

Patient Re- Admittance Analysis

Marc Navia
Data Science Team 1

Executive summary

Introduction

Our analytics team was given four datasets from a large hospital system on various data about patients, hospitals, and geography. The hospital system is having a problem with patients being readmitted back into their system after being discharged and is looking for insights on what factors could be causing this. Throughout this process, the analytics team will conduct Exploratory data analysis to get a deeper insight into the data. From those insights, multiple supervised machine learning algorithms such as logistic regression and decision trees will be used to predict the probability of a patient being readmitted.

Problem Statement

The hospital is looking for a way to find the probability of a patient being readmitted back into their system after being discharged. With the passage of the Affordable Care Act in 2012 came with the Hospital Readmission Reduction Program (HRRP). This program financially penalized hospital with a higher than expected risk-standardized 30-day readmission rates. This penalty will continue to grow with every passing year so reducing patient readmittance is a priority. With consumers becoming more conscious of their purchasing decision hospital could lose business to other hospital system who have a much lower readmittance rate due to the quality of care.

Data and Analysis

Before we begin any analysis, we should understand the data that we are working with. We were given four datasets that have information about the demographics of the patient, the condition of the patient in the hospital and information about the initial hospital stay. There also a dataset that has information about the geographical location of the patient. From this data, we will have to do data preprocessing to ensure the quality and integrity of the data. With the data finally clean and joined together, analysis can be done. Some of the modeling techniques that will be used are logistic regressions and decision trees. To gain further insight we will use visual analysis on the data using Tableau. This can help visualize our data and create a story for the insights that we found when exploring the datasets.

Findings/Recommendations

From the logistic regression model there were many insights that were gained. First, we found three variables with powerful predicting power. Those three variables were order total charges, the procedure long description and the diagnosis long description. For every \$1000 increase in the order total charge, the likelihood that a patient is readmitted back into the hospital increases by 1.32%. For the diagnosis long description patients with the conditions such *EMPHYSEMATOUS BLEB* and *OTHER STAPHYLOCOCCUS PNEUMONIA* were 17%

more likely to be readmitted back. Patients with *PERICARDIOCENTESIS* and *BIOPSY OF ABDOMINAL WALL* procedures were 10% and 13% respectively more likely to be readmitted back into the hospital.

Our recommendation for the hospital is to use the logistic regression model right after the patient has finished their operations. By inputting the necessary variables that were used in our logistic model the probability of the patient being readmitted will be established. Our data science team recommend a cutoff point of 50% to determine whether or not a patient will be readmitted or not. So any patient who gets a probability of $\geq 50\%$ should be given additional care and treatment to ensure that their risk of being readmitted back into the system is lower.

Limitations and Next Steps

Some of the limitations that could affect our modeling is the amount of data that we have received. While 160,000 record may seem like a lot to some, having a larger data set can allow the model to learn a lot better. There could be also important variables that have strong explanatory value for our target variable that has been left out of the dataset. The next steps for improving our analysis is to go out and gather more data. Through a limitation that can occur from this is that if we have got too much data we might not have enough computing power to make a proper analysis without investing more capital in improving our assets.

Appendix A: Technical Write-Up

Loading in the Data

A new SAS Project was created to begin the data mining and modeling process for the hospital data. The first task was to connect Enterprise Miner to the Library in SAS Studio. To do this the file path was taken from SAS studio and inserted into the project start code using a LIBNAME statement.

After the connection has been made, the data needs to be loaded into the SAS project data source. Using the data source wizard, the 90% hospital data set was loaded in using all the default wizard settings.

Once the data source has the desired dataset, the diagram can finally be created. The diagram will be used for some quick data exploration, modeling pre-processing and the logistic regression. When the diagram is created, the next thing that was done is the loading in of the data. From the data source tab, the table that we loaded in was drag onto the diagram. A new data node appears ending our data loading process

Exploratory Data Analysis

Our data science team used two tools to conduct exploratory data analysis on our dataset, SAS Studio and SAS Enterprise Miner. SAS Studio was used as the main tool for EDA since it provided more options with its plot. On SAS Studio histograms, scatter plots and bar and box plots were created for variables that we found interesting. Enterprise Miner was mostly used for quick summary statistics and also its multiplot node that provided insight in how the x variables and the target variable related.

Modeling Pre-processing

Input Data Node

In this node most of the work will be done in the Edit Variables dropdown selection. The main purpose of this node was to adjust the metadata of our dataset that we brought in. The variables that were brought in were carefully chosen to be either inputted or rejected. Most of the work was done by the wizard itself but some changes had to be made done manually. The target variable had to be manually chosen and variables such as longitude and latitude had to be rejected. After all the roles were assigned, our team went over each variable level to ensure that they were correctly chosen.

Multiplot Node

This node was chosen by our data science team as a way to explore our data through stacked bar charts to the target variable. Before using the multiplot node the statistic property was changed to percent for better readability. All of the other properties were left as default and the node was run. From this node we were able to gain additional insight of how our x variables are driving our target variable.

StatsExplore Node

For this node all of the properties were left to default. The main purpose of this node was to add additional exploratory data analysis through descriptive statistics. From the results we are able to know what categorical variables that have potential significance through a Chi-square plot. It also gives another bar graph that shows the possible worth of each x variable whether it is categorical or interval through log worth. Lastly, we are given descriptive statistics such as mean, median and mode while also additional information like the amount of missing records.

Data Partition Node

Before we can create the model, we must split up the data into two different sets, the training data set and the validation training set. By splitting our main dataset into two, we will be able to better assess the models that we create in the future with the use of the validation data set. Our models will learn from the training set and then validated in our validation data set. Our team choose to split the data 70/30. To do this we had to change the properties on the Data partition node by setting the training property field to 70% and the validation property field to 30%.

Impute Node

With the dataset split into two the final thing that must be done to finish the model preprocessing is filling in and removing missing data. The impute node provides many options for filling in missing data points. For filling in the missing categorical data points, the default input method property was changed to tree. We choose tree as our input method because it is a multi-layer approach that finds the closest neighbor based on certain rules and fills it in for the missing data point. For our interval data we change the default input method property to median instead of mean. We choose median since it isn't sensitive to outliers like mean is.

Modeling: Logistic Regression

Since the variable that we are predicting is binary, logistic regression made the most sense to be the model of choice for prediction. The logistic regression gives each record a percentage of how likely it is to be 1. For our problem it will give us the percentage of how likely a patient will be readmitted back into the hospital.

Regression Node

This node provides the choice of either linear or logistic regression. For our case we will use the logistic regression due to reasons mentioned above. Next, we choose stepwise for our selection model. Stepwise is a mixture of both the backwards and forwards method. We choose stepwise due to the nature of it removing variables that lose significance as it adds additional variables. Lastly, we use validation misclassification for the model selection criterion since we have a validation dataset and we are also dealing with a binary prediction problem.

After preparing the desired settings for our regression model we run the logistic regression analysis. From the model we get a misclassification rate of 14.92% for the training dataset and 15.06% for the validation dataset. The gap between the training and validation dataset is very tight showing that our model is strong and verifies that our model has been trained right.

The model found 16 variables that were statistically significant in predicting whether or not a patient would be readmitted or not. The four strongest drivers for our model in their respective order are `order_total_charges`, `Procedure_long_desc`, `diagnosis_long_desc` and `standard_orders_used`. These variables were identified based on how large their chi square value was. One thing that is possibly worrisome is the possible strength of the `order_total_charges` variable. Its chi-square value is significantly larger than its counterparts but after removing the variable from the model it caused a 4% increase in the model classification rate.

The variables that are important to our model are:

Intercept, DIAGNOSIS_LONG_DESC DISCHARGED_TO, Disch_Nurse_ID, HOSPITAL_ICU_DAYS, IMP_Department, IMP_Num_Chronic_Cond, IMP_order_total_charges, LENGTH_OF_STAY, PROCEDURE_LONG_DESC, STATECODE, Standard_Orders_Used, gender op_visits6, operationcount, race_cd

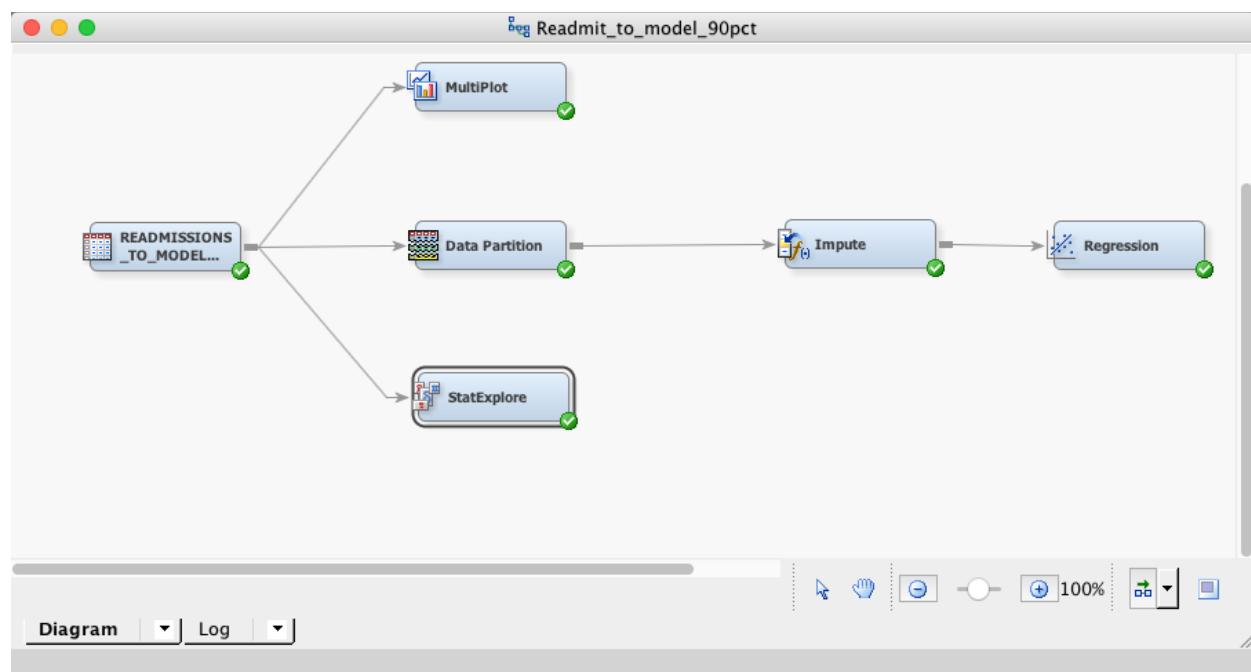
For our lift chart we looked at the 20% decile point to find the true predictive power of our model. From the lift chart we find that the cumulative lift at the 20% decile point was 2.8. That means that our model is 2.8 times more likely to predict whether a patient is readmitted back into the hospital compared to random guessing.

Appendix B: Screenshots

LIBRARY CONNECTION

```
1 Libname readmit '/home/manavia0/Hospital Data/';
2 run;
3
```

FINAL DIAGRAM



DATA SOURCE



BUS458Hospital



Data Sources



READMIT_TO_MODEL_90PCT

INPUT DATA METADATA

Name	Role	Level	Report	Order	Drop	Lower Limit	Upper Limit
ActualPropoRejected		Unary	No	No		.	.
ADMIT_DATERejected		Interval	No	No		.	.
AllocProportRejected		Unary	No	No		.	.
CITY_Rejected		Nominal	No	No		.	.
County_namRejected		Nominal	No	No		.	.
DAYS_ICU_Input		Interval	No	No		.	.
Department_Input		Nominal	No	No		.	.
Diagnosis_GInput		Nominal	No	No		.	.
DIAGNOSIS_IInput		Nominal	No	No		.	.
Disch_NurseInput		Interval	No	No		.	.
DISCHARGE_Rejected		Interval	No	No		.	.
DISCHARGEDInput		Nominal	No	No		.	.
DOCTOR_Input		Interval	No	No		.	.
ENCOUNTER_ID		Interval	No	No		.	.
gender_Input		Binary	No	No		.	.
HOSPITAL_Input		Nominal	No	No		.	.
ICU_DAYS_Input		Interval	No	No		.	.
LENGTH_OF_Input		Interval	No	No		.	.
Num_ChroniInput		Nominal	No	No		.	.
NUMBER_CHIRRejected		Unary	No	No		.	.
op_visits6_Input		Interval	No	No		.	.
operationcoInput		Nominal	No	No		.	.
order_set_usInput		Binary	No	No		.	.
order_total_cInput		Interval	No	No		.	.
PATIENT_NURejected		Interval	No	No		.	.
PatientAge_Input		Interval	No	No		.	.
PROCEDURE_Input		Nominal	No	No		.	.
race_cd_Input		Nominal	No	No		.	.
readmit_dateRejected		Interval	No	No		.	.
readmit_day:Rejected		Nominal	No	No		.	.
readmit_discRejected		Interval	No	No		.	.
readmit_nurTarget		Binary	No	No		.	.
REGION_Rejected		Nominal	No	No		.	.
SampleSize_Rejected		Unary	No	No		.	.
SamplingWeiRejected		Unary	No	No		.	.
SelectionProlRejected		Unary	No	No		.	.
Standard_OriInput		Binary	No	No		.	.
STATECODE_Input		Nominal	No	No		.	.
Total_Rejected		Unary	No	No		.	.
X_Rejected		Interval	No	No		.	.
Y_Rejected		Interval	No	No		.	.
ZIP_Rejected		Interval	No	No		.	.

STATSEXPLORER NODE PROPERTIES & RESULTS

General

Node ID	Stat
Imported Data	...
Exported Data	...
Notes	...

Train

Variables	...
-----------	-----

Data

Number of Observations	100000
Validation	No
Test	No

Standard Reports

Interval Distribution	Yes
Class Distributions	Yes
Level Summary	Yes
Use Segment Variables	No
Cross-Tabulation	...

Variable Selection

Hide Rejected Variables	Yes
Number of Selected	1000

Chi-Square Statistics

Chi-Square	Yes
Interval Variables	No
Number of Bins	5

Correlation Statistics

Correlations	Yes
Pearson Correlations	Yes
Spearman Correlation	No

Results - Node: StatExplore Diagram: ReadmitData

File Edit View Window

Variable Worth

Target = readmit_number

Variable

Output

```

1  *
2 User: manavia
3 Date: September 24, 2018
4 Time: 22:22:45
5 *
6 * Training Output
7 *
8
9
10
11
12
13
14
15
16

```

Variable Summary

Role	Measurement	Frequency
	Level	Count

Chi-Sq...

Chi-Square Plot

Target = readmit_number
Data Role = TRAIN

MULTIPLY NODE PROPERTIES & RESULTS

General

Node ID	Plot
Imported Data	...
Exported Data	...
Notes	...

Train

Variables	...
Type of Charts	Bar Charts
Bar Chart Options	
Graph Orientation	Vertical
Include Missing Value	Yes
Interval Target Charts	Mean
Show Values	Yes
Statistic	Percent
Numeric Threshold	20
Scatter Options	
Confidence Interval	Yes
Regression Equation	No
Regression Type	Linear

Results - Node: MultiPlot Diagram: ReadmitData

File Edit View Window

Train Graphs

DAYS_ICU by readmit_number

DAYS_ICU by readmit_number.png

Output

```

1 *-----
2 User: manavia0
3 Date: September 24, 2018
4 Time: 22:23:11
5 *-----
6 * Training Output
7 *-----
8
9
10
11
12 Variable Summary
13
14 Role Measurement Level Frequency
15      Count
16

```

DATA PARTITION NODE PROPERTIES AND RESULTS

Property	Value
General	
Node ID	Part
Imported Data	...
Exported Data	...
Notes	...
Train	
Variables	...
Output Type	Data
Partitioning Method	Default
Random Seed	12345
Data Set Allocations	
- Training	70.0
- Validation	30.0
- Test	0.0
Report	
Interval Targets	Yes
Class Targets	Yes

Results - Node: Data Partition Diagram: ReadmitData

File Edit View Window

Output

```

48
49
50
51 Summary Statistics for Class Targets
52
53 Data=DATA
54
55      Variable    Numeric   Formatted   Frequency
56          Value        Value       Count      Percent   Label
57
58 readmit_number     0           0        116804    80.5767
59 readmit_number     1           1        28156     19.4233
60
61
62 Data=TRAIN
63
64      Variable    Numeric   Formatted   Frequency
65          Value        Value       Count      Percent   Label
66
67 readmit_number     0           0        81762     80.5775
68 readmit_number     1           1        19708     19.4225
69
70
71 Data=VALIDATE
72
73      Variable    Numeric   Formatted   Frequency
74          Value        Value       Count      Percent   Label
75
76 readmit_number     0           0        35042     80.5748
77 readmit_number     1           1        8448      19.4252
78

```

IMPUTATION NODE PROPERTIES AND RESULTS

Property	Value
General	
Node ID	Impt
Imported Data	...
Exported Data	...
Notes	...
Train	
Variables	...
Nonmissing Variables	No
Missing Cutoff	50.0
Class Variables	
Default Input Method	Tree
Default Target Method	None
Normalize Values	Yes
Interval Variables	
Default Input Method	Median
Default Target Method	None
Default Constant Value	
Default Character Value	
Default Number Value	.
Method Options	
Random Seed	12345
Tuning Parameters	...
Tree Imputation	...
Score	
Hide Original Variables	Yes
Indicator Variables	
Type	None
Source	Imputed Variables
Role	Rejected
Report	
Validation and Test Data	No
Distribution of Missing	No

Results - Node: Impute Diagram: ReadmitData

File Edit View Window

Imputation Summary

Variable Name	Impute Method	Imputed Variable	Impute Value	Role	Measurement Level	Label	Number of Missing for TRAIN
Department	TREE	IMP_Depa...	.INPUT	NOMINAL			10975
Num_Chronic_Cond	TREE	IMP_Num...	.INPUT	NOMINAL			129
order_total_charges	MEDIAN	IMP_order...	27839	INPUT	INTERVAL		38

Output

```

33
34
35
36
37      Imputation Summary
38      Number Of Observations
39
40
41      Variable Name      Impute      Imputed Variable      Impute Value      Role      Measurement      Label
42      Method            Method        Variable           Value          Level
43
44      Department        TREE         IMP_Department       .           INPUT      NOMINAL
45      Num_Chronic_Cond  TREE         IMP_Num_Chronic_Cond   .           INPUT      NOMINAL
46      order_total_charges MEDIAN      IMP_order_total_charges 27839      INPUT      INTERVAL
47
48

```

LOGISTIC REGRESSION PROPERTIES

Property	Value
General	
Node ID	Reg
Imported Data	...
Exported Data	...
Notes	...
Train	
Variables	...
Equation	
Main Effects	Yes
Two-Factor Interactions	No
Polynomial Terms	No
Polynomial Degree	2
User Terms	No
Term Editor	...
Class Targets	
Regression Type	Logistic Regression
Link Function	Logit
Model Options	
Suppress Intercept	No
Input Coding	Deviation
Model Selection	
Selection Model	Stepwise
Selection Criterion	Validation Misclassification
Use Selection Defaults	Yes
Selection Options	...
Optimization Options	
Technique	Default
Default Optimization	Yes
Max Iterations	0
Max Function Calls	0
Maximum Time	1 Hour
Convergence Criteria	
Uses Defaults	Yes
Options	...
Output Options	

LOGISTIC REGRESSION FINAL EQUATION

5964	Intercept		1	-19.5169	14.7737	1.75	0.1865	0.000	
5965	DIAGNOSIS_LONG_DESC	ACUTE COMBINED SYSTOLIC AND DIAS	1	6.5869	4.8240	1.86	0.1721	725.500	
5966	DIAGNOSIS_LONG_DESC	ACUTE DIASTOLIC HEART FAILURE	1	6.1003	4.8234	1.60	0.2060	445.978	
5967	DIAGNOSIS_LONG_DESC	ACUTE MYOCARDIAL INFARCTION SUBE	1	-3.4637	30.2388	0.01	0.9088	0.031	
5968	DIAGNOSIS_LONG_DESC	ACUTE MYOCARDIAL INFARCTION UNSP	1	-4.1960	33.8253	0.02	0.9013	0.015	
5969	DIAGNOSIS_LONG_DESC	ACUTE ON CHRONIC COMBINED SYSTOL	1	5.9362	4.8233	1.51	0.2184	378.501	
5970	DIAGNOSIS_LONG_DESC	ACUTE ON CHRONIC DIASTOLIC HEART	1	5.8906	4.8232	1.49	0.2220	361.639	
5971	DIAGNOSIS_LONG_DESC	ACUTE ON CHRONIC SYSTOLIC HEART	1	5.8120	4.8232	1.45	0.2282	334.281	
5972	DIAGNOSIS_LONG_DESC	ACUTE SYSTOLIC HEART FAILURE	1	6.1548	4.8235	1.63	0.2020	470.968	
5973	DIAGNOSIS_LONG_DESC	BACTERIAL PNEUMONIA UNSPECIFIED	1	6.4029	4.8254	1.76	0.1845	603.565	
5974	DIAGNOSIS_LONG_DESC	BRONCHIECTASIS WITH ACUTE EXACER	1	7.3629	4.8243	2.33	0.1270	999.000	
5975	DIAGNOSIS_LONG_DESC	BRONCHIECTASIS WITHOUT ACUTE EXA	1	-3.7048	55.2117	0.00	0.9465	0.025	
5976	DIAGNOSIS_LONG_DESC	BRONCHOPNEUMONIA ORGANISM UNSPEC	1	-3.6790	48.6173	0.01	0.9397	0.025	
5977	DIAGNOSIS_LONG_DESC	CHRONIC AIRWAY OBSTRUCTION NOT E	1	6.8857	4.8264	2.04	0.1537	978.196	
5978	DIAGNOSIS_LONG_DESC	CHRONIC DIASTOLIC HEART FAILURE	1	5.7274	4.8248	1.41	0.2352	307.157	
5979	DIAGNOSIS_LONG_DESC	CHRONIC OBSTRUCTIVE ASTHMA UNSPE	1	-3.1359	68.1002	0.00	0.9633	0.043	
5980	DIAGNOSIS_LONG_DESC	CHRONIC OBSTRUCTIVE ASTHMA WITH	1	6.8937	4.8242	2.04	0.1530	986.086	
5981	DIAGNOSIS_LONG_DESC	CHRONIC SYSTOLIC HEART FAILURE	1	8.5474	4.8242	3.14	0.0764	999.000	
5982	DIAGNOSIS_LONG_DESC	CONGESTIVE HEART FAILURE UNSPECI	1	6.4324	4.8231	1.78	0.1823	621.672	
5983	DIAGNOSIS_LONG_DESC	EMPHYSEMATOUS BLEB	1	17.7112	31.6387	0.31	0.5756	999.000	
5984	DIAGNOSIS_LONG_DESC	FEEDING DIFFICULTIES AND MISMANA	1	-3.3164	45.4300	0.01	0.9418	0.036	
5985	DIAGNOSIS_LONG_DESC	HEART FAILURE UNSPECIFIED	1	-3.9690	43.4717	0.01	0.9273	0.019	
5986	DIAGNOSIS_LONG_DESC	HYPERTENSIVE HEART AND CHRONIC K	1	5.9480	4.8235	1.52	0.2175	382.990	
5987	DIAGNOSIS_LONG_DESC	HYPERTENSIVE HEART DISEASE MALIG	1	-3.3570	38.1381	0.01	0.9299	0.035	
5988	DIAGNOSIS_LONG_DESC	HYPERTENSIVE HEART DISEASE UNSPE	1	6.0028	4.8240	1.55	0.2134	404.571	
5989	DIAGNOSIS_LONG_DESC	HYPOTOTASSEMIA	1	-3.0067	48.5763	0.00	0.9506	0.049	
5990	DIAGNOSIS_LONG_DESC	Hosp 46 DISEASES DUE TO OTHER MY	1	-3.6754	62.3966	0.00	0.9530	0.025	
5991	DIAGNOSIS_LONG_DESC	LEFT HEART FAILURE	1	7.9840	4.8270	2.74	0.0981	999.000	
5992	DIAGNOSIS_LONG_DESC	METHICILLIN SUSCEPTIBLE PNEUMONI	1	5.9922	4.8242	1.54	0.2142	400.290	
5993	DIAGNOSIS_LONG_DESC	OBSTRUCTIVE CHRONIC BRONCHITIS W	1	6.9201	4.8233	2.06	0.1514	999.000	
5994	DIAGNOSIS_LONG_DESC	OTHER AMYLOIDOSIS	1	-104.4	46.0843	5.14	0.0234	0.000	
5995	DIAGNOSIS_LONG_DESC	OTHER CHRONIC BRONCHITIS	1	-3.3540	46.9611	0.01	0.9431	0.035	
5996	DIAGNOSIS_LONG_DESC	OTHER CHRONIC Hosp 46 HEART DISE	1	-6.1515	67.9756	0.01	0.9279	0.002	
5997	DIAGNOSIS_LONG_DESC	OTHER EMPHYSEMA	1	7.0512	4.8247	2.14	0.1439	999.000	
5998	DIAGNOSIS_LONG_DESC	OTHER SEPTICEMIA DUE TO GRAMNEGA	1	-3.7742	51.0398	0.01	0.9411	0.023	
5999	DIAGNOSIS_LONG_DESC	OTHER STAPHYLOCOCCUS PNEUMONIA	1	17.8657	24.2953	0.54	0.4621	999.000	
6000	DIAGNOSIS_LONG_DESC	PNEUMOCOCCAL PNEUMONIA STREPTOCO	1	4.8554	4.8242	1.01	0.3142	128.427	
6001	DIAGNOSIS_LONG_DESC	PNEUMONIA DUE TO ANAEROBES	1	-1.6802	44.1689	0.00	0.9697	0.186	
6002	DIAGNOSIS_LONG_DESC	PNEUMONIA DUE TO HEMOPHILUS INFL	1	-5.7026	55.6960	0.01	0.9184	0.003	
6003	DIAGNOSIS_LONG_DESC	PNEUMONIA DUE TO KLEBSIELLA PNEU	1	-2.2983	50.2247	0.00	0.9635	0.100	

6004	IAGNOSIS_LONG_DESC	PNEUMONIA DUE TO MYCOPLASMA PNEU	1	-4.1183	67.4421	0.00	0.9513	0.016
6005	IAGNOSIS_LONG_DESC	PNEUMONIA DUE TO OTHER GRAMNEGAT	1	-4.6273	48.3374	0.01	0.9237	0.010
6006	IAGNOSIS_LONG_DESC	PNEUMONIA DUE TO OTHER SPECIFIED	1	-1.5702	48.2324	0.00	0.9740	0.208
6007	IAGNOSIS_LONG_DESC	PNEUMONIA DUE TO OTHER STREPTOCO	1	-4.7950	43.9089	0.01	0.9130	0.008
6008	IAGNOSIS_LONG_DESC	PNEUMONIA DUE TO PARAINFLUENZA V	1	-2.4541	33.0397	0.01	0.9408	0.086
6009	IAGNOSIS_LONG_DESC	PNEUMONIA DUE TO PSEUDOMONAS	1	6.4779	4.8256	1.80	0.1795	650.622
6010	IAGNOSIS_LONG_DESC	PNEUMONIA DUE TO RESPIRATORY SYN	1	-3.7882	37.5719	0.01	0.9197	0.023
6011	IAGNOSIS_LONG_DESC	PNEUMONIA DUE TO UNSPECIFIED STR	1	6.2452	4.8272	1.67	0.1958	515.546
6012	IAGNOSIS_LONG_DESC	PNEUMONIA ORGANISM UNSPECIFIED	1	5.9934	4.8232	1.54	0.2140	400.779
6013	IAGNOSIS_LONG_DESC	RHEUMATIC HEART FAILURE CONGESTI	1	3.9171	4.8261	0.66	0.4170	50.255
6014	IAGNOSIS_LONG_DESC	UNSPECIFIED DIASTOLIC HEART FAIL	1	5.8578	4.8241	1.47	0.2246	349.949
6015	IAGNOSIS_LONG_DESC	UNSPECIFIED SYSTOLIC HEART FAILU	1	-3.8387	36.3153	0.01	0.9158	0.022
6016	ISCHARGED_TO	AGNST MEDICAL ADVICE	1	1.3769	0.1021	181.81	<.0001	3.963
6017	ISCHARGED_TO	CHG TO LTAC	1	-0.9808	0.2555	14.74	0.0001	0.375
6018	ISCHARGED_TO	HOME HEALTH AGENCY	1	0.8439	0.0546	239.22	<.0001	2.326
6019	ISCHARGED_TO	HOSPICE (HOME)	1	1.4404	0.0789	333.50	<.0001	4.222
6020	ISCHARGED_TO	HOSPICE - MEDICAL INP	1	3.1789	0.2298	191.40	<.0001	24.020
6021	ISCHARGED_TO	INTERMEDIATE CARE	1	1.4086	0.0782	324.77	<.0001	4.090
6022	ISCHARGED_TO	OTHER ACUTE HOSP	1	-1.1433	0.2223	26.44	<.0001	0.319
6023	ISCHARGED_TO	OTHER DEATH	1	0.7080	0.0790	80.32	<.0001	2.030
6024	ISCHARGED_TO	REHAB HOSPITAL	1	-8.8148	.	.		0.000
6025	ISCHARGED_TO	ROUTINE DSCHG, HOME	1	0.8501	0.0523	264.49	<.0001	2.340
6026	isch_Nurse_ID		1	6.281E-7	9.26E-8	46.01	<.0001	0.0549
6027	OSPITAL	Hosp 1	1	0.0808	0.0584	1.91	0.1668	1.084
6028	OSPITAL	Hosp 10	1	-0.0171	0.0583	0.09	0.7696	0.983
6029	OSPITAL	Hosp 11	1	0.0410	0.0578	0.50	0.4784	1.042
6030	OSPITAL	Hosp 12	1	0.0592	0.0575	1.06	0.3025	1.061
6031	OSPITAL	Hosp 13	1	0.0794	0.0569	1.95	0.1630	1.083
6032	OSPITAL	Hosp 14	1	-0.0135	0.0584	0.05	0.8173	0.987
6033	OSPITAL	Hosp 15	1	0.0728	0.0583	1.56	0.2120	1.076
6034	OSPITAL	Hosp 16	1	0.0507	0.0581	0.76	0.3828	1.052
6035	OSPITAL	Hosp 17	1	0.0919	0.0567	2.63	0.1050	1.096
6036	OSPITAL	Hosp 18	1	0.1702	0.0568	8.98	0.0027	1.186
6037	OSPITAL	Hosp 19	1	0.0933	0.0572	2.66	0.1030	1.098
6038	OSPITAL	Hosp 2	1	0.1794	0.0568	9.97	0.0016	1.196
6039	OSPITAL	Hosp 20	1	0.0283	0.0581	0.24	0.6265	1.029
6040	OSPITAL	Hosp 21	1	-0.0267	0.0581	0.21	0.6461	0.974
6041	OSPITAL	Hosp 22	1	0.0379	0.0573	0.44	0.5086	1.039
6042	OSPITAL	Hosp 23	1	0.1015	0.0577	3.09	0.0788	1.107
6043	OSPITAL	Hosp 24	1	-0.0236	0.0576	0.17	0.6814	0.977
6044	OSPITAL	Hosp 25	1	-0.0466	0.0580	0.65	0.4218	0.954
6045	OSPITAL	Hosp 26	1	-0.0496	0.0581	0.73	0.3933	0.952
6046	OSPITAL	Hosp 27	1	-0.1238	0.0590	4.40	0.0360	0.884
6047	OSPITAL	Hosp 28	1	-0.0323	0.0570	0.32	0.5703	0.968
6048	HOSPITAL	Hosp 29	1	-0.0176	0.0591	0.09	0.7659	0.983
6049	HOSPITAL	Hosp 3	1	0.0978	0.0577	2.88	0.0897	1.103
6050	HOSPITAL	Hosp 30	1	-0.0741	0.0585	1.61	0.2049	0.929
6051	HOSPITAL	Hosp 31	1	-0.1293	0.0594	4.73	0.0296	0.879
6052	HOSPITAL	Hosp 32	1	-0.0964	0.0580	2.76	0.0968	0.908
6053	HOSPITAL	Hosp 33	1	-0.1418	0.0600	5.58	0.0182	0.868
6054	HOSPITAL	Hosp 34	1	-0.1596	0.0595	7.19	0.0073	0.852
6055	HOSPITAL	Hosp 35	1	-0.1069	0.0589	3.29	0.0697	0.899
6056	HOSPITAL	Hosp 36	1	-0.0566	0.0590	0.92	0.3375	0.945
6057	HOSPITAL	Hosp 37	1	-0.0939	0.0593	2.50	0.1135	0.910
6058	HOSPITAL	Hosp 38	1	-0.2425	0.0604	16.10	<.0001	0.785
6059	HOSPITAL	Hosp 39	1	-0.0680	0.0847	0.64	0.4219	0.934
6060	HOSPITAL	Hosp 4	1	0.0992	0.0570	3.03	0.0819	1.104
6061	HOSPITAL	Hosp 5	1	0.1000	0.0582	2.95	0.0859	1.105
6062	HOSPITAL	Hosp 6	1	-0.00022	0.0579	0.00	0.9969	1.000
6063	HOSPITAL	Hosp 7	1	0.0839	0.0586	2.05	0.1518	1.088
6064	HOSPITAL	Hosp 8	1	-0.0533	0.0596	0.80	0.3712	0.948
6065	ICU_DAYS		1	-0.0981	0.0174	31.88	<.0001	-0.2066
6066	IMP_Department	CHILDRENS	1	2.2178	8.9766	0.06	0.8049	9.187
6067	IMP_Department	GENERAL MED	1	2.2089	8.9766	0.06	0.8056	9.106
6068	IMP_Department	GENERAL SURG	1	1.5415	8.9819	0.03	0.8637	4.671
6069	IMP_Department	HEART	1	2.3374	8.9766	0.07	0.7946	10.355
6070	IMP_Department	NEUROSCIENCES	1	1.8546	8.9781	0.04	0.8363	6.389
6071	IMP_Department	ONCOLOGY	1	2.4598	8.9766	0.08	0.7841	11.702
6072	IMP_Department	OTHER TRANSPL	1	-7.4473	44.1169	0.03	0.8659	0.001
6073	IMP_Department	PSYCH	1	-8.2302	57.3683	0.02	0.8859	0.000
6074	IMP_Num_Chronic_Cond	0	1	0.1501	0.0221	46.08	<.0001	1.162
6075	IMP_Num_Chronic_Cond	1	1	0.0844	0.0213	15.77	<.0001	1.088
6076	IMP_Num_Chronic_Cond	2	1	0.0339	0.0260	1.69	0.1935	1.034
6077	IMP_Num_Chronic_Cond	3	1	0.3689	0.0362	103.78	<.0001	1.446
6078	IMP_order_total_charges		1	0.000132	1.347E-6	9569.34	<.0001	0.6153
6079	LENGTH_OF_STAY		1	0.1465	0.0164	80.04	<.0001	0.3304
6080	PROCEDURE_LONG_DESC	??	1	0.4472	0.0422	112.47	<.0001	1.564
6081	PROCEDURE_LONG_DESC	ALCOHOL DETOXIFICATION	1	-9.9562	37.3466	0.07	0.7898	0.000
6082	PROCEDURE_LONG_DESC	AQUAPHERESIS	1	0.5689	0.1141	24.84	<.0001	1.766
6083	PROCEDURE_LONG_DESC	ARTERIAL CATHETERIZATION	1	0.6540	0.1577	17.19	<.0001	1.923
6084	PROCEDURE_LONG_DESC	ARTHROCENTESIS	1	-9.0773	34.0567	0.07	0.7898	0.000
6085	PROCEDURE_LONG_DESC	ARTIFICIAL PACEMAKER RATE CHECK	1	0.9443	0.1357	48.43	<.0001	2.571
6086	PROCEDURE_LONG_DESC	ATRIAL CARDIOVERSION	1	0.8250	0.1660	24.71	<.0001	2.282
6087	PROCEDURE_LONG_DESC	AUTOMATIC IMPLANTABLE CARDIOVERT	1	0.6215	0.0957	42.20	<.0001	1.862
6088	PROCEDURE_LONG_DESC	BIOPSY OF ABDOMINAL WALL OR UMBI	1	13.9280	24.3887	0.33	0.5679	999.000
6089	PROCEDURE_LONG_DESC	BIOPSY OF BONE MARROW	1	2.4424	0.1699	206.57	<.0001	11.501
6090	PROCEDURE_LONG_DESC	BIOPSY OF LYMPHATIC STRUCTURE	1	1.5472	0.2672	33.54	<.0001	4.698
6091	PROCEDURE_LONG_DESC	BIOPSY OF SKIN AND SUBCUTANEOUS	1	-10.3556	51.9133	0.04	0.8419	0.000
6092	PROCEDURE_LONG_DESC	CARDIOVASCULAR AND HEMATOPOIETIC	1	-0.3083	0.2818	1.20	0.2740	0.735

6093	PROCEDURE_LONG_DESC	CLOSED ENDOSCOPIC BIOPSY OF BRON	1	0.5457	0.1063	26.37	<.0001	1.726
6094	PROCEDURE_LONG_DESC	CLOSED ENDOSCOPIC BIOPSY OF LUNG	1	0.7520	0.0862	76.18	<.0001	2.121
6095	PROCEDURE_LONG_DESC	CLOSED ENDOSCOPIC BIOPSY OF RECT	1	101.8	.	.	.	999.000
6096	PROCEDURE_LONG_DESC	CLOSURE OF SKIN AND SUBCUTANEOUS	1	12.3887	23.0306	0.29	0.5906	999.000
6097	PROCEDURE_LONG_DESC	COMBINED RIGHT AND LEFT HEART CA	1	-12.2357	49.3337	0.06	0.8041	0.000
6098	PROCEDURE_LONG_DESC	CONTINUOUS INVASIVE MECHANICAL V	1	-0.0104	0.0826	0.02	0.8997	0.990
6099	PROCEDURE_LONG_DESC	CONTROL OF EPISTAXIS BY ANTERIOR	1	-11.2410	42.4921	0.07	0.7914	0.000
6100	PROCEDURE_LONG_DESC	DIAGNOSTIC ULTRASOUND OF HEART	1	0.2915	0.0425	47.14	<.0001	1.338
6101	PROCEDURE_LONG_DESC	ENTERAL INFUSION OF CONCENTRATED	1	-0.1003	0.2031	0.24	0.6215	0.905
6102	PROCEDURE_LONG_DESC	ESOPHAGOGASTRODUODENOSCOPY EGD W	1	-9.8443	45.5288	0.05	0.8288	0.000
6103	PROCEDURE_LONG_DESC	EXCISION OR DESTRUCTION OF OTHER	1	23.7674	60.6661	0.15	0.6952	999.000
6104	PROCEDURE_LONG_DESC	FLEXIBLE SIGMOIDOSCOPY	1	2.4801	0.2236	123.04	<.0001	11.942
6105	PROCEDURE_LONG_DESC	HEMODIALYSIS	1	0.4037	0.0638	40.01	<.0001	1.497
6106	PROCEDURE_LONG_DESC	INCISION OF GUM OR ALVEOLAR BONE	1	-8.2984	48.3394	0.03	0.8637	0.000
6107	PROCEDURE_LONG_DESC	INCISION WITH REMOVAL OF FOREIGN	1	-10.6986	45.8111	0.05	0.8153	0.000
6108	PROCEDURE_LONG_DESC	INFUSION OF VASOPRESSOR AGENT	1	-10.1887	35.6187	0.08	0.7748	0.000
6109	PROCEDURE_LONG_DESC	INJECTION OF ANTIBIOTIC	1	-0.6348	0.1190	28.45	<.0001	0.530
6110	PROCEDURE_LONG_DESC	INJECTION OF ANTICOAGULANT	1	0.1842	0.0888	4.31	0.0380	1.202
6111	PROCEDURE_LONG_DESC	INJECTION OF INSULIN	1	0.3859	0.0912	17.89	<.0001	1.471
6112	PROCEDURE_LONG_DESC	INJECTION OF OTHER ANTIINFECTIVE	1	0.8769	0.0867	102.20	<.0001	2.403
6113	PROCEDURE_LONG_DESC	INJECTION OF STEROID	1	-0.0662	0.1413	0.22	0.6396	0.936
6114	PROCEDURE_LONG_DESC	INJECTION OR INFUSION GAMMA GLOB	1	2.0506	0.2041	100.95	<.0001	7.773
6115	PROCEDURE_LONG_DESC	INJECTION OR INFUSION OF CANCER	1	22.4593	48.3599	0.22	0.6423	999.000
6116	PROCEDURE_LONG_DESC	INJECTION OR INFUSION OF ELECTRO	1	-8.6468	46.1445	0.04	0.8514	0.000
6117	PROCEDURE_LONG_DESC	INJECTION OR INFUSION OF OTHER T	1	0.9098	0.0567	257.02	<.0001	2.481
6118	PROCEDURE_LONG_DESC	INSERTION OF ENDOTRACHEAL TUBE	1	2.0747	0.1777	136.38	<.0001	7.962
6119	PROCEDURE_LONG_DESC	INSERTION OF INTERCOSTAL CATHETE	1	-10.1700	24.1133	0.18	0.6732	0.000
6120	PROCEDURE_LONG_DESC	INSERTION OF OTHER NASOGASTRIC T	1	-8.8158	45.8931	0.04	0.8477	0.000
6121	PROCEDURE_LONG_DESC	LARYNGOSCOPY AND OTHER TRACHEOSC	1	11.7406	19.1838	0.37	0.5405	999.000
6122	PROCEDURE_LONG_DESC	LEFT HEART CARDIAC CATHETERIZATI	1	0.9583	0.2967	10.43	0.0012	2.607
6123	PROCEDURE_LONG_DESC	OTHER BRONCHOSCOPY	1	-9.8906	38.7914	0.07	0.7987	0.000
6124	PROCEDURE_LONG_DESC	OTHER ENDOVASCULAR REPAIR OF ANE	1	-9.8233	47.6690	0.04	0.8367	0.000
6125	PROCEDURE_LONG_DESC	OTHER EXCISION OF PERIANAL TISSU	1	-9.6763	48.4823	0.04	0.8418	0.000
6126	PROCEDURE_LONG_DESC	OTHER GASTROSTOMY	0	0
6127	PROCEDURE_LONG_DESC	OTHER LAVAGE OF BRONCHUS AND TRA	1	2.6404	0.2628	100.92	<.0001	14.019
6128	PROCEDURE_LONG_DESC	OTHER RADIOTHERAPEUTIC PROCEDURE	1	-12.4261	40.2761	0.10	0.7577	0.000
6129	PROCEDURE_LONG_DESC	PARENTERAL INFUSION OF CONCENTRA	1	3.1117	0.3606	74.47	<.0001	22.460
6130	PROCEDURE_LONG_DESC	PERCUTANEOUS ABDOMINAL DRAINAGE	1	-8.9378	33.3929	0.07	0.7890	0.000
6131	PROCEDURE_LONG_DESC	PERCUTANEOUS ENDOSCOPIC GASTROST	1	-9.4086	63.0002	0.02	0.8813	0.000
6132	PROCEDURE_LONG_DESC	PERICARDIOCENTESIS	1	10.0041	22.8712	0.19	0.6618	999.000
6133	PROCEDURE_LONG_DESC	REPLACEMENT OF TRACHEOSTOMY TUBE	1	-10.7492	48.8855	0.05	0.8260	0.000
6134	PROCEDURE_LONG_DESC	RIGHT HEART CARDIAC CATHETERIZAT	1	2.3575	0.2730	74.56	<.0001	10.565
6135	PROCEDURE_LONG_DESC	SPINAL TAP	1	1.3582	0.1515	80.35	<.0001	3.889
6136	PROCEDURE_LONG_DESC	TOTAL TEMPORARY TRACHEOSTOMY	1	2.8386	0.2561	122.81	<.0001	17.091
6137	ROCEDURE_LONG_DESC	THORACENTESIS	1	-0.0990	0.0947	1.09	0.2960	0.906
6138	ROCEDURE_LONG_DESC	THORACOSCOPIC DECORTICATION OF L	1	-11.4392	96.8400	0.01	0.9060	0.000
6139	ROCEDURE_LONG_DESC	THORACOSCOPIC DRAINAGE OF PLEURA	1	-8.9022	34.2392	0.07	0.7949	0.000
6140	ROCEDURE_LONG_DESC	TRANSFUSION OF PACKED CELLS	1	-9.9185	47.8263	0.04	0.8357	0.000
6141	ROCEDURE_LONG_DESC	TRANSJUGULAR LIVER BIOPSY	1	-9.7735	51.1226	0.04	0.8484	0.000
6142	ROCEDURE_LONG_DESC	VENOUS CATHETERIZATION FOR RENAL	1	0.3494	0.1649	4.49	0.0342	1.418
6143	ROCEDURE_LONG_DESC	VENOUS CATHETERIZATION NOT ELSEW	0	0
6144	TATECODE	AL	1	2.2207	0.0485	2092.75	<.0001	9.214
6145	TATECODE	AR	1	-7.3859	56.9528	0.02	0.8968	0.001
6146	TATECODE	FL	1	2.1158	.	.	.	8.297
6147	TATECODE	GA	1	2.5355	0.0495	2623.29	<.0001	12.622
6148	TATECODE	IL	1	-0.1639	0.2240	0.54	0.4644	0.849
6149	TATECODE	MO	1	3.8101	61.8959	0.00	0.9509	45.153
6150	TATECODE	MS	1	-8.7092	43.3773	0.04	0.8409	0.000
6151	TATECODE	TN	1	5.1037	61.7124	0.01	0.9341	164.631
6152	TATECODE	TX	1	-0.7766	0.2002	15.05	0.0001	0.460
6153	standard_Orders_Used	H	1	0.2931	0.0112	680.53	<.0001	1.341
6154	ender	F	1	-0.0604	0.00966	39.08	<.0001	0.941
6155	p_visits6		1	0.0116	0.00311	13.87	0.0002	0.0189
6156	perationcount	0	1	1.5011	10.6965	0.02	0.8884	4.487
6157	perationcount	1	1	1.6327	10.6965	0.02	0.8787	5.118
6158	perationcount	2	1	1.5437	10.6965	0.02	0.8852	4.682
6159	perationcount	3	1	1.7961	10.6966	0.03	0.8667	6.026
6160	perationcount	4	1	1.9875	10.6972	0.03	0.8526	7.297
6161	ace_cd	Asi	1	0.2018	0.0296	46.36	<.0001	1.224
6162	ace_cd	Black	1	-0.2226	0.0257	75.20	<.0001	0.800

LOGISTIC REGRESSION STEPWISE SUMMARY

Summary of Stepwise Selection							
Step	Entered	Effect	DF	Number In	Score Chi-Square	Wald Chi-Square	Validation Misclassification Rate
1		IMP_order_total_charges	1	1	17535.1847	<.0001	0.1619
2		PROCEDURE_LONG_DESC	64	2	3517.3682	<.0001	0.1554
3		DIAGNOSIS_LONG_DESC	49	3	2704.1232	<.0001	0.1532
4		Standard_Orders_Used	1	4	925.9924	<.0001	0.1548
5		STATECODE	9	5	597.4952	<.0001	0.1530
6		LENGTH_OF_STAY	1	6	510.2683	<.0001	0.1507
7		DISCHARGED_TO	10	7	441.2856	<.0001	0.1523
8		IMP_Num_Chronic_Cond	4	8	171.0664	<.0001	0.1521
9		race_cd	2	9	82.8759	<.0001	0.1515
10		operationcount	5	10	73.0216	<.0001	0.1512
11		IMP_Department	8	11	81.4028	<.0001	0.1525
12		gender	1	12	37.2852	<.0001	0.1528
13		ICU_DAYS	1	13	31.4106	<.0001	0.1517
14		Disch_Nurse_ID	1	14	14.0644	0.0002	0.1519
15		op_visits6	1	15	12.8329	0.0003	0.1510
16		HOSPITAL	38	16	64.6522	0.0045	0.1507

LOGISTIC REGRESSION LIFT CHART

