
# CAPSTONE PROJECT 3 – PACKAGE PRICING AT MISSION HOSPITAL

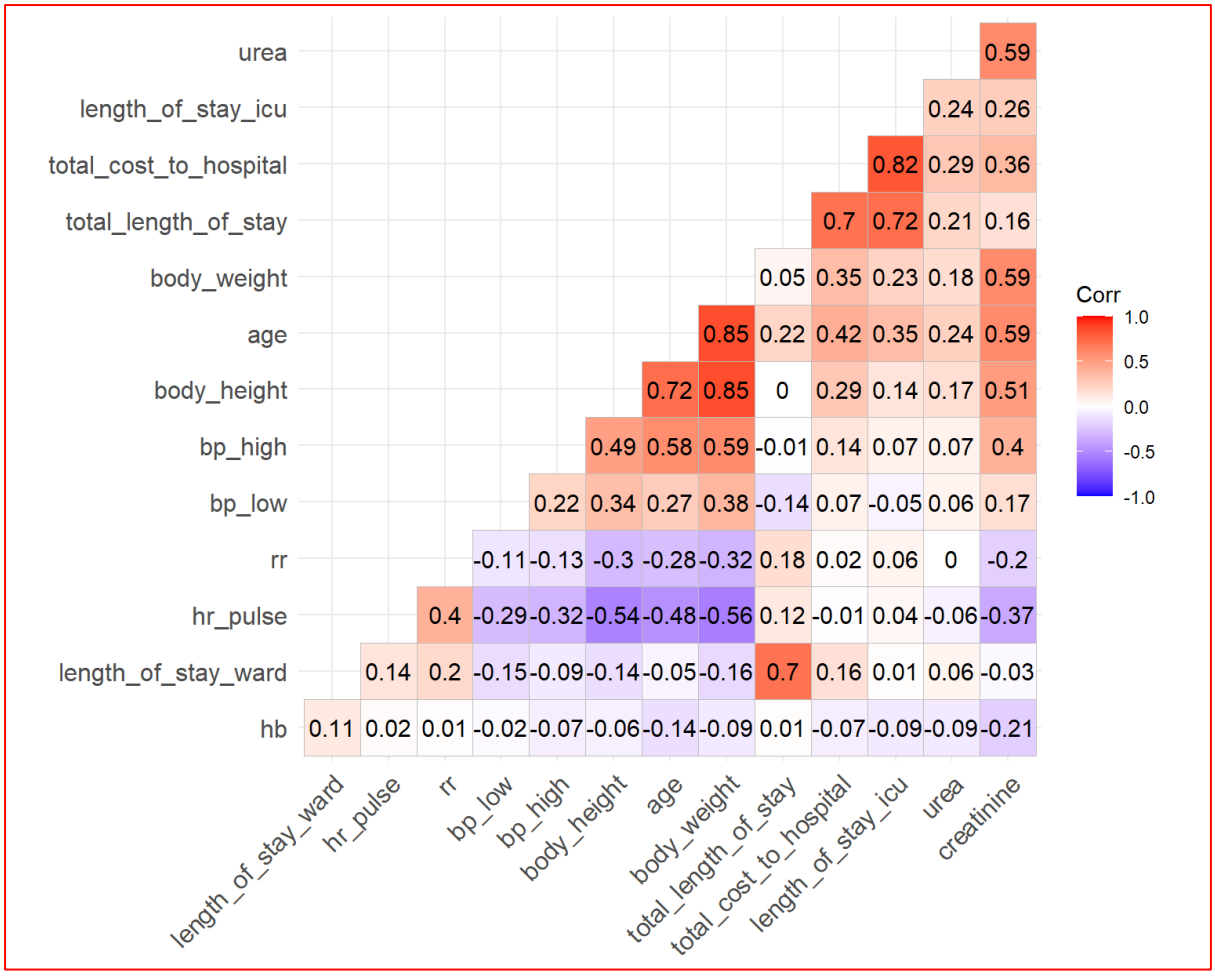
**Q:** The objective is to develop a package pricing strategy for treating ailments of economically disadvantaged individuals, considering market demand and competition. This pricing strategy should be tailored to specific treatments and account for regional variations in pricing due to different state policies.

Furthermore, the package pricing strategy should address the perception of medical tourism and consider its impact on the healthcare sector. It is important to analyze and understand the differences between traditional and conventional strategies in order to support the transition to the conventional mode of package pricing.

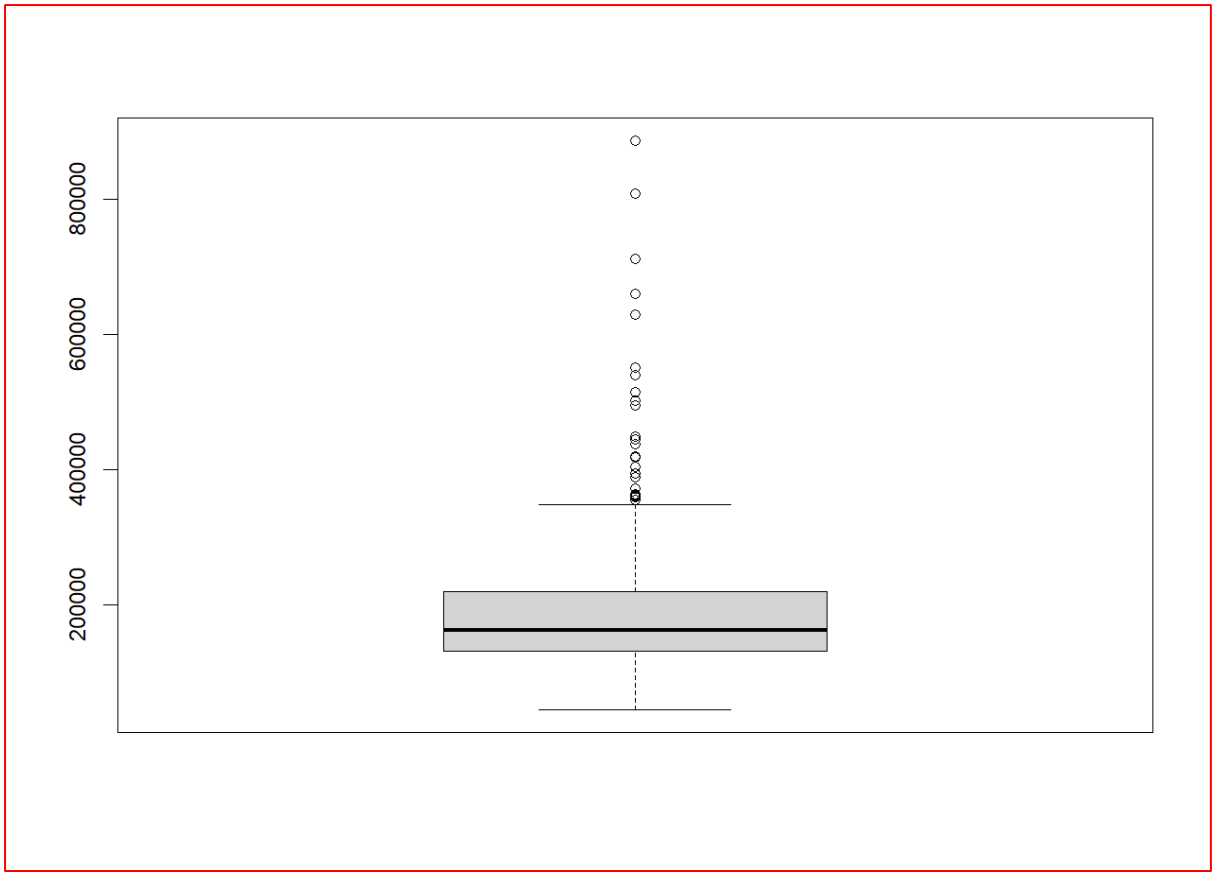
**Sol:-**

```
# Let's read the data;
data_MH <- read.csv("clipboard", sep = "\t", header = T)
# Install some packages to do this analysis;
install.packages("dplyr")
library(dplyr)

install.packages("mice")
library(mice)

install.packages("ggcorrplot")
library(ggcorrplot)

install.packages("fastDummies")
library(fastDummies)

install.packages("lmtest")
library(lmtest)

install.packages("lm.beta")
library(lm.beta)

install.packages("car")
library(car)
```

```
install.packages("janitor")
library(janitor)

# Pre processing the data;

data1_MH <- clean_names(data_MH)
data1_MH$past_medical_history_code[which(is.na(data1_MH$past_medical_histoty_code))]
<- "None"
data2_MH <- data1_MH %>% mutate_if(is.character, as.factor)
data3_MH <- subset(data2_MH, select = -c(cost_of_implant))
# MICE imputation to convert all NA Values
set.seed(1234)
data4_MH <- mice(data3_MH)
# From the imputed values, we have chosen the variable 'Creatinine' for mean comparison
with the imputed values.
mean(data3_MH$creatinine, na.rm = T)
[1] 0.7469767
imp.values <- data4_MH$imp$creatinine
mean(imp.values$"1")
[1] 0.630303
> mean(imp.values$"2")
[1] 0.7242424
> mean(imp.values$"3")
[1] 0.6333333
> mean(imp.values$"4")
[1] 0.6272727
> mean(imp.values$"5")
[1] 0.6969697

# Since the mean value has most proximity to the 1st imputed value, we are selecting that
imputation
data5_MH <- complete(data4_MH, 1)
# Pre-processing of data
names(data5)[23] <- "implant_used"
data5_MH$bp_low <- as.numeric(data5_MH$bp_low)
data6_MH <- subset(data5_MH, select = -c(sl))
# Correlation matrix for creating correlation plot
data6.1 <- data6_MH %>% select_if(is.numeric)
cor.matrix <- cor(data6.1)
```




```
# Creating fast dummies of categorical variables
data7_MH <- dummy_cols(data6_MH, remove_most_frequent_dummy = T,
remove_selected_columns = T)
# Outlier detection and removal from dependent variable
outlier <- boxplot(data7_MH$total_cost_to_hospital)$out
> length(outlier)
[1] 24
> outlier.data <- data7_MH[which(data7_MH$total_cost_to_hospital %in% outlier),]
> data8_MH <- data7_MH[-which(data7_MH$total_cost_to_hospital %in% outlier),]
> mode11 <- lm(total_cost_to_hospital ~., data8_MH)
> options(scipen = 100)
```




```
> summary(mode11)
Call:
lm(formula = total_cost_to_hospital ~ ., data = data8_MH)
Residuals:
   Min     1Q Median     3Q    Max
-63942 -16879      0  14833  92114
```

```
## Coefficients:
##                              Estimate Std. Error t value
## (Intercept)                   54750.27   30864.34   1.774
## age                             169.28     268.14   0.631
## body_weight                     -26.85     289.19  -0.093
## body_height                     158.60     109.71   1.446
## hr_pulse                         81.23     141.11   0.576
## bp_high                          10.65     126.05   0.084
## bp_low                          293.45     340.27   0.862
## rr                            -1240.58     725.12  -1.711
## hb                             -773.45     763.61  -1.013
## urea                            -27.06     199.48  -0.136
## creatinine                    20219.81    8849.99   2.285
```

```
## total_length_of_stay                  444.89   10240.39   0.043
## length_of_stay_icu                   17674.46   10335.92   1.710
## length_of_stay_ward                   4977.71   10145.10   0.491
## gender_F                             732.94    4719.94   0.155
## marital_status_MARRIED                -8416.46   10596.59  -0.794
## key_complaints_code_ACHD              -5601.42    7930.75  -0.706
## `key_complaints_code_CAD-DVD`          9255.87   10807.93   0.856
## `key_complaints_code_CAD-SVD`         -43194.20   24456.59  -1.766
## `key_complaints_code_CAD-TVD`           -85.25   11194.69  -0.008
## `key_complaints_code_CAD-VSD`         -10271.08   22180.11  -0.463
## `key_complaints_code_OS-ASD`            240.09    8809.88   0.027
## `key_complaints_code_other- respiratory`  3421.12    7967.69   0.429
## `key_complaints_code_other-general`     -39374.88   23354.88  -1.686
## `key_complaints_code_other-nervous`      1187.39   14364.91   0.083
## `key_complaints_code_other-tertalogy`    31663.00    8392.36   3.773
## `key_complaints_code_PM-VSD`            25371.32   13868.09   1.829
## key_complaints_code_RHD                -4181.06    9791.36  -0.427
## past_medical_history_code_Diabetes1      8471.51   12390.38   0.684
## past_medical_history_code_Diabetes2     46771.15   19210.71   2.435
## past_medical_history_code_hypertension1   2939.12   10837.34   0.271
## past_medical_history_code_Hypertension1  -23568.57   19932.47  -1.182
## past_medical_history_code_hypertension2  -11397.01   10192.43  -1.118
## past_medical_history_code_hypertension3    796.33   17357.20   0.046
## past_medical_history_code_other         -11950.97    8606.25  -1.389
## mode_of_arrival_AMBULANCE               47526.71   32619.11   1.457
## mode_of_arrival_TRANSFERRED            -30276.07   16215.94  -1.867
## state_at_the_time_of_arrival_CONFUSED    -5647.34   45179.48  -0.125
## type_of_admsn_EMERGENCY                -65286.02   31980.45  -2.041
## implant_used_Y                       78650.77    8473.03   9.282
##                                      Pr(>|t|)
## (Intercept)                             0.077733 .
## age                                    0.528626
## body_weight                              0.926126
## body_height                              0.149986
## hr_pulse                                0.565557
```

```
## bp_high                                  0.932780
## bp_low                                   0.389598
## rr                                     0.088789 .
## hb                                       0.312448
## urea                                     0.892237
## creatinine                               0.023470 *
## total_length_of_stay                     0.965394
## length_of_stay_icu                       0.088951 .
## length_of_stay_ward                      0.624258
## gender_F                                 0.876766
## marital_status_MARRIED                   0.428067
## key_complaints_code_ACHD                 0.480900
## `key_complaints_code_CAD-DVD`            0.392892
## `key_complaints_code_CAD-SVD`            0.079028 .
## `key_complaints_code_CAD-TVD`            0.993932
## `key_complaints_code_CAD-VSD`            0.643857
## `key_complaints_code_OS-ASD`             0.978288
## `key_complaints_code_other- respiratory` 0.668153
## `key_complaints_code_other-general`      0.093503 .
## `key_complaints_code_other-nervous`      0.934212
## `key_complaints_code_other-tertalogy`    0.000217 ***
## `key_complaints_code_PM-VSD`             0.068947 .
## key_complaints_code_RHD                  0.669867
## past_medical_history_code_Diabetes1      0.495014
## past_medical_history_code_Diabetes2      0.015861 *
## past_medical_history_code_hypertension1  0.786539
## past_medical_history_code_Hypertension1  0.238564
## past_medical_history_code_hypertension2  0.264946
## past_medical_history_code_hypertension3  0.963457
## past_medical_history_code_other          0.166621
## mode_of_arrival_AMBULANCE                0.146815
## mode_of_arrival_TRANSFERRED              0.063484 .
## state_at_the_time_of_arrival_CONFUSED    0.900662
## type_of_admsn_EMERGENCY                  0.042635 *
## implant_used_Y                         < 0.0000000000000002 ***
```

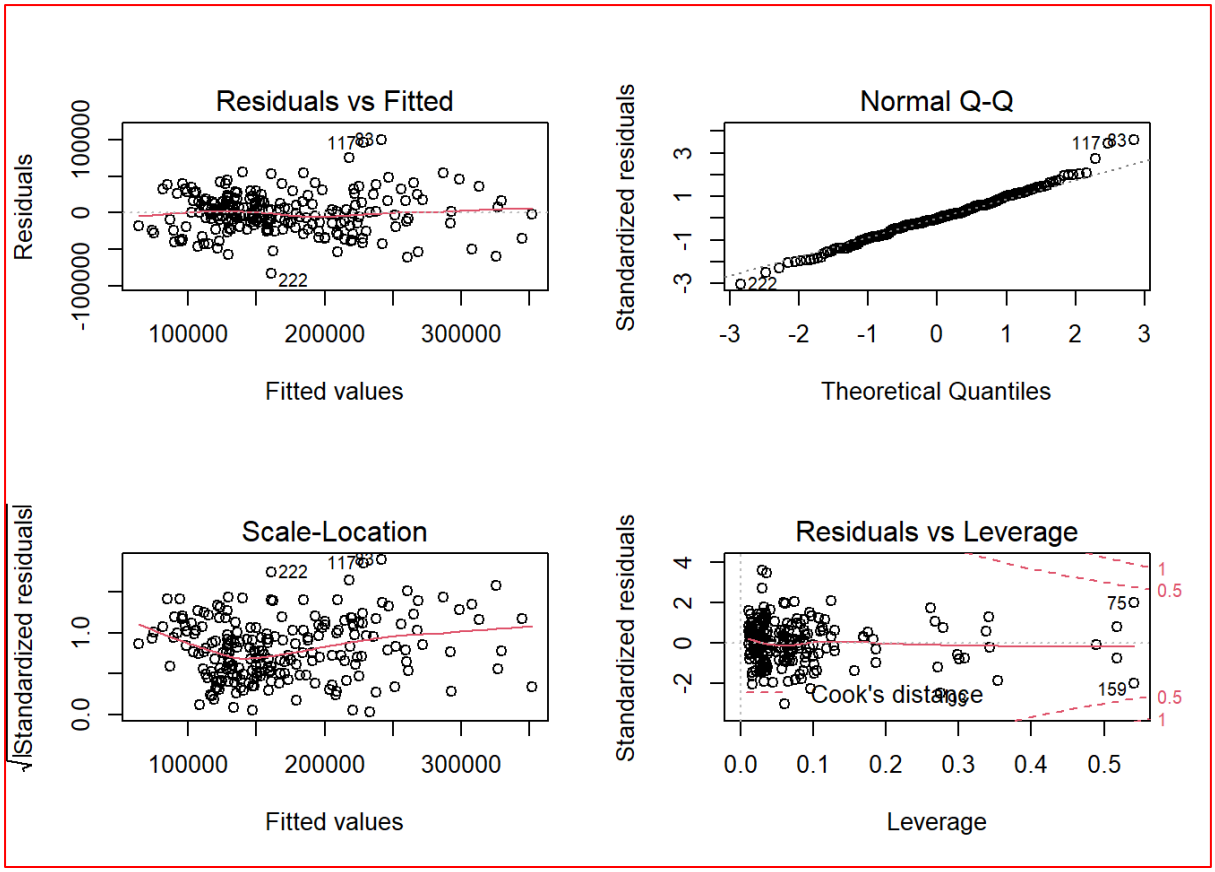
```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# Stepwise regression to identify the significant variables in the model
> mode12 <- step(mode11, trace = 0)
> summary(mode12)

Call:
lm(formula = total_cost_to_hospital ~ body_height + bp_low +
    rr + hb + creatinine + length_of_stay_icu + length_of_stay_ward +
    key_complaints_code_ + `key_complaints_code_CAD-DVD` + `key_complaints_code_CAD-
SVD` +
    `key_complaints_code_other-general` + `key_complaints_code_other-tertalogy` +
    `key_complaints_code_PM-VSD` + past_medical_history_code_Diabetes2 +
    past_medical_history_code_Hypertension1 + past_medical_history_code_hypertension2 +
    past_medical_history_code_other + mode_of_arrival_AMBULANCE +
    mode_of_arrival_TRANSFERRED + type_of_admsn_EMERGENCY + implant_used_y_n_Y,
    data = data8_MH)

Residuals:
   Min     1Q Median     3Q    Max
-66256 -14930   -120  15521  96471
```

```
## Coefficients:
##                                  Estimate Std. Error t value
## (Intercept)                       60894.73   20123.04   3.026
## body_height                         146.17      66.74   2.190
## rr                                -1277.62     659.11  -1.938
## creatinine                        19466.47    4845.00   4.018
## length_of_stay_icu                18431.69    1050.34  17.548
## length_of_stay_ward                5319.80     611.65   8.697
## `key_complaints_code_CAD-DVD`     11315.37    7625.61   1.484
## `key_complaints_code_CAD-SVD`    -39429.64   20902.43  -1.886
## `key_complaints_code_other-general`  -34886.66   21159.77  -1.649
## `key_complaints_code_other-tertalogy`  29709.33    6886.53   4.314
## `key_complaints_code_PM-VSD`      26046.62   12445.49   2.093
## past_medical_history_code_Diabetes2   43754.53   15289.69   2.862
## past_medical_history_code_Hypertension1 -24938.34   17751.26  -1.405
```

```
## past_medical_history_code_other        -11872.14    7883.21  -1.506
## mode_of_arrival_AMBULANCE               54838.55   30379.38   1.805
## mode_of_arrival_TRANSFERRED            -31423.99   14729.27  -2.133
## type_of_admsn_EMERGENCY                -72979.91   29536.34  -2.471
## implant_used_Y                          76820.61    5584.35  13.756
##                                        Pr(>|t|)
## (Intercept)                             0.00279 **
## body_height                             0.02964 *
## rr                                      0.05394 .
## creatinine                         0.00008231642922645 ***
## length_of_stay_icu                  < 0.0000000000000002 ***
## length_of_stay_ward                 0.00000000000000108 ***
## `key_complaints_code_CAD-DVD`           0.13937
## `key_complaints_code_CAD-SVD`           0.06065 .
## `key_complaints_code_other-general`     0.10073
## `key_complaints_code_other-tertalogy`  0.00002485689935067 ***
## `key_complaints_code_PM-VSD`            0.03759 *
## past_medical_history_code_Diabetes2     0.00465 **
## past_medical_history_code_Hypertension1 0.16156
## past_medical_history_code_other         0.13360
## mode_of_arrival_AMBULANCE               0.07252 .
## mode_of_arrival_TRANSFERRED             0.03407 *
## type_of_admsn_EMERGENCY                 0.01429 *
## implant_used_Y                      < 0.0000000000000002 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 28140 on 206 degrees of freedom
## Multiple R-squared:  0.8192, Adjusted R-squared:  0.8043
## F-statistic: 54.91 on 17 and 206 DF,  p-value: < 0.00000000000000022
```

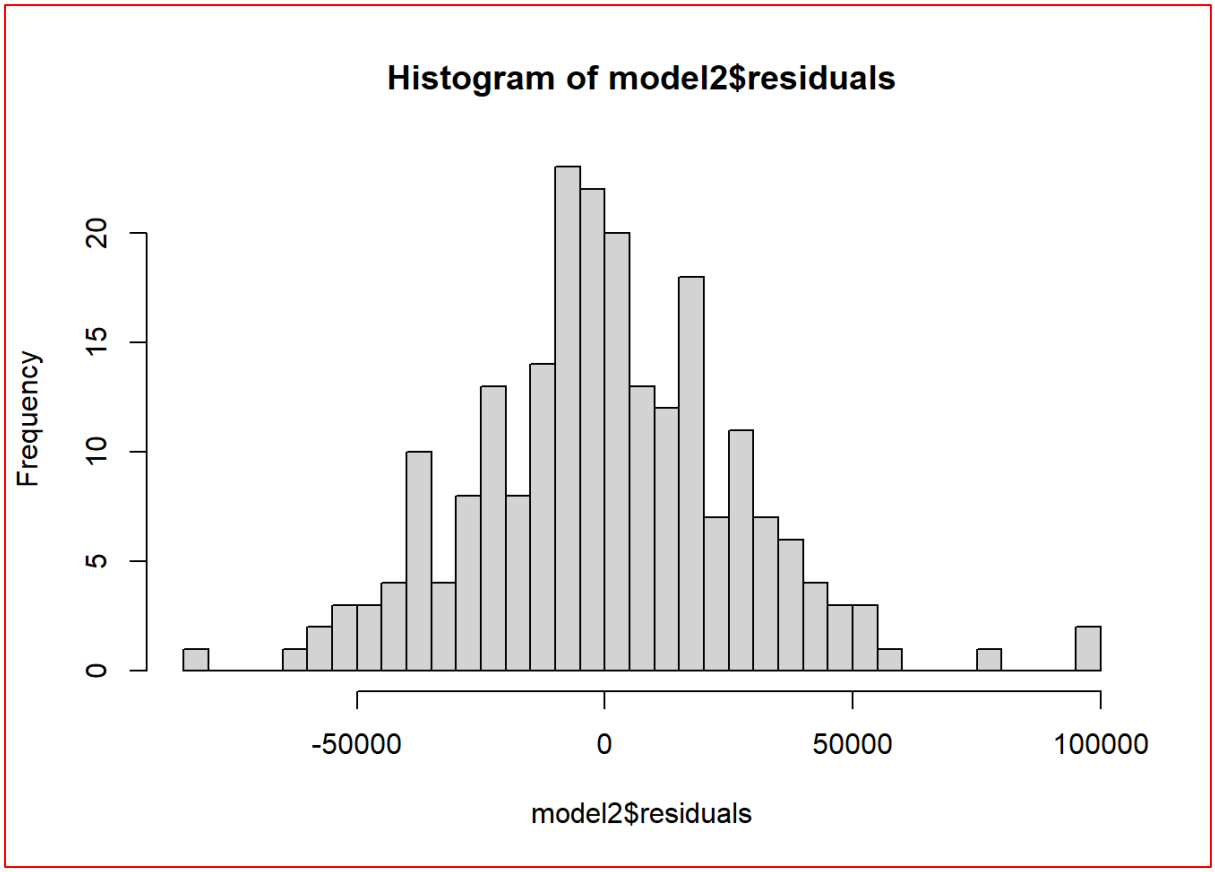
```
> par(mfrow = c(2,2))
> plot(mode12)
 72, 135, 142
```




> bptest(mode12)

studentized Breusch-Pagan test

data: mode12
BP = 43.247, df = 21, p-value = 0.002923
# Conducting Bruesh-Pagan test for checking heteroscedasticity
bptest(model2)
# studentized Breusch-Pagan test
data: model2
BP = 30.557, df = 17, p-value = 0.0226
# Getting the standard normal values of regression coefficients to identify the order of significance
par(mfrow= c(1,1))
hist(model2$residuals, breaks = 30)




```
library(lm.beta)
coeff <- lm.beta(model2)
std.coeff <- data.frame(coeff$standardized.coefficients)
std.coeff
```

```
##                                 coeff.standardized.coefficients
## (Intercept)                                          NA
## body_height                                    0.09025878
## rr                                            -0.06988629
## creatinine                                     0.15915995
## length_of_stay_icu                             0.59033228
## length_of_stay_ward                            0.27803070
## `key_complaints_code_CAD-DVD`                  0.05411276
## `key_complaints_code_CAD-SVD`                 -0.05843636
## `key_complaints_code_other-general`           -0.05170347
## `key_complaints_code_other-tertalogy`          0.13347249
## `key_complaints_code_PM-VSD`                   0.06625582
## past_medical_history_code_Diabetes2            0.09129211
```

```
## past_medical_history_code_Hypertension1          -0.04516408
## past_medical_history_code_other                  -0.04675378
## mode_of_arrival_AMBULANCE                         0.24071796
## mode_of_arrival_TRANSFERRED                      -0.06556492
## type_of_admsn_EMERGENCY                          -0.32787041
## implant_used_Y                                    0.45894933
```

```
# Checking multicollinearity
vif <- data.frame(vif(model2))
vif
```

```
##                                vif.model2.
## body_height                       1.935469
## rr                                1.481242
## creatinine                        1.788177
## length_of_stay_icu                1.289602
## length_of_stay_ward               1.164468
## `key_complaints_code_CAD-DVD`     1.515444
## `key_complaints_code_CAD-SVD`     1.093562
## `key_complaints_code_other-general`  1.120655
## `key_complaints_code_other-tertalogy`  1.090756
## `key_complaints_code_PM-VSD`      1.142085
## past_medical_history_code_Diabetes2  1.159703
## past_medical_history_code_Hypertension1  1.177712
## past_medical_history_code_other   1.098273
## mode_of_arrival_AMBULANCE        20.264324
## mode_of_arrival_TRANSFERRED       1.076247
## type_of_admsn_EMERGENCY          20.065060
## implant_used_Y                    1.268380
```

```
# Creating Training and test data
index <- sample(1:nrow(data8), 0.80*(nrow(data8)))
train_data <- data8[index,]
test_data <- data8[-index,]
# Creating model using training data
```

```
model3 <- lm(total_cost_to_hospital ~., train_data)
summary(model3)
```

```
##
## Call:
## lm(formula = total_cost_to_hospital ~ ., data = train_data)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -74361 -16277  -1021  16750  98322
##
## Coefficients:
##                                 Estimate Std. Error t value
## (Intercept)                      63658.99   37427.60   1.701
## age                                179.00     331.45   0.540
## body_weight                        112.12     342.78   0.327
## body_height                        134.46     135.38   0.993
## hr_pulse                           120.41     165.49   0.728
## bp_high                             84.75     161.20   0.526
## bp_low                             149.15     428.55   0.348
## rr                               -2112.67     899.73  -2.348
## hb                                 -743.55     903.77  -0.823
## urea                                -40.22     238.80  -0.168
## creatinine                        20591.93   11018.67   1.869
## total_length_of_stay               1801.76   11095.76   0.162
## length_of_stay_icu                16019.11   11266.66   1.422
## length_of_stay_ward                4091.76   10991.51   0.372
## gender_F                           -425.35    5733.43  -0.074
## marital_status_MARRIED           -15808.27   13072.06  -1.209
## key_complaints_code_ACHD          -2888.34    9608.64  -0.301
## `key_complaints_code_CAD-DVD`     12171.25   12967.32   0.939
## `key_complaints_code_CAD-SVD`    -48579.52   26482.56  -1.834
## `key_complaints_code_CAD-TVD`     1866.71   13541.02   0.138
## `key_complaints_code_CAD-VSD`    -11804.45   23759.70  -0.497
## `key_complaints_code_OS-ASD`      -3623.44   10449.68  -0.347
## `key_complaints_code_other- respiratory`  3953.95    9988.24   0.396
```

```
## `key_complaints_code_other-general`      -35389.31   25279.64  -1.400
## `key_complaints_code_other-nervous`       1482.35   16901.96   0.088
## `key_complaints_code_other-tertalogy`    32526.94   10070.20   3.230
## `key_complaints_code_PM-VSD`             30238.11   18038.72   1.676
## key_complaints_code_RHD                 -5806.00   11711.22  -0.496
## past_medical_history_code_Diabetes1      -1283.97   16255.29  -0.079
## past_medical_history_code_Diabetes2      58785.13   24715.09   2.379
## past_medical_history_code_hypertension1   3732.02   12800.26   0.292
## past_medical_history_code_Hypertension1  -24411.24   22023.71  -1.108
## past_medical_history_code_hypertension2  -14788.93   13160.41  -1.124
## past_medical_history_code_hypertension3      77.17   18868.69   0.004
## past_medical_history_code_other         -12209.15    9754.71  -1.252
## mode_of_arrival_AMBULANCE                55586.89   35531.88   1.564
## mode_of_arrival_TRANSFERRED             -44336.48   24991.05  -1.774
## state_at_the_time_of_arrival_CONFUSED   -16002.21   52088.37  -0.307
## type_of_admsn_EMERGENCY                 -74634.31   34485.79  -2.164
## implant_used_Y                           81968.11    9973.07   8.219
##                                        Pr(>|t|)
## (Intercept)                             0.09121 .
## age                                     0.59001
## body_weight                             0.74409
## body_height                             0.32234
## hr_pulse                                0.46809
## bp_high                                 0.59989
## bp_low                                  0.72834
## rr                                      0.02028 *
## hb                                      0.41207
## urea                                    0.86648
## creatinine                              0.06375 .
## total_length_of_stay                    0.87124
## length_of_stay_icu                      0.15732
## length_of_stay_ward                     0.71026
## gender_F                                0.94097
## marital_status_MARRIED                  0.22859
## key_complaints_code_ACHD                0.76417
```

```
## `key_complaints_code_CAD-DVD`                  0.34956
## `key_complaints_code_CAD-SVD`                  0.06873 .
## `key_complaints_code_CAD-TVD`                  0.89055
## `key_complaints_code_CAD-VSD`                  0.62010
## `key_complaints_code_OS-ASD`                  0.72930
## `key_complaints_code_other- respiratory`          0.69281
## `key_complaints_code_other-general`             0.16377
## `key_complaints_code_other-nervous`             0.93024
## `key_complaints_code_other-tertalogy`           0.00155 **
## `key_complaints_code_PM-VSD`                  0.09593 .
## key_complaints_code_RHD                       0.62084
## past_medical_history_code_Diabetes1             0.93716
## past_medical_history_code_Diabetes2             0.01874 *
## past_medical_history_code_hypertension1          0.77106
## past_medical_history_code_Hypertension1          0.26960
## past_medical_history_code_hypertension2          0.26306
## past_medical_history_code_hypertension3          0.99674
## past_medical_history_code_other                 0.21281
## mode_of_arrival_AMBULANCE                       0.11999
## mode_of_arrival_TRANSFERRED                     0.07824 .
## state_at_the_time_of_arrival_CONFUSED           0.75914
## type_of_admsn_EMERGENCY                         0.03216 *
## implant_used_Y                      0.000000000000129 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 30810 on 139 degrees of freedom
## Multiple R-squared:  0.8239, Adjusted R-squared:  0.7745
## F-statistic: 16.68 on 39 and 139 DF,  p-value: < 0.00000000000000022
```

```
> mode14 <- step(mode13, trace = 0)
> summary(mode14)

Call:
lm(formula = total_cost_to_hospital ~ body_height + hr_pulse +
    rr + hb + creatinine + length_of_stay_icu + length_of_stay_ward +
```

```
    key_complaints_code_ + key_complaints_code_ACHD + `key_complaints_code_CAD-SVD`
+
    `key_complaints_code_other-tertalogy` + `key_complaints_code_PM-VSD` +
    past_medical_history_code_hypertension1 + mode_of_arrival_AMBULANCE +
    mode_of_arrival_TRANSFERRED + state_at_the_time_of_arrival_CONFUSED +
    type_of_admsn_EMERGENCY + implant_used_y_n_Y, data = train_data)

Residuals:
   Min     1Q Median     3Q    Max
-62297 -14711    471  14209 103666
```

```
## Coefficients:
##                                  Estimate Std. Error t value
## (Intercept)                       77505.98   23332.76   3.322
## body_height                         180.91      74.66   2.423
## rr                                -2273.86     764.80  -2.973
## creatinine                        18362.56    5444.29   3.373
## length_of_stay_icu                18482.24    1209.90  15.276
## length_of_stay_ward                5692.97     687.81   8.277
## `key_complaints_code_CAD-SVD`       -47367.06   21820.58  -2.171
## `key_complaints_code_other-general`   -35057.81   22131.67  -1.584
## `key_complaints_code_other-tertalogy`  30585.50    8150.25   3.753
## `key_complaints_code_PM-VSD`          32859.88   15969.00   2.058
## past_medical_history_code_Diabetes2    54088.27   18379.04   2.943
## mode_of_arrival_AMBULANCE             55114.65   31645.29   1.742
## mode_of_arrival_TRANSFERRED          -43879.89   22437.95  -1.956
## type_of_admsn_EMERGENCY              -74398.91   30758.42  -2.419
## implant_used_Y                     77777.92    6288.27  12.369
##                                  Pr(>|t|)
## (Intercept)                       0.001103 **
## body_height                         0.016471 *
## rr                                0.003392 **
## creatinine                          0.000928 ***
## length_of_stay_icu                < 0.0000000000000002 ***
## length_of_stay_ward                0.0000000000000419 ***
## `key_complaints_code_CAD-SVD`       0.031388 *
## `key_complaints_code_other-general`   0.115107
```

```
## `key_complaints_code_other-tertalogy`            0.000242 ***
## `key_complaints_code_PM-VSD`                     0.041199 *
## past_medical_history_code_Diabetes2              0.003723 **
## mode_of_arrival_AMBULANCE                        0.083447 .
## mode_of_arrival_TRANSFERRED                      0.052210 .
## type_of_admsn_EMERGENCY                          0.016665 *
## implant_used_Y                    < 0.0000000000000002 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 29320 on 164 degrees of freedom
## Multiple R-squared:  0.8119, Adjusted R-squared:  0.7959
## F-statistic: 50.57 on 14 and 164 DF,  p-value: < 0.00000000000000022
```

```
# Using the model to predict values of test data and comparing the results
test_data$Predicted <- predict(model4, test_data)
comparison <- data.frame(Actual_cost =test_data$total_cost_to_hospital, Predicted_cost=
test_data$Predicted)
comparison
```

```
##    Actual_cost Predicted_cost
## 1     341109.0      273140.85
## 2     144037.2      126197.74
## 3     164962.0      107180.31
## 4     120131.0      133664.99
## 5     138923.0      156152.22
## 6     122892.0      111192.02
## 7     142552.0      134051.78
## 8     109085.8      106938.44
## 9     125643.0      118862.96
## 10    128196.0       95529.98
## 11    109085.8      106938.44
## 12    125643.0      120518.31
## 13    294615.9      296296.33
## 14    156576.9      154020.51
## 15    109575.6      167453.97
```

```
## 16    201219.0      202625.43
## 17    214679.0      270628.50
## 18    189701.5      178381.03
## 19    139723.0      152435.75
## 20    119685.6      117874.87
## 21    276458.0      260702.84
## 22    150337.0      130609.74
## 23    139067.0      146234.87
## 24    127899.0      110775.28
## 25    146355.0      170204.38
## 26     97060.8      119870.27
## 27    106070.0      111186.56
## 28    140372.0      155320.25
## 29    138769.4      128346.19
## 30     77241.0       74704.68
## 31     49700.0       73587.04
## 32    137273.0      100162.45
## 33    193543.0      228469.83
## 34    191102.0      234498.35
## 35    132585.0      160565.91
## 36    170654.0      162985.88
## 37    174074.0      128594.76
## 38    210622.0      210393.38
## 39     46093.0       50167.98
## 40    188824.0      196148.22
## 41    146700.0      144152.96
## 42    149462.0      158009.78
## 43    186450.0      227407.55
## 44    132997.0      132721.96
## 45    248112.0      265147.35
```

```
# Checking the validity of model using MAPE and RMSE
mape <- mean(abs(comparison$Actual_cost-
comparison$Predicted_cost)/comparison$Actual_cost)
mape
```

```
## [1] 0.1253493
```

```
1-mape
## [1] 0.8746507
rmse <- sqrt(mean(comparison$Actual_cost- comparison$Predicted_cost)^2)
rmse
## [1] 1226.584
```

***Conclusion:-***

The MAPE value of 0.1253493 indicates that, on average, the absolute percentage difference between the predicted values and the actual values is approximately 12.53%. This value represents the average relative error in percentage terms.

1. The MAPE value suggests that, on average, the model's predictions deviate by around 12.53% from the actual values.
2. The RMSE value indicates that the average prediction error, in absolute terms, is approximately 1226.584. This value gives an idea of the magnitude of the errors made by the model.

*Presented by;*
*Sarveswara Sarma Nemani*
17.06.2023

******************Thank You************************