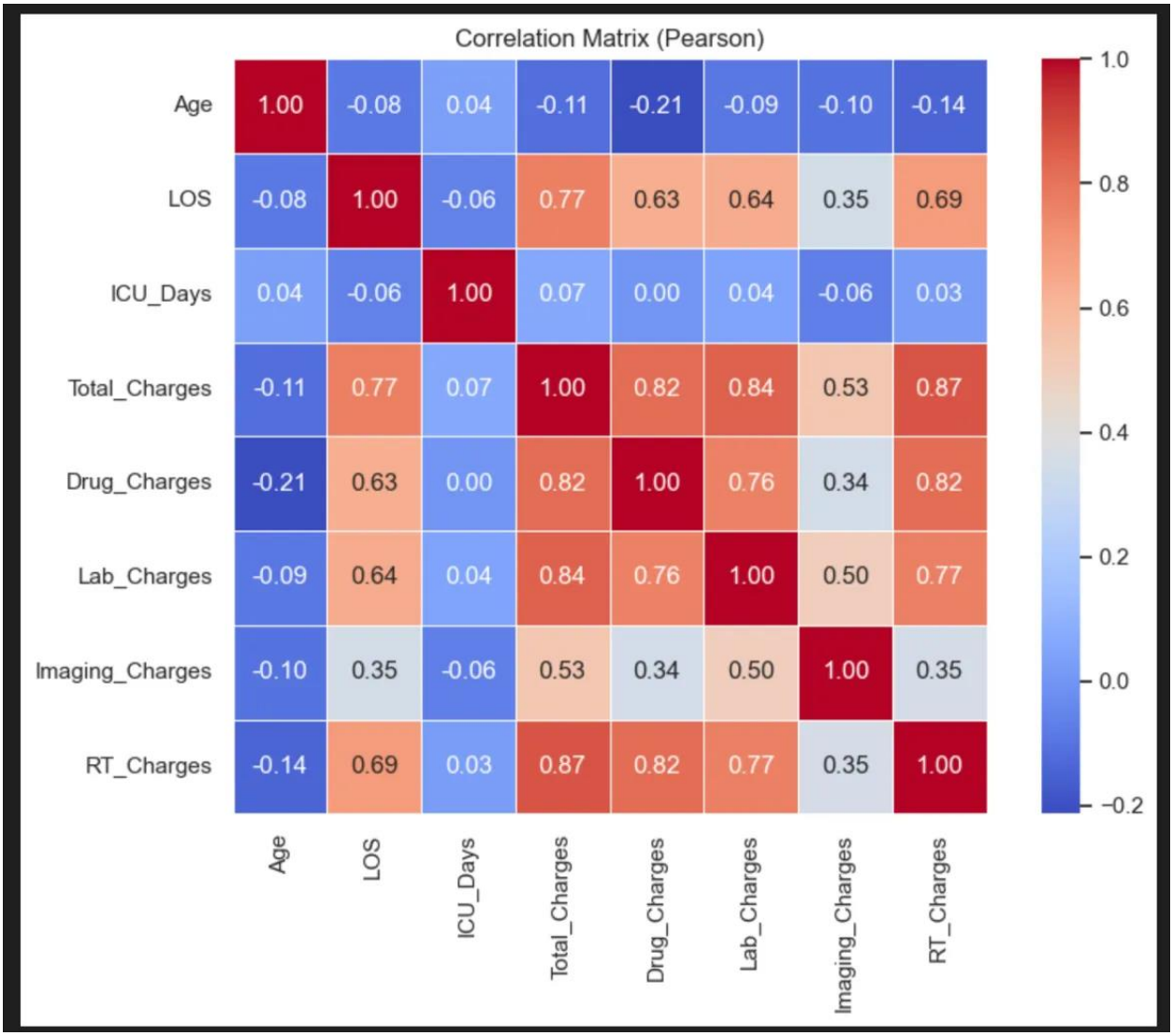


Figure 6: Correlation matrix

The Strongest Predictor: RT Charges

The greatest predictor in our dataset is respiratory treatment (RT) charges, which show the highest association with total charges (0.87). According to this remarkably strong correlation, RT charges alone account for almost 76% of the variation in overall charges ($r^2 = 0.76$).

Other Strong Predictors:

Beyond RT charges, several other variables show strong predictive power:

Lab Charges (r = 0.84): The complexity of patient treatment and the cost of diagnostic testing are reflected in the strong correlation between laboratory testing costs and overall charges.

Drug Charges (r = 0.82): Another important predictor is pharmaceutical costs, which show that overall costs are driven by medication-intensive treatments.

Length of Stay (r = 0.77): As discussed, stay duration remains a powerful predictor of total costs.

Imaging Charges (r = 0.53): Diagnostic imaging costs show a moderate positive relationship with total charges

Variables with Minimal Predictive Value:

Interestingly, some variables show little to no predictive utility:

Age (r = -0.11): There is almost no correlation between the age of the patient and the overall charges, indicating that the severity and kind of medical problem, rather than age, determines costs.

ICU Days (r = 0.07): The low correlation in intensive care utilization is probably caused by the small number of patients in this sample who needed ICU services.

Building a Prediction Model

These results allow us to develop a reliable approach for forecasting future patient total charges. Since RT charges have a correlation of 0.87, this would be the most basic single-variable prediction. We can reasonably predict a patient's total charges if their RT charges are known or can be estimated.

A multiple regression model that includes RT charges, lab charges, and drug charges would probably yield even higher accuracy for a more thorough prediction because these three variables capture various aspects of the intensity of care and have strong correlations with total charges (0.87, 0.84, and 0.82, respectively).

Additionally, length of stay is a practical indicator that can be forecast based on diagnosis and treatment plans, making it especially helpful for early cost estimation. The length of stay alone accounts for almost 59% of the variation in charges, with a correlation of 0.77.

Conclusions

This analysis successfully identifies key predictors of hospital charges, with RT charges emerging as the strongest single predictor (r = 0.87). The typical patient is a 75-year-old individual staying approximately 4.4 days with average charges of $6,446, though

substantial variation exists. Gender does not significantly influence costs, while length of stay shows a strong positive relationship with charges.

The analysis has successfully identified major determinants of hospital charges, with RT charges emerging as the greatest single predictor ($r = 0.87$). Although there is significant variety, the average patient consists of a 75-year-old who stays for about 4.4 days and charges $6,446. While duration of stay has a substantial positive correlation with charges, gender has little to no effect on expenses.