Recommendation Pricing Model at General Hospital

Mohammad Abu Talha Sunny

Package Pricing at General Hospitals

Package Pricing: A patient is provided a tailored quote on treatment cost at the time of admission for a group of related services, based on the expected costs for a clinically defined episode of care

Why Package Pricing at General Hospital?

Increase customer confidence and make pricing policies more transparent

Dr. Eric at General Hospital was at a crossroad to:

- Decide whether to use package pricing or traditional pricing
- Design a strategy as an accurate approach to predict the package price at time of admission
- How to use package pricing as a competitive strategy

Conceptual Model

Medical Data

Key complaint codes

Past medical history code

Implant (Y/N)

Personal Data

Age

Gender

BMI

Marital Status

Stay at hospital

Total Length Of Stay

Length of stay-ICU

Length of stay-Ward

Mode Of Arrival

State at Arrival

Type Of Admission

Symptoms

HR Pulse

BP -high

BP-low

RR

HB

Urea

Creatinine

The five assumptions of regression analysis

- **Linear relationship:** There exists a linear relationship between the independent and dependent variable
- **No or little multicollinearity**: more than two explanatory variables should not be highly linearly related
- **No autocorrelation**: no correlation between residuals
- **Normality**: The residuals of the model are normally distributed
- **Homoscedasticity**: The residuals have constant variance at every level of independent variable

Data Preparation

Age Categories (per case study appendix)

age <10: Child

age 11-25: Young Adult

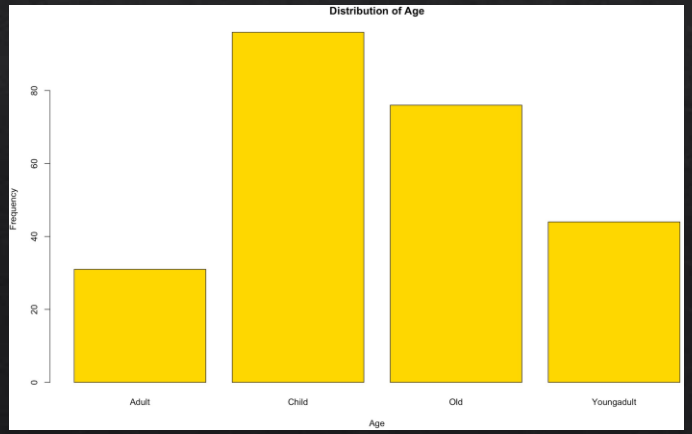
age 26-50: Adult

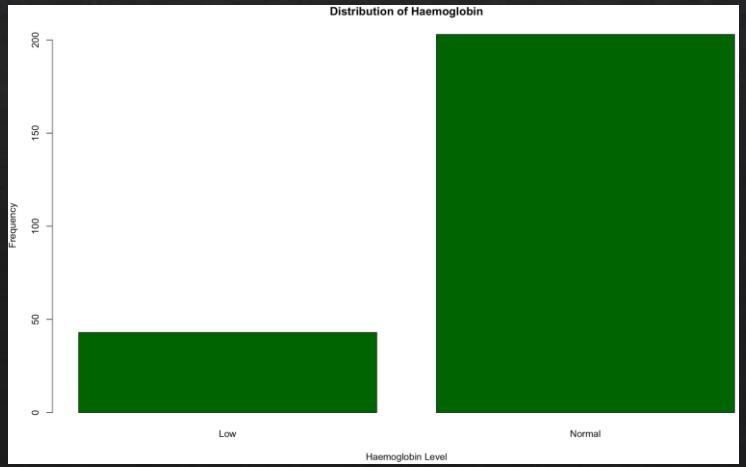
age >50: Old

Haemoglobin (per subject matter knowledge)

"normal": Female 12 to 15.5 and Men 13 to 17.5

Any value outside these limits will be "abnormal"



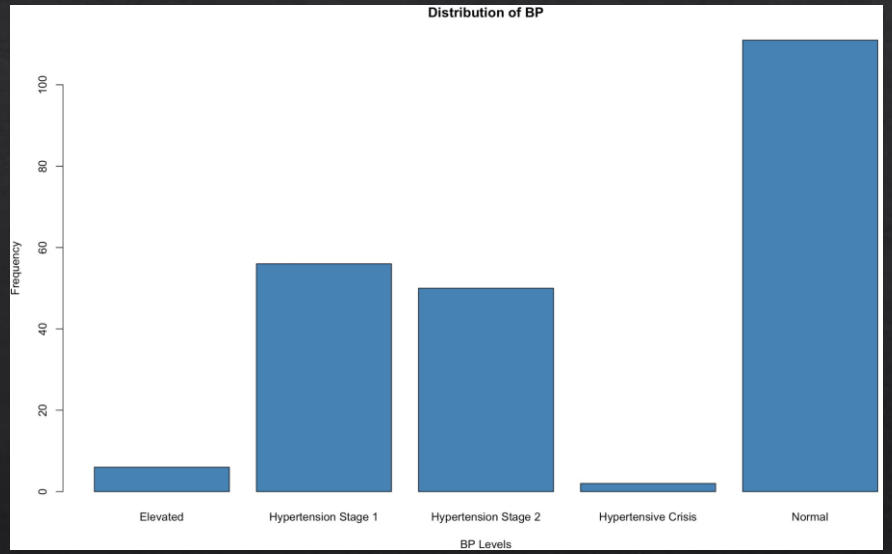


Derived Variables

BP Ranges (per subject matter knowledge)

Blood Pressure Category	Systolic mm Hg (upper #)		Diastolic mm Hg (lower #)
Normal	less than 120	and	less than 80
Elevated	120-129	and	less than 80
High Blood Pressure (Hypertension) Stage 1	130-139	or	80-89
High Blood Pressure (Hypertension) Stage 2	140 or higher	or	90 or higher
Hypertensive Crisis (Seek Emergency Care)	higher than 180	and/or	higher than 120

Source: American Heart Association



Urea Categories (Per subject matter knowledge)

Female: 6 to 21 mg/dl: Normal

Male: Urea 7 to 20 mg/dl: Normal

Urea >20 mg/dl: Abnormal

Data Preparation

BMI Categories (Per subject matter knowledge)

BMI <18.5: Underweight

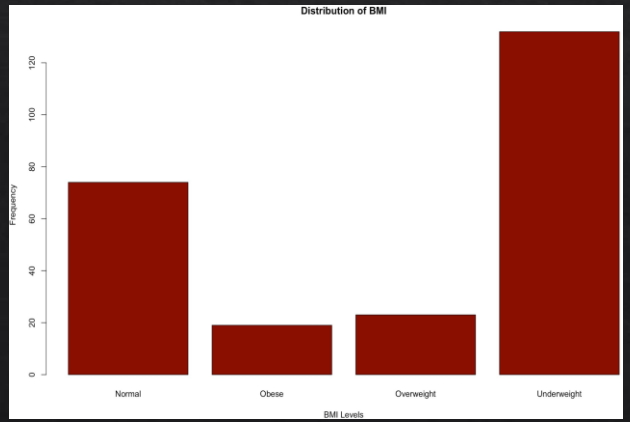
BMI <25: Normal

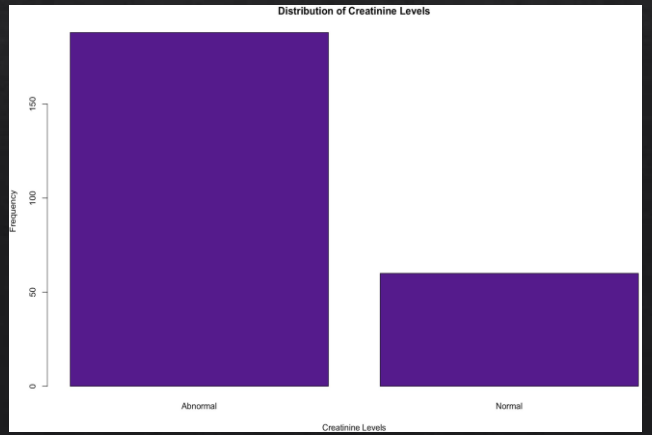
BMI 25 - 30: Overweight

BMI >30: Obese

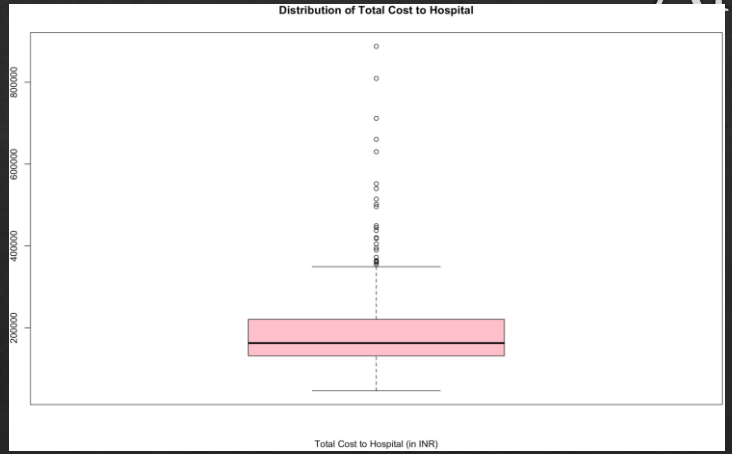
Creatinine Categories (Per subject matter knowledge)

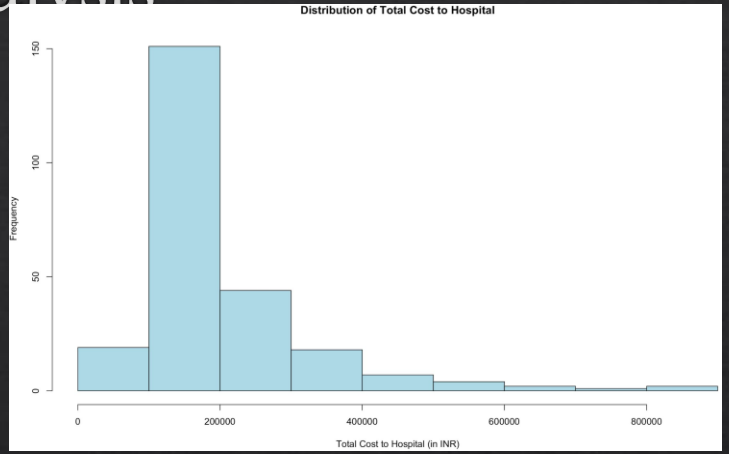
Age <3 & creatinine: 0.3-0.7	Normal
Age: 3-18 & creatinine: 0.5-1.0	Normal
Age >18 & Female & creatinine: 0.6 - 1.1	Normal
Age >18 & Male & creatinine: 0.9 - 1.3	Normal
Else:	Abnormal

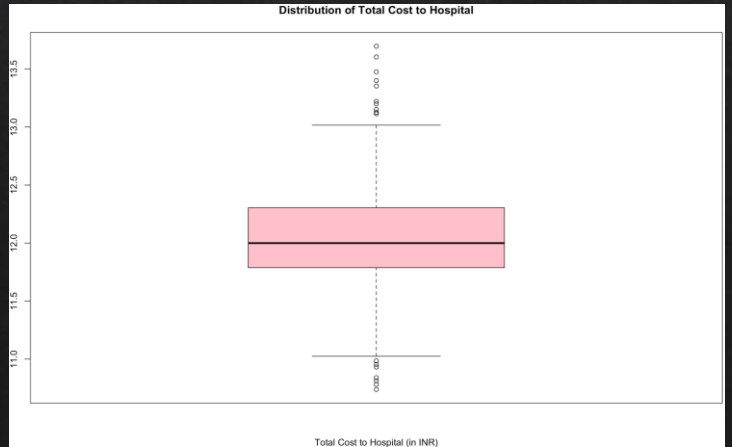


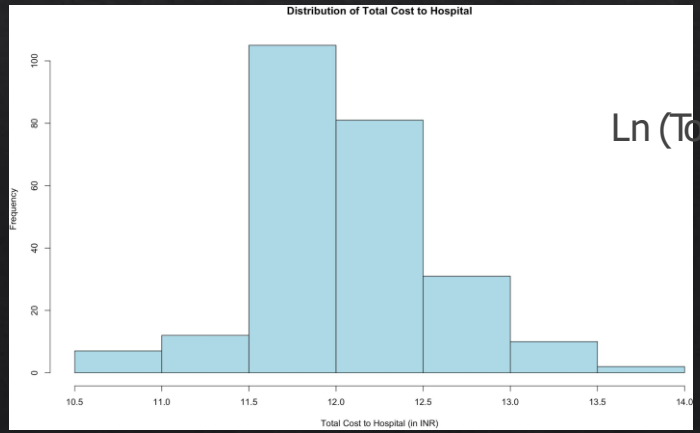


Dependent Variable - Univariate Analysis





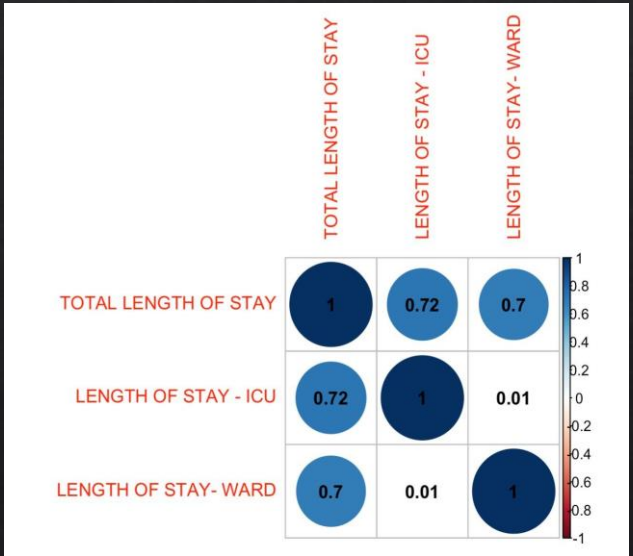


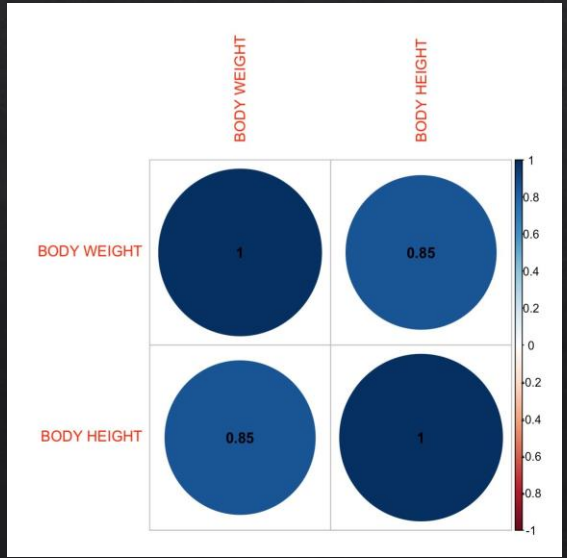


Ln (Total Cost)

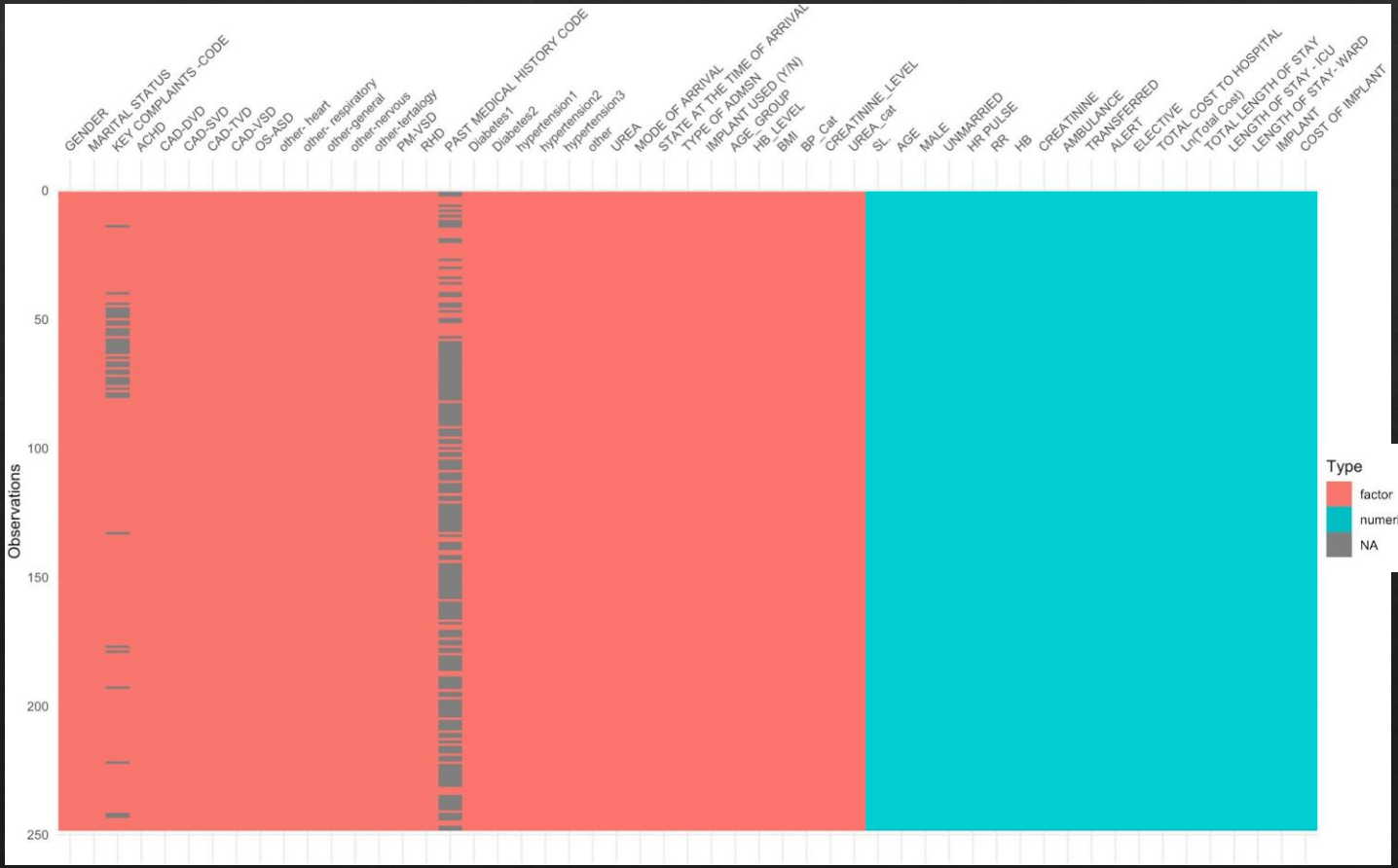
Data Preparation

- Total Length of Stay is highly correlated with dependent variable, Length of Stay in the ICU and Ward
- Body Height and Body Weight are highly correlated → created new variable BMI to avoid multicollinearity
- BP-High and BP-Low are correlated as well → defined new variable with both BP variables





Handling NULL values



- BP imputed 'Normal' for null values with juvenile patients (age <10)
- Urea Imputed 'Normal' for 11 null values and outliers (Assumption: Urea measurement is not critical for that patient)

Statistical Tests & Variable Reduction

Statistical tests were performed on variables and removed following variables on account of statistical insignificance

other-heart

other-nervous

other-respiratory

Diabetes1

Hypertension2

Hypertension3

Haemoglobin

PM-VSD

Gender

CAD-SVD

CAD-VSD

Creatinine

Urea

Impact of Body Weight on Total Cost

- Body weight and total cost relationship
- Equation:
  Ln(Total Cost) = 11.745+0.0084 (Body Weight)
- With every unit increase in the weight, there will be 0.84% increase in the logarithmic total cost of the treatment
- The average cost for a patient weighing 50 kg is
  $ 198,723
  =198723*0.0084
  =$ 1669.27
- A patient weighing 51 kg will pay $ 1,669 more than a patient weighing 50 kg

```
> linear <- lm(data$`Ln(Total Cost)`~data$`BODY WEIGHT`)
> summary.fit<- summary(linear)
> summary.fit

Call:
lm(formula = data$`Ln(Total Cost)` ~ data$`BODY WEIGHT`)

Residuals:
     Min       1Q   Median       3Q      Max
-1.35444 -0.28017 -0.02519  0.23823  1.51239

Coefficients:
                    Estimate Std. Error t value             Pr(>|t|)
(Intercept)        11.745190   0.056638 207.372 < 0.0000000000000002 ***
data$`BODY WEIGHT`  0.008442   0.001285   6.568        0.000000000301 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.4671 on 246 degrees of freedom
Multiple R-squared:  0.1492,    Adjusted R-squared:  0.1457
F-statistic: 43.14 on 1 and 246 DF,  p-value: 0.0000000003015
```

Multiple regression models to identify significant predictors

```
Call:
lm(formula = `Ln(Total Cost)` ~ MALE + UNMARRIED + BP_Cat + Diabetes1 +
    Diabetes2 + hypertension1 + hypertension2 + hypertension3 +
    other + ACHD + `CAD-DVD` + `CAD-SVD` + `CAD-TVD` + `OS-ASD` +
    `CAD-VSD` + `PM-VSD` + RHD + `other- respiratory` + `other-general` +
    ACHD + `CAD-DVD` + `CAD-SVD`, data = train.data)

Residuals:
     Min       1Q   Median       3Q      Max
-0.95413 -0.21414 -0.00657  0.21789  1.13104

Coefficients:
                                Estimate Std. Error t value             Pr(>|t|)
(Intercept)                     12.19355    0.18734  65.087 < 0.0000000000000002 ***
MALE                            -0.03801    0.06837  -0.556              0.57902
UNMARRIED                       -0.17177    0.08688  -1.977              0.04974 *
BP_CatHypertension Stage 1      -0.12165    0.19774  -0.615              0.53929
BP_CatHypertension Stage 2      -0.13300    0.19542  -0.681              0.49712
BP_CatNormal                    -0.08730    0.19674  -0.444              0.65783
Diabetes11                      -0.15693    0.16261  -0.965              0.33598
Diabetes21                       0.36842    0.18668   1.974              0.05016 .
hypertension11                  -0.02541    0.12111  -0.210              0.83408
hypertension21                  -0.25074    0.13321  -1.882              0.06161 .
hypertension31                   0.11730    0.23591   0.497              0.61970
other1                          -0.11613    0.12753  -0.911              0.36385
ACHD1                           -0.23655    0.12783  -1.851              0.06608 .
`CAD-DVD`1                       0.37528    0.12559   2.988              0.00325 **
`CAD-SVD`1                       0.31614    0.30623   1.032              0.30346
`CAD-TVD`1                       0.29642    0.12563   2.360              0.01950 *
`OS-ASD`1                       -0.10582    0.13418  -0.789              0.43148
`CAD-VSD`1                      -0.08119    0.39667  -0.205              0.83808
`PM-VSD`1                       -0.08649    0.20355  -0.425              0.67148
RHD1                             0.44661    0.10177   4.388             0.0000207 ***
`other- respiratory`1           -0.16659    0.12474  -1.336              0.18361
`other-general`1                -1.02470    0.41983  -2.441              0.01575 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

```
Call:
lm(formula = `Ln(Total Cost)` ~ BMI + AGE_GROUP + AMBULANCE +
    `COST OF IMPLANT` + `LENGTH OF STAY - ICU` + `LENGTH OF STAY- WARD` +
    `IMPLANT USED (Y/N)` + `MODE OF ARRIVAL` + `STATE AT THE TIME OF ARRIVAL`,
    data = train.data)

Residuals:
     Min       1Q   Median       3Q      Max
-0.91991 -0.09279  0.04334  0.14326  0.87071

Coefficients: (1 not defined because of singularities)
                                          Estimate   Std. Error t value             Pr(>|t|)
(Intercept)                           11.380417358  0.086036034 132.275 < 0.0000000000000002 ***
BMIObese                              -0.007735352  0.078117731  -0.099              0.921228
BMIOverweight                          0.008266533  0.069928652   0.118              0.906026
BMIUnderweight                        -0.048527874  0.067687767  -0.717              0.474317
AGE_GROUPAdult                         0.027943650  0.082957340   0.337              0.736617
AGE_GROUPOld                           0.193258934  0.074330352   2.600              0.010075 *
AGE_GROUPYoungadult                    0.029905539  0.057434923   0.521              0.603208
AMBULANCE                             -0.103882607  0.067215034  -1.546              0.123928
`COST OF IMPLANT`                      0.000005026  0.000001419   3.543              0.000502 ***
`LENGTH OF STAY - ICU`                 0.083440314  0.005200084  16.046 < 0.0000000000000002 ***
`LENGTH OF STAY- WARD`                 0.034338061  0.005112132   6.717       0.000000000222 ***
`IMPLANT USED (Y/N)`Y                  0.205508982  0.083255249   2.468              0.014479 *
`MODE OF ARRIVAL`TRANSFERRED          -0.265754721  0.132008724  -2.013              0.045547 *
`MODE OF ARRIVAL`WALKED IN                    NA           NA      NA                    NA
`STATE AT THE TIME OF ARRIVAL`CONFUSED 0.183725073  0.267501571   0.687              0.493057
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.2562 on 185 degrees of freedom
  (1 observation deleted due to missingness)
Multiple R-squared:  0.7395,    Adjusted R-squared:  0.7212
F-statistic: 40.39 on 13 and 185 DF,  p-value: < 0.00000000000000022
```

Recommended Model

Length of stay - ICU

Length of stay - Ward

Mode of Arrival: Transferred

Rheumatic Heart Disease (RHD)

Coronary Artery Disease - Double Vessel Disease (CAD-DVD)

Coronary Artery Disease - Triple Vessel Disease (CAD-TVD)

Cost of Implant

Age group: OLD

Unmarried

Model	Adjusted R-squared	RMSE
1: medical data	29.87%	0.5043
2: others	84.28%	0.2572
Ensemble	80.64%	0.2329

Recommended Model & Inferences

	97.5 %
(Intercept)	11.710760576637
UNMARRIED	0.081623656367
ACHD1	-0.081805955343
`CAD-DVD`1	0.293856441737
`CAD-TVD`1	0.333433720176
`other-general`1	-0.437994768502
`COST OF IMPLANT`	0.000009997105
`LENGTH OF STAY - ICU`	0.086768261208
`LENGTH OF STAY- WARD`	0.043175192270
AGE_GROUPAdult	0.095868638563
AGE_GROUPOld	0.244623495655
AGE_GROUPYoungadult	0.102291052774
`MODE OF ARRIVAL`TRANSFERRED	0.101009214297
`MODE OF ARRIVAL`WALKED IN	0.153457432745
RHD1	0.245653824249

Ln(Total Cost) ~a_1*Diabetes2 +a_2*ACHD + a_3*CAD-DVD +a_4*CAD-TVD +a_5*other-general + a_6*COST OF IMPLANT + a_7*LENGTH OF STAY - ICU + a_8*LENGTH OF STAY- WARD +a_9*IMPLANT USED (Y/N) +a_{10}*AGE_GROUP

For a patient with CAD-DVD, predicted cost of treatment can increase by 29.38% when compared with a person who does not come in with CAD-DVD

If a patient spends one more day in the ICU, the cost increases by 8.67% and a day more in ward increases the cost by 4.31%

A patient aged 51 years will have the predicted cost of treatment increased by 24.4% when compared to a patient of 50 years

Should General Hospital adopt Package Pricing?

Potential Advantages	Affected Party
Decreased health care costs and improved care coordination	Payers, Patients
Discourage unnecessary care	Payers, Patients
Strong incentive to avoid complications and readmissions	Payers, Patients
Increase transparency for costs of care	Payers, Patients
Expanded referral base and increased market share due to preferred agreements	Providers

Potential Disadvantages	Affected Party
Difficulty defining discrete episodes of care for chronic conditions	Providers
Potential avoidance of necessary specialty care	Providers, Patients
May encourage unnecessary episodes of care	Payers, Patients
Unclear accounting for value of academic endeavors (teaching, research)	Providers
Implementation challenges	Payers, Providers