Analysis of patient satisfaction survey data

by

Jason M. Gaumer

A.A., Johnson County Community College, 2012
B.S., Kansas State University, 2014

A REPORT

submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

Department of Statistics
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2019

Approved by:

Major Professor
Dr. Karen Keating

Abstract

We analyzed a dataset provided by an anonymous hospital in the Midwest, for the purpose of identifying characteristics that affect two response variables of interest: Topbox Overall score and Advocacy. Topbox Overall score is when patients rate the hospital as a 9 or 10 for an overall patient satisfaction score. Advocacy is when patients say “Yes” they would recommend the hospital to a close family member or friend. Since both Topbox Overall score and Advocacy are binary variables, we will use a logistic model for each response.

The dataset contains 434 observations and 21 potential predictors. Most predictors are on an ordinal scale and contain many missing values. Ordinal predictors were converted to a Likert scale and treated as numeric reducing the number of parameters required to fit the logistic models. Missing values were examined to determine the cause of missingness, and most were found to be missing because they were not applicable. These missing values were changed to zero on the Likert scale, which allowed the affected observations to remain in the analysis. In total, 16 observations were removed from the analysis due to missing values leaving 418 observations to be used in the model building process.

We used several different variable selection techniques to generate suitable models for the two distinct response variables: Topbox Overall score and Advocacy. These techniques were needed to identify a parsimonious model. Forward selection and backward elimination were used with a penalized AIC. These are two common techniques for variable selection. Variable selection was also performed using backward elimination via the p-value approach. For this technique the p-value was computed using the chi-squared distribution. Different techniques were used to determine if the results could be replicated. The same models were identified using all three techniques. After the reduced models were identified, two processes were used for

model checking: Cook’s distances and the Hosmer-Lemeshow test. The Cook’s distances identified no influential points or outliers, and the Hosmer-Lemeshow test indicated that the logistic models were appropriate for both response variables.

The variable selection process resulted in three predictors for the Topbox Overall score and two predictors for Advocacy. Using these predictors, a full interaction model was generated for each response. None of the interactions were significant, so the additive models were accepted as the final models. For Topbox Overall score, the three predictors identified were clear communication by nurses, received care within 30 minutes of arriving in the emergency room, and the doctors spent enough time with the patients. For Advocacy two predictors were identified the doctors listened carefully and nurses spent enough time with patients.

In the two models both had predictors that involved the doctors and nurses, but the variables were not exactly the same. Variables related to communication and time spent with the patient were important themes for both models. Timeliness of care had a greater impact on Topbox Overall score than on Advocacy.

Table of Contents

List of Figures

List of Tables

Chapter 1 - Introduction

We base our analysis on a data set from an anonymous hospital in the Midwest that would like to improve their patients' experience in the emergency room. The hospital tracks their patients' experience via a survey. The survey results are important internally, because they provide direct feedback on patients' experiences and the quality of the care provided.

The results that are of interest to the hospital for this analysis are Advocacy for their facility and the Topbox Overall score. Advocacy is when patients say they would recommend the hospital to a close family member or friend. Topbox Overall score is when a patient gives a nine or ten rating, the two highest overall scores. This overall metric is used by the hospital as one method to track patient satisfaction and the quality of care provided.

From the survey and other hospital records, twenty-one predictor variables are available for analysis. They include questions about staff members' communication, timeliness of care, courtesy of staff, and staff's respect towards patients. The hospital would like to understand which of the predictors have the greatest influence on the Topbox Overall score and Advocacy. Our analysis takes two parallel tracks, one for each response variable. Since both responses are binary, we use a logistic model for each response.

Chapter 2 - Methods

2.1 Data Description

An anonymous hospital in the Midwest provided patient satisfaction data for patients who visited the emergency room during the six-month period from January to June 2017. The data was collected via an emailed survey, which was administered by an independent vendor. We received approval from Kansas State University's Institutional Review Board (IRB) to access the data, but due to confidentiality constraints the data cannot be made public.

A total of 434 patients responded to the survey over the six-month time period. The survey contains basic patient information and responses to questions about patients' experience in the emergency room. The data also contain two distinct response variables: Topbox Overall score and Advocacy. Both are binary variables and will be analyzed via logistic models. A complete list of the responses, predictors, and the levels are provided in Appendix A – Data Dictionary.

The purpose of the analysis is to identify which predictors are significant for estimating the probability of each of the two response variables. One difficulty in applying logistic models to this data set is the high proportion of observed "successes" for Advocacy. Ninety-one percent said that they would recommend the hospital to a close family member or friend (Advocacy) compared to sixty-six percent of the patients who responded to the survey rated the hospital with a nine or ten Topbox Overall score. With Topbox there is a more balanced distribution of the responses between having a Topbox Overall score and not having one. The relatively small proportion of observations (9%) that did not Advocate for the hospital makes it more difficult to draw conclusions about the reasons why they did not advocate. Since the data is sparse the decision was made not to partition the data into multiple groups and set up training and test data sets.

We will consider using all the predictors to estimate the probability of the two responses: Topbox Overall score and Advocacy. If we treated all the variables as categorical and included the two-way interactions, the model would contain 1705 parameters. With only 434 observations, it is not possible to fit a logistic model with this number of parameters. If the interactions are excluded, we would have only 60 parameters in an additive model, but we felt the number of observations was insufficient to estimate 60 parameters, given that only 9% of the observations are "failures" with respect to Advocacy. For this reason, we applied a Likert scale to the ordinal variables and treated the Likert score as numeric. This reduced the number of model parameters to 30.

Of the twelve ordinal variables, six of them have four levels and the other six have five levels. Tables 2.1 and 2.2 provide the Likert scales applied to these variables.

Responses in Data Set	**Scores**
Blank	0
No	1
Yes Somewhat	2
Yes Definitely	3

Table 2.1: Values for the four level Likert scale

Responses in Data Set	**Scores**
Blank	0
Never	1
Sometimes	2
Usually	3
Always	4

Table 2.2: Values for the five level Likert scale

The survey was voluntary, and respondents were not required to answer every question. Of the twelve ordinal variables, three of them had many missing values. This was a concern for data analysis, because observations with missing values cannot be used to fit a model. We wanted to include as much of the data as possible for analysis, and so we reviewed each question to see if we could determine why the values were missing.

One of the ordinal predictor variables, told purpose of new medications, had 279 missing values (64.3% of the observations were missing). These missing values have a clear explanation that comes from another question asking if they were given a new medication. The 279 survey respondents that did not answer the first question were also not given a new medication, so it was not applicable to them. In this case a score of zero was assigned for these missing responses, because there was a clear explanation for the missing value. These records were then included in the analysis.

Another predictor variable with many missing values was: given as much information as wanted on test results (17.5% of the observations were missing). This question is not applicable if the patient was not given any tests during their emergency room visit. Of the seventy-six patients that did not respond to this question, seventy-three patients did not receive a test, as recorded in the visit information that was received with the survey results. For these patients, the missing values were changed to zero. The other three patients did not respond to the question and had a test. The records for these three patients were removed since we did not know why the values were missing.

The final predictor variable with many missing values was: doctors and nurses did everything to help with pain (19.6% of the observations were missing). For this question there is not another question or other clinical visit information that provides clarification about the patient's pain experience during their stay. Without that clarifying information from the data, it is difficult to determine why there are missing values. Looking at other research articles regarding pain experience in the emergency room, one study identified that seven out of ten emergency room patients are seeking treatment because of pain [1]. In our dataset 19.6% of the patients did not respond to the question about their experience regarding pain. The decision was made to leave the observations in the dataset because it is plausible that the lack of response was because these people were not experiencing pain. A zero was applied to these 85 missing values.

The rest of the variables had less than 2% of the observations missing and the reasons were not clear why they were missing. These remaining records were removed from the data set. In the entire cleaning process, a total of 16 records were removed leaving 418 records to be used in the analysis. Appendix B – Missing Values summarizes the missing values in the raw data listing the possible predictors, the number of observed values, number of missing values, and the number of observations after modifying the data.

2.2 Building the Logistic Models

2.2.1 Variable Selection and Model Building

Since both responses are binary in this project, logistic models will be used. Within the data set there are 21 predictor variables. With this number of variables, it would be challenging to calculate expected probabilities. A subset of these variables needs to be identified to develop the final models. Ideally, each final model would contain two or three variables. To achieve this goal, we employed three different variable selection techniques for each response variable.

Variable selection was performed using the statistical program R [2]. Two functions in R were utilized: step and fastbw. The step function was used to perform forward selection and backward elimination. The fastbw function was used to perform backward elimination using p-values.

The first two methods employ a penalized Akaike's Information Criterion (AIC). Forward selection (AIC) and backward elimination (AIC) are performed via the step function in R. The value for K is the penalty incorporated into the AIC score. Higher values for K produce more parsimonious models. The default value for K is 2, which is a true AIC. As part of our analysis, multiple values of K are trialed during variable selection until there is one variable left in the model. The third method of variable selection, backward elimination (LRT), is implemented via the R function fastbw. It utilizes the chi-squared likelihood ratio test (LRT) to calculate the p-value associated with removing a variable from the model. Five specific thresholds (0.1, 0.05, 0.01, 0.005, 0.001) for the p-values are trialed to see if the same variables can be identified using all three methods.

Our goal is to have a consistent K level for Topbox Overall score and Advocacy. Once the K is trialed for each response until there is only a single variable selected, then the results for the two response variables are compared and one K is selected across the two models.

After the variables are identified to include in the additive models, we test to see if an interaction model is needed. A likelihood ratio test is used to compare the interaction and additive models. If the interaction model is significant then it will be selected; however, if the interaction model is not significant the additive model is selected.

2.2.2 Model Evaluation: Hosmer-Lemeshow

To make sure the model is appropriate; the Hosmer-Lemeshow process is used. It is a process described by Michael H. Kutner et al. [3]. The estimated probabilities of success are calculated for every observation. Next the observations are ordered from smallest to largest using the estimated probabilities, and then the ordered values are divided into five to ten groups that are of approximately the same size. The expected and observed values are calculated by summing up the number of successes for the observed values and adding up the probabilities of success for the expected values. Using the observed and expected values, a Pearson chi-squared test statistic is used to determine if the logistic response function is appropriate. The degrees of freedom for the chi-squared test statistic are the number of groups minus two. The 5% level of significance is used as the threshold.

2.2.3 Model Evaluation: Cook's Distances

After verifying that the logistic model is appropriate, we use the Cook's distances to check for influential points and outliers for the logistic model. The Cook's distances are compared to the F-distribution with the numerator degrees of freedom equal to number of parameters in the model and the denominator degrees of freedom is the difference between the sample size and the number of parameters in the model. If the percentile is less than 10%, this indicates that the observation has little effect on the parameter estimates. This process is described by Michael H. Kutner et al [3]. If there is an influential point, we can remove that point and refit the model. If the parameter estimates change dramatically for the refit model, we will consider removing that observation.

2.2.4 Model Interpretation

After checking for outliers and influential points, we can interpret the logistic models. First, we will estimate the probabilities for each level of the predictor variables. The estimated probabilities will be placed in tables and displayed visually through plots. From the plots and tables, response combinations that result in high or low probabilities can be identified. The shapes of the curves in the plots help us identify dramatic increases (or decreases) in the probabilities of the responses. This process helps identify the combination that results in the highest probability of the responses for Topbox Overall score and Advocacy. We will calculate the change in the probability of the response variable with a one level increase in each predictor variable while keeping the other variables constant. This will show us the impact each change has on the response. We will focus on the largest probability changes for Topbox and Advocacy.

Chapter 3 - Results

3.1 Logistic Model with Topbox Overall Score as the Response

3.1.1 Model Selection Results

Three methods of variable selection were used: forward selection (AIC), backward elimination (AIC), and backward elimination (LRT). First, forward selection (AIC) was performed and various levels of the penalty (K) were trialed. With the lowest value of K (K=2), twelve variables remained in the model. As K was increased the number of variables decreased. When K was between 10 and 20, three variables were consistent in the model. They were “clear communication by nurses” (“nurses communication”), “received care within 30 minutes of arriving in ER” (“care within 30 minutes”), and “doctors spent enough time with patient” (“doctors time”). When K=40 only one variable remained in the model “doctors time.” The results of the variables selected when K was varied with forward selection using step are provided below in Table 3.1. All 21 predictors were considered, but only the ones that were selected are included in the table.

	Values of K Trialed							
	2	3	4	5	6-9	10-20	25, 30, 35	40
doctors time	X	X	X	X	X	X	X	X
care within 30 minutes	X	X	X	X	X	X	X	
nurses communication	X	X	X	X	X	X		
charge of care	X	X	X	X	X			
nurses time	X	X	X	X				
doctors listen	X	X	X					
Age	X	X	X					
Pain	X	X						
Class	X	X						
Test	X	X						
information test	X	X						
time before asked	X							

Table 3.1: Forward selection (AIC) for Topbox Overall score

For backward elimination using a penalized AIC, we trialed values of K ranging from 2 to 40. When K is between 14 and 20, the same three variables remained. They were “nurses communication,” “care within 30 minutes,” and “doctors time.” When K=40 only one variable, “doctors time,” remained in the model. The results of the backward elimination using penalized AIC are provided in Table 3.2.

	Values of K Trialed								
	2	3	4	5	6-9	10-13	14-20	25, 30	40
doctors time	X	X	X	X	X	X	X	X	X
care within 30 minutes	X	X	X	X	X	X	X	X	
nurses communication	X	X	X	X	X	X	X		
information test	X	X	X	X	X	X			
test	X	X	X	X	X	X			
nurses time	X	X	X	X	X				
purpose new meds	X	X		X					
meds	X	X		X					
doctors listen	X	X	X						
age	X	X	X						
charge of care	X								
time before asked	X								
pain	X								
class	X								

Table 3.2: Backward elimination (AIC) for Topbox Overall score

For backward elimination (LRT) we trialed values of alpha ranging from 0.1 to 0.001. When alpha was 0.1, there were nine variables in the model. With an alpha of 0.001 there were only two variables remaining in the model. The results of backward elimination (LRT) are provided in Table 3.3.

	Values of Alpha Trialed				
	0.1	0.05	0.01	0.005	0.001
doctors time	X	X	X	X	X
care within 30 minutes	X	X	X	X	X
nurses communication	X	X	X		
information test	X	X			
Test	X	X			
nurses time	X	X			
purpose new meds	X				
Meds	X				
Age	X				

Table 3.3: Backward elimination (LRT) for Topbox Overall score

When looking across the three methods, the same three variables were selected in each method. They are: "nurses communication," "care within 30 minutes," and "doctors time." Consistent variable selection using all three methods provides a sense of confidence in the final model. Using these three variables creates the largest model where all the methods agree on the variables. Additionally, both forward selection (AIC) and backward elimination (AIC) use approximately the same value for K. Different variables were selected for a fourth predictor depending on the technique used. This further confirms that these three variables were the "best" for the model. The model was finalized with the three common "best" variables.

Next, we need to consider an interaction model to determine if it is a better fit than the additive model that has been developed. A model was fit with all interactions, both the three-way interaction and all two-way interactions. A likelihood ratio test was performed using the chi-squared distribution. The chi-squared test statistic was 4.82, with 4 degrees of freedom. The p-value is 0.3063, indicating that the additive model adequately fits the data and interaction terms were not required for this model.

Based on this analysis, the logistic model for Topbox Overall score is an additive model with three predictors: nurses communication, care within 30 minutes, and doctors time. Estimated coefficients and the standard errors for this model are provided in Table 3.4.

Coefficients	**Estimates**	**Standard Errors (SE)**
Intercept	-13.28	1.597
Nurses Communication	1.558	0.356
Care Within 30 minutes	2.244	0.402
Doctors Time	2.448	0.303

Table 3.4: Estimated coefficients and standard errors: Topbox model

3.1.2 Model Evaluation: Hosmer-Lemeshow

To make sure the logistic model is appropriate, the Hosmer-Lemeshow process was used. This process is described in the methods section (2.2.2). Table 3.5 provides the intervals the probabilities are divided into, the number of observations in each interval, and the observed and expected values for each interval.

Interval for π	n_i	**Observed**	**Expected**
(0.00, 0.04]	47	0	0.357
[0.04, 0.25)	52	8	7.327
[0.25, 0.55)	40	19	20.951
[0.55, 0.73)	39	26	24.857
[0.73, 1.00)	240	223	222.508

Table 3.5: Hosmer-Lemeshow summary table for Topbox Overall score

The group with the interval from 0.73 to 1.00 had 240 observations, which is larger than any of the other groups. This is because all the observations in this group had the same estimated probability of 0.927, so it was impossible to split them into separate groups. The calculated chi-squared test statistic for this dataset is 0.654. We compared that to the chi-squared distribution with three degrees of freedom. The critical value at the 5% level of significance is

7.81 and the p-value for this test statistic is 0.884. This indicates that the logistic response function is appropriate and supports use of this logistic model.

3.1.3 Model Evaluation: Cook's Distances

After the logistic model was validated as appropriate, we next checked for influential points using Cook's distances. There are four parameters in the logistic model. We compared the largest Cook's distance, 0.1061, to the F-distribution (4, 418-4). This resulted in a percentile of 0.0197. Since this value is less than 0.1, the most extreme observation has little apparent influence on the model fit and this indicates that none of the observations have any apparent influence. No influential points were identified to be removed from the model.

3.1.4 Interpretation of Logistic Model for Topbox Score

After model selection for the Topbox Overall score, three predictor variables are found to be significant. These predictors are summarized below.

nurses communication

> During this emergency room visit, how often did nurses explain things in a way you could understand?
>
> Never / Sometimes / Usually / Always

doctors time

> During this emergency room visit, did doctors spend enough time with you?
>
> No / Yes somewhat / Yes definitely

care within 30 minutes

> During this emergency room visit did you get care within 30 minutes of getting to the emergency room?
>
> No / Yes

The estimated probabilities for Topbox Overall score are calculated using the reduced logistic model. Separate plots and tables were created for care received within 30 minutes and care not received within 30 minutes. These plots can be seen in Figures 3.1 and 3.2 respectively. The data used to create the plots is displayed in Tables 3.6 and 3.7 respectively. The tables provide the estimated probabilities of Topbox Overall score for the various levels of time doctors spent with patients and nurses communication.

Doctors time is very important for the Topbox Overall score. The highest probability of a Topbox Overall score is 0.927 when doctors time is rated at the highest level. When it drops down one level to “somewhat,” the highest probability drops to 0.523. When the doctor does not spend enough time with the patient, the largest probability of Topbox Overall score is only 0.087.

Nurses communication is also very important for the Topbox Overall score. The highest probability of 0.927 occurs when nurses communication is at the highest level. If the nurse never explained things in a way the patient could understand, the largest probability of Topbox Overall score is only 0.106. This increases to 0.727 if the nurse “usually” explained things well. It is important to note, however, that even when nurse’s communication is at the highest level when doctors time is “no,” the probability remains very low at or below 0.087.

All three factors are important when estimating Topbox probabilities. It is not surprising that the highest probability (0.927) occurs when all three factors are at their highest levels and the lowest probability (0.0001) occurs when all three factors are at their lowest levels. Since there are more observations at the higher levels, we focus our interpretation on those levels. When all three factors are at their highest levels, the estimated probability of Topbox is 0.927. If the nurses communication drops one level (from “always” to “usually”), the probability changes from 0.927 to 0.727. If the doctors time is lowered one level (from “yes definitely” to “yes somewhat”), the maximum probability is reduced from 0.927 to 0.523. If the patient does not receive care within 30 minutes, the maximum probability is reduced from 0.927 to 0.574. While the smallest of these decreases occur for nurses communication, this does not indicate that this factor is less important than the others. This is due, in part, to the fact that nurses communication has four levels, while doctors time has only three levels and care within 30 minutes has only two levels.

If the doctor does not spend enough time with the patient, the highest probability is only 0.087. If the nurses never communicate with the patient, the highest probability is 0.106. If care is not received within 30 minutes, the highest probability is 0.574. From these figures, it would seem that a positive experience with the medical staff is more important than the timeliness of care.

When considering one level changes for both the highest and the second highest probability of Topbox Overall score, doctors time and care received within 30 minutes have the largest impact. The largest one level increase in the probability of Topbox Overall score is with a change in “doctors time.” The change is when care is within 30 minutes, nurses “usually” communicate and doctors time changes from “yes somewhat” to “yes definitely.” There is a 0.540 absolute increase in probability of a Topbox Overall score. Doctors time appears to be the most influential factor for increasing the probability of a Topbox Overall score.

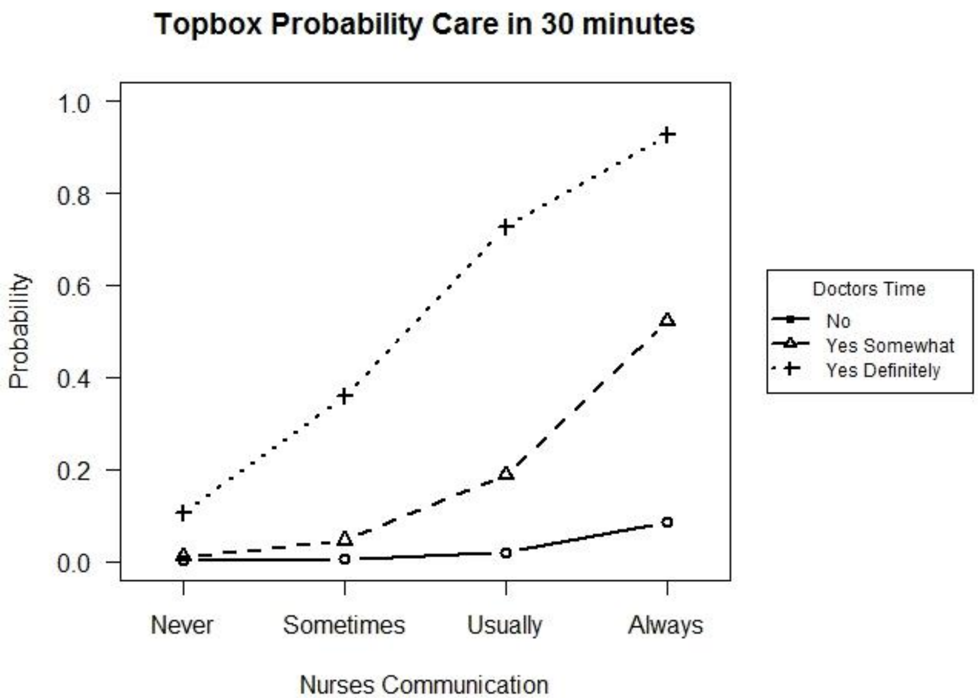

Figure 3.1: Plot of Topbox Overall score: "care within 30 minutes" is "Yes"

<table>
<tr><td></td><td></td><th colspan="3">Doctors Time</th></tr>
<tr><td></td><td></td><th>No</th><th>Yes Somewhat</th><th>Yes Definitely</th></tr>
<tr><th rowspan="4">Nurses Communication</th><th>Never</th><td>0.0009
N=2</td><td>0.0101
N=0</td><td>0.1058
N=1</td></tr>
<tr><th>Sometimes</th><td>0.0042
N=8</td><td>0.0463
N=5</td><td>0.3597
N=0</td></tr>
<tr><th>Usually</th><td>0.0195
N=3</td><td>0.1874
N=20</td><td>0.7273
N=16</td></tr>
<tr><th>Always</th><td>0.0865
N=8</td><td>0.5226
N=40</td><td>0.9268
N=240</td></tr>
</table>

Table 3.6: Probabilities of Topbox Overall score: "care within 30 minutes" is "Yes"

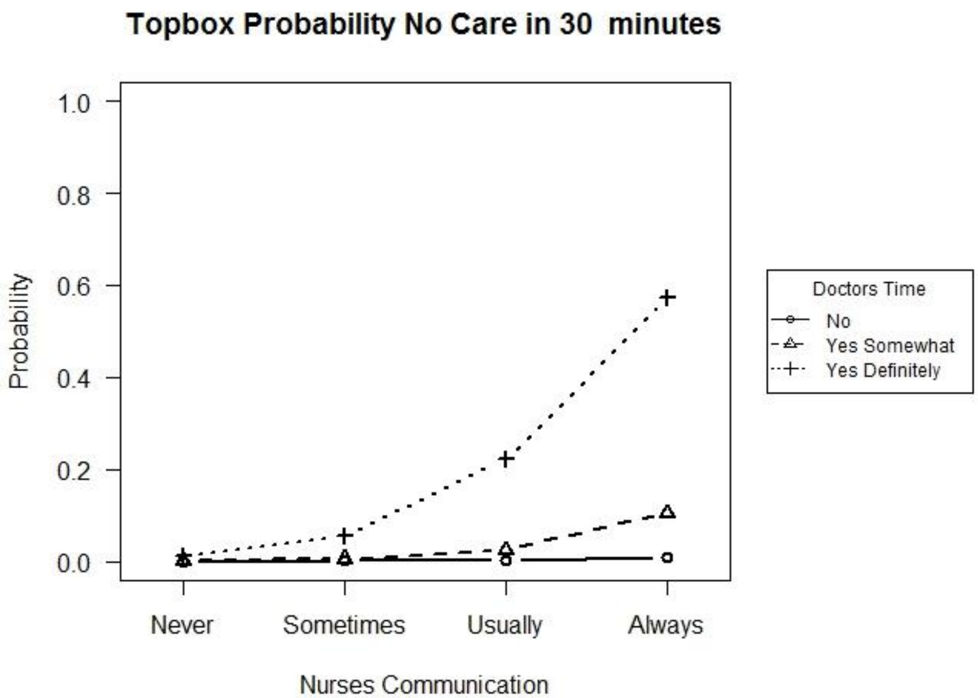

Figure 3.2: Plot of Topbox Overall score: “care within 30 minutes” is “No”

Doctors Time

		No	**Yes Somewhat**	**Yes Definitely**
Nurses Communication	**Never**	0.0001 N=5	0.0011 N=0	0.0124 N=0
	Sometimes	0.0004 N=5	0.0051 N=1	0.0564 N=1
	Usually	0.0021 N=11	0.0240 N=8	0.2212 N=6
	Always	0.0100 N=4	0.1044 N=11	0.5741 N=23

Table 3.7: Probabilities of Topbox Overall score: “care within 30 minutes” is “No”

3.2 Logistic Model with Advocacy as the Response

3.2.1 Model Selection Results

The first step in model selection for Advocacy was variable selection. Three methods of variable selection were used including forward selection (AIC), backward elimination (AIC), and backward elimination (LRT). For forward selection (AIC), various levels of the penalty K were trialed. With the lowest value of K (K=2), eight variables remained in the model. As K was increased, the number of variables decreased. When K was between 7 and 16, two variables were consistent in the model. They were "doctors listened carefully" ("doctors listen") and "nurses spent enough time with patient" ("nurses time"). When K=17 the only variable that remained in the model was "doctors listen." The results of the variables selected when K was varied with forward selection (AIC) are provided in Table 3.8. All 21 predictors were considered, but only the ones that were selected are included in the table.

	Values of K Trialed				
	2	3	4-6	7-16	17
doctors listen	X	X	X	X	X
nurses time	X	X	X	X	
Pain	X	X	X		
nurses communication	X	X			
Class	X	X			
time before asked	X				
Meds	X				
purpose new meds	X				

Table 3.8: Forward selection (AIC) for Advocacy

For backward elimination using a penalized AIC, we trialed values of K ranging from 2 to 17. When K is between 7 and 16, the same two variables remained. They were "doctors listen" and "nurses time." At K=17 only one variable remained, so no further levels of K were trialed. The results of the backward elimination using penalized AIC are provided in Table 3.9.

All 21 predictors were considered, but only the ones that were selected are included in the table. The same variables were selected with forward selection when K was set at the same level between 7 and 16.

	Values of K Trialed				
	2	3	4-6	7-16	17
doctors listen	X	X	X	X	X
nurses time	X	X	X	X	
Pain	X	X	X		
courtesy nurses	X	X			
nurses communication	X	X			
Class	X	X			
doctors time	X	X			
coverage area	X				

Table 3.9: Backward elimination (AIC) for Advocacy

For backward elimination using the likelihood ratio tests (LRT), we trialed values of alpha ranging from 0.1 to 0.001. The results of the backward elimination (LRT) are provided in Table 3.10. All 21 predictors were considered, but only the ones that were selected are included in the table. In order to have at least two variables for consideration, the final variables selected from backward elimination (LRT) were "doctors listen" and "nurses time."

	Values of Alpha Trialed				
	0.1	0.05	0.01	0.005	0.001
doctors listen	X	X	X	X	X
nurses time	X	X			

Table 3.10: Backward elimination (LRT) for Advocacy

When looking across the three methods of variable selection, two variables were selected in each method. They are "doctors listen" and "nurses time." Consistent variable selection using all three methods gives a greater confidence in the final model. Additionally, all three methods identified these same "best" variables when K=15 for forward selection (AIC) and backward elimination (AIC). Because backward elimination (LRT) only identified these two variables, no other variables were considered for the final model. The model was finalized with these two "best" variables identified in all three methods.

Next, we need to consider an interaction model to determine if it is a better fit than the additive model that has been developed. A model was fit with the two-way interaction. A likelihood ratio (LR) test was performed using the chi-squared distribution. The chi-squared test statistic was 1.81 with degrees of freedom equal to 1. The p-value is 0.179 indicating that the additive model adequately fits the data and interaction terms were not required for this model.

Based on this analysis, the logistic model for Advocacy is an additive model with two predictors: doctors listen and nurses time. Estimated coefficients and the standard errors for this model are provided in Table 3.11.

Coefficients	Estimates	Standard Errors (SE)
Intercept	4.684	0.777
Doctors Listen	1.306	0.215
Nurses Time	1.210	0.304

Table 3.11: Estimated coefficients and standard errors for Advocacy

3.2.2 Model Evaluation: Hosmer-Lemeshow

The Hosmer-Lemeshow process was used to see if the logistic model was appropriate. This process is described in the methods section (2.2.2). Table 3.12 provides a summary of the intervals the probabilities are divided into, the number of observations in each interval, and the observed and expected values for each interval.

Interval for π	n_i	Observed	Expected
(0.00, 0.83)	47	21	19.786
[0.83, 0.95)	51	43	45.500
[0.95, 0.96)	41	38	38.974
[0.96, 1.00)	279	277	274.739

Table 3.12: Hosmer-Lemeshow summary table for Advocacy

The group with the interval from 0.96 to 1.00 had 279 observations, which is larger than any of the other groups. This is because all the observations in this group had the same estimated probability of 0.985, so it was impossible to split them into separate groups. The calculated chi-squared test statistic for this dataset is 0.255. We compared that to the chi-squared distribution with two degrees of freedom. The critical value at the 5% level of significance is 5.99 and the p-value for this test statistic is 0.880. This indicates that the logistic response function is appropriate and supports use of this logistic model.

3.2.3 Model Evaluation: Cook's Distances

After the logistic model was validated as appropriate, we next checked for influential points using Cook's distances. There are three parameters in the logistic model. We compared the largest Cook's distance, 0.075, to the F-distribution (3, 418-3). This resulted in a percentile of 0.0265. Since this value is less than 0.1, the most extreme observation has little apparent influence on the model fit and this indicates that none of the observations have any apparent influence. No influential points were identified to be removed from the model.

3.2.4 Interpretation of Logistic Model for Advocacy

After model selection for Advocacy, two predictor variables are found to be significant. These are summarized below.

doctors listen

During this emergency room visit, how often did doctors listen carefully to you?

Never / Sometimes / Usually / Always

nurses time

During this emergency room visit, did nurses spend enough time with you?

No / Yes somewhat / Yes definitely

The estimated probabilities for Advocacy are calculated using the additive logistic model. The probabilities are shown in Table 3.13 and are graphed in Figure 3.3. The lines representing the different levels of doctors listen do not cross since we are fitting an additive model that does not allow the lines to cross.

The highest probability of advocacy (0.985) occurs when the nurse “definitely” spent enough time with the patient and the doctor “always” listened carefully. In fact, the probability of advocacy remains at 0.95 or higher if either factor drops by one level. When there is a reduction of two levels for either nurses time or doctors listen, the probability of advocacy is in the range of 0.826 to 0.852.

When doctors listen is “always,” the probability remains above 0.852 regardless of how much time nurses spent with the patient. An increase in nurses time when doctors listen is rated at the highest level has a small impact. The largest possible increase in the probability of advocacy for a one level nursing change is only 0.099. This increase occurs when nurses time increases from “no” to “yes somewhat.” With a two-level change when nurses time changes from “no” to “yes definitely,” the increase in probability is only 0.133. The small increase is not surprising because when the doctors are “always’ listening the probability is already in the higher range.

Nurses time has a large impact on the change in probability when doctors listen is rated at a lower level. For example, when nurses time changes from “no” to “yes definitely,” the increase in probability is 0.529 when doctors “sometimes” listen. There is an increase of 0.289 when doctors listen “sometimes,” and nurses time increased from “no” to “yes somewhat.” There is a similar increase of 0.285 when doctors do not listen and nurses time increased from “yes somewhat” to “yes definitely.”

The largest one level change in the probability of Advocacy is with a change in doctors listening. The increase in probability is 0.312 occurring when nurses time is “no” and doctors listen goes from “sometimes” to “usually.” This indicates that doctors listening is an important factor for increasing the probability of Advocacy, especially when there is not enough time with the nurses. The largest one level change for nurses time is 0.289 which occurs when doctors listen is “sometimes,” and nurses time goes from “no” to “yes somewhat.” This is close to the largest one level change for doctors listening (0.312) and indicates that nurses time is also an important factor, especially when doctors are not listening.

We also see the potential for doctors listening and nurses time to balance each other. If the rating for either doctors listen or nurses time drops by one level and the other one goes up one level then the probability will remain about the same. We see this when looking at the estimated coefficients for the logistic model (Table 3.11) where doctors listen and nurses time have similar coefficient estimates (1.31 and 1.21, respectively). This supports that a one-level change in either doctors listening or nurses time will have a similar impact to the probability of Advocacy. Since doctors listen has a slightly larger estimate, it will have a slightly greater impact.

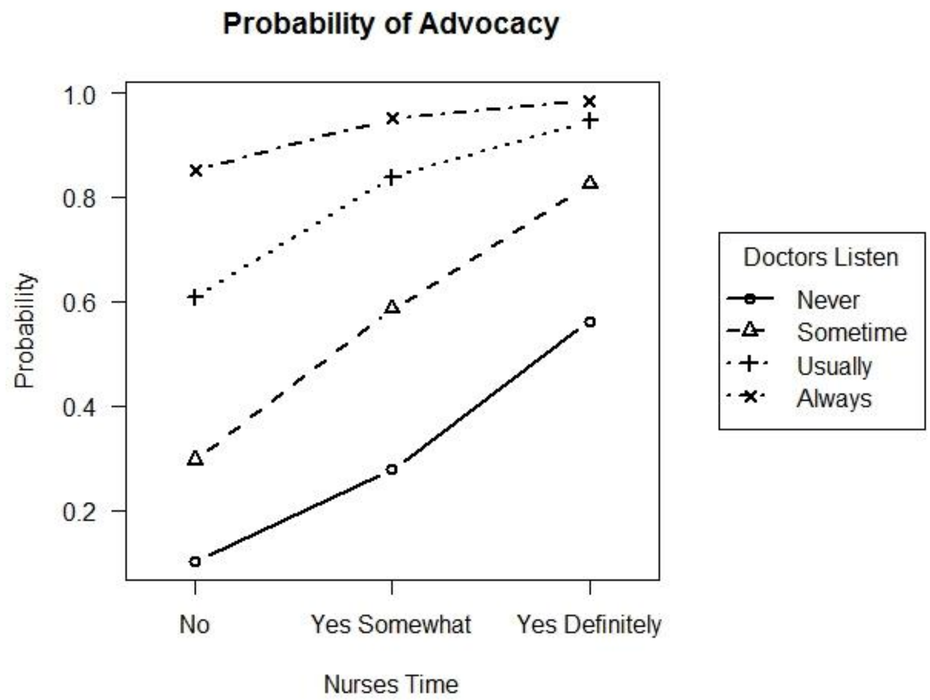

Figure 3.3: Plot of Advocacy Probabilities

		Doctors Listen			
		Never	**Sometimes**	**Usually**	**Always**
Nurses Time	**No**	0.1025 N=7	0.2966 N=10	0.6087 N=7	0.8516 N=3
	Yes Somewhat	0.2769 N=8	0.5856 N=6	0.8391 N=23	0.9506 N=41
	Yes Definitely	0.5621 N=5	0.8257 N=4	0.9459 N=25	0.9847 N=279

Table 3.13: Probabilities of Advocacy

3.3 Compare Topbox Overall Score and Advocacy Models

3.3.1 Compare the Variables in each Model

Since both Topbox Overall score and Advocacy are measuring a general form of patient satisfaction, the final step in our analysis is to compare the two models. First, let us consider the predictor variables that are in the two logistic models. Table 3.14 shows the predictors in the logistic models for Topbox Overall score and Advocacy, along with the specific questions that the patients were asked.

Topbox Overall Score	Advocacy
Doctor time *During the emergency room visit, did doctors spend enough time with you?*	**Nurses time** *During the emergency room visit, did nurses spend enough time with you?*
Nurses communication *During the emergency room visit, how often did nurse explain things in a way you could understand?*	**Doctor listen** *During the emergency room visit, how often did doctor listen carefully to you?*
Care within 30 minutes *During this emergency room visit did you get care within 30 minutes of getting to the emergency room?*	

Table 3.14: Predictor variables for Topbox Overall score and Advocacy

From the analyses performed, the variables for Topbox Overall Score and Advocacy are different. This indicates that there is no one strong variable that is important to both analyses. Since a question about doctors and a question about nurses is in both analyses, this indicates that both doctors and nurses are important to patients' experiences and likelihood to recommend the facility. Nurses communication, which is included in the Topbox Overall score, relates to providing explanations that the patient or family understand. This would include explanations regarding tests and discharge instructions. For Advocacy, nurses time was more important than nurses communication. Nurses time could include communication with the patient but could also include the frequency with which the nurse checked on the patient and family, perhaps to counsel or reassure them.

The two models also incorporated slightly different predictors related to the doctor's activities. For the Topbox Overall Score, the time the doctor spent with the patient was deemed important, but for Advocacy the degree to which the doctor listened to the patient was deemed important. As with the nurses, these variables are closely related but subtly different, since doctors can spend time with a patient explaining or performing procedures (and not explicitly listening).

3.3.2 K and Alpha Differences

When choosing the final model via one of the two penalized AIC methods, a constant penalty K was desired for both forward selection and backward elimination. It was also desired to have parsimonious models with only two or three variables. A K of 15 in both forward selection (AIC) and backward elimination (AIC) would achieve these goals with final models containing three or less variables in each model.

For the model selection process that used likelihood ratio tests, different levels of alpha were used for the final models for Topbox Overall score and Advocacy. Five levels of alpha were trialed for each response. In backward elimination (LRT) an alpha of 0.1 or 0.05 for Advocacy and an alpha of 0.01 for Topbox Overall score selected the same variables in each model as was found with a K of 15. The alphas may be different between the two analyses because there are a different number of parameters in the models. Alpha can be affected by the number of parameters.

3.3.3 Pearson Chi-squared test for Topbox Overall score and Advocacy

Since both Topbox Overall score and Advocacy are measuring the level with which the patient is satisfied with the hospital, we next explored the relationship between these two measures using an ordinary Pearson chi-squared test for association. The contingency table, shown in Table 3.15, produced the test statistic 77.2 with a corresponding p-value less than 2.2×10^{-16}. Such a small p-value indicates there is a strong association between Topbox Overall score and Advocacy.

		Advocacy	
		Yes	**No**
Topbox Overall Score	**Yes**	275	1
	No	104	38

Table 3.15: Contingency table for Topbox and Advocacy

Chapter 4 - Discussion

This study analyzed survey responses for 418 patients experiences during their emergency room visit. There were twenty-one predictor variables, which included demographic information, physician care, nursing care, and patient experience. As part of the analysis, we wanted to identify which variables had the greatest influence on patient satisfaction (Top Box overall score) and willingness to recommend the facility (Advocacy).

4.1 Why are we using variable selection?

The reason we decided to use variable selection instead of alternative available techniques was the sample size and the distribution of the results. For example, one data analysis technique we considered using was Lasso. With this technique the data is randomly divided into three equal size groups. This works well when the data is uniformly distributed across groups. That was not the case for this data. For example, with Advocacy only 9% of the patients said they would not recommend the hospital ("no" for Advocacy). Due to the sparse data, the response might not be represented well in the three groups created. Variable selection allows all the data to be used to fit the model. Another factor supporting the use of variable selection was the goal to create parsimonious models. We wanted only two or three parameters in the model so the model was understandable and could provide more useful feedback for the hospital to focus improvement efforts. Using variable selection, we are able to create models of the desired size.

4.2 Limitations

Based on the data collected there are some limitations on the research. The data is collected from an emailed survey, and patients that did not respond to the survey are not represented in the results.

Another limitation is the sample size. The time frame for the data collection could not be started earlier, because the survey questions were modified starting in January 2017. For the agreed upon reporting period, January to June 2017, there were 418 observations available for analysis after data cleaning. This limits what techniques to use for selecting variables for the model and the types of models to consider.

A final limitation for this study is its repeatability. The hospital reviews its survey processes at least annually and changes the survey questions. This limits the ability to repeat the survey and compare the results over time. If we were to do this analysis again, we would need to see what questions and responses are comparable over time.

4.3 Discussion on Future Work

The approach described in this paper is one way to analyze the data. We could try different scorings systems for the Likert variables and see how that impacts the results. We could compare Likert scoring systems to treating the variables as nominal categories. We could consider different methods of managing the missing values. Another area that could be investigated is considering two-way and three-way interactions among the predictors during model selection and seeing how that impacts the analysis. We could also look at different types of models. Some other models to consider are machine learning, multinomial models, and cumulative logistic models. For example, decision trees in machine learning use the information learned about that individual from all his/her responses to see how they would rate the hospital. The analysis could be repeated with a larger sample size, which would allow use of other techniques, such as Lasso. Additionally, this analysis could be repeated with the same methods and see if there are changes over time. Data could be obtained from other surveying methods like a questionnaire at time of discharge for all emergency room patients. This could be helpful to evaluate impact of improvements being trialed. It would be valuable to research why these selected variables are so important to patients and what impacts patient perceptions of these key variables.

Chapter 5 - Conclusions

Two logistic models were fit to two response variables: patient satisfaction (Topbox Overall score) and willingness to recommend the facility (Advocacy). After cleaning the data, three methods of model selection were performed: backward elimination (AIC), forward selection (AIC), and backward elimination (LRT). The resulting reduced models were evaluated for validity. Interaction terms were considered, but it was judged that they did not significantly improve the models, so the additive models were used as the final models. These models were checked with Cook's distances and the Hosmer-Lemeshow test, and no issues were found.

The hospital providing this data requested analysis to better understand what variables positively impact patients' satisfaction and willingness to recommend the hospital. Understanding these variables can help direct the hospital to identify areas to focus on to improve their patients' experience. Improvements in the patients' experience should result in improved ratings on these two satisfaction responses.

For Topbox Overall score, the final model included variables related to the amount of time doctors spend with patients, nurses' communication with patients, and whether or not care is received within 30 minutes. The Advocacy model includes two variables: doctors listening carefully and nurses spending enough time with patients. The variables are different in the two models, but have common themes related to time and communication. For Advocacy, both doctors and nurses had a similar impact on the response and can provide a balancing effect for each other. They are both vital to the patient's experience.

In order to increase patient satisfaction and advocacy, the hospital should focus on ways to improve staff (physicians and nurses) communication and increase the time spent with patients. It will be also important to consider what factors impact patients' perception of communication and time spent with them. For Topbox Overall score improvement a focus on timeliness of care within 30 minutes is also very important.

For future studies there are opportunities to look at different scoring systems, the impact of interactions between variables and other data model types. There are clinical research opportunities to understand why these predictor variables are important and what impacts perceptions of these variables.

Chapter 6 - Bibliography

[1] L. Keating and S. Smith, "Acute Pain in the Emergency Department: The Challenges," *Rev Pain,* vol. 5, no. 3, pp. 13-17, 2011.

[2] "R Core Team (2017). R: A language and environment for statical computing. R Foundation for Statistical Computing," [Online]. Available: https://www.r-project.org/.

[3] M. H. Kutner, C. J. Nachtsheim, J. Neter and W. Li, Applied Linear Statistical Models (5th ed.), New York: McGrew-Hill Irwin, 2005.

[4] D. Pregibon, "Logistic Regression Diagnostics," *The Annals of Statistics,* vol. 9, no. 4, pp. 705-724, 1981.

[5] W. Johnson, "Influence Measures for Logistic Regression: Another Point of View," vol. 72, pp. 59-65, 1985.

[6] A. Agresti, An Introduction to Categorical Data Analysis (2nd ed.), New Jersey: Wiley-Interscience, 2007.

[7] R. Christensen, Log-Linear Models and Logistic Regression (2nd ed.), New York: Springer, 1997.

Appendix A - Data Dictionary

Potential predictors (variable name in R)	**Possible values**
Time before someone asked about reason for visit (time.before.asked)	Less than 5 minutes 5 to 15 minutes More than 15 minutes
Received care within 30 minutes of arriving in ER (care.within.30)	Yes, No Blank
Test (test)	Yes No
Meds (meds)	Yes No
Coverage Area (coverage.area)	P=Primary S=Secondary O=Outside
Shift (shift)	A is from 7am to 3 pm B is from 3 pm to 11 pm C is from 11 pm to 7 am
Gender (gender)	M=Male F=Female
Class (class)	1 = Medicare 2 = Managed Care 3 = Traditional Indemnity/Commercial 4 = Medicaid 6 = Self-Pay/None
Age (age)	A=18 to 24 B=25 to 44 C=45 to 64 D=65 and older
Three Level Likert Scale	
Told purpose of new medications (purpose.new.meds)	1=No 2=Yes Somewhat 3=Yes Definitely 0=Blank
Doctors and nurses did everything to help with pain (pain)	
Given as much information as wanted on test results (information.test)	
Kept informed on who was in-charge of care (charge.of.care)	
Nurses spent enough time with patient (nurses.time)	
Doctors spent enough time with patient (doctors.time)	

Four Level Likert Scale	
Courtesy / respect of nurses (courtesy.nurses)	1=Never 2=Sometimes 3=Usually 4=Always 0=Blank
Nurses listened carefully (nurses.listen)	
Clear communication by nurses (nurses.communication)	
Courtesy / respect of doctors (courtesy.doctors)	
Doctors listened carefully (doctors.listen)	
Clear communication by doctors (doctors.communication)	

Table A.1: Data dictionary of covariate variable

Response variables (Variable name in R)	**Possible variables**
Patient Advocacy (likelihood to recommend)	No=Definitely no No=Probably no Yes=Probably yes Yes=Definitely yes
Topbox Overall score	Yes= 9, 10-Best No=0-Worst possible,1,2,3,4,5,6,7,8

Table A.2: Data dictionary of the three response variables

Appendix B - Missing Values

Potential predictors (variable name in R)	Observed Values	Missing Values	Number of Unresolved Missing Values
Time before someone asked about reason for visit (time.before.asked)	433	1	1
Received care within 30 minutes of arriving in ER (care.within.30)	433	1	1
Test (test)	434	0	0
Meds (meds)	434	0	0
Coverage Area (coverage.area)	434	0	0
Shift (shift)	434	0	0
Gender (gender)	434	0	0
Class (class)	434	0	0
Age (age)	434	0	0
Told purpose of new medications (purpose.new.meds)	155	279	0
Doctors and nurses did everything to help with pain (pain)	349	85	0
Given as much information as wanted on test results (information.test)	358	76	3
Kept informed on who was in-charge of care (charge.of.care)	432	2	2
Nurses spent enough time with patient (nurses.time)	430	4	4
Doctors spent enough time with patient (doctors.time)	430	4	4
Courtesy / respect of nurses (courtesy.nurses)	432	2	2
Nurses listened carefully (nurses.listen)	433	1	1
Clear communication by nurses (nurses.communication)	433	1	1
Courtesy / respect of doctors (courtesy.doctors)	429	5	5
Doctors listened carefully (doctors.listen)	428	6	6
Clear communication by doctors (doctors.communication)	430	4	4

Table B.1: Summary of missing predictors

Appendix C - Glossary

Advocacy

Advocacy is when patients say yes, they would recommend the hospital to a close family member or friend.

Likert Scale

Likert Scale is a scoring system used for ordinal data. It takes ordinal categorical data and assigns numbers to the levels.

Overall Score

Overall Score is a discrete score between 0 and 10. The score has values 0, 1, 2, …, 9, 10, Zero is the worst possible score and ten is the best possible score.

Topbox Overall Score

Topbox overall score is when patients rate the hospital as a 9 or a 10 in overall score.

Appendix D - R Code

Data Cleaning Code

```
#Read in data
df1<-read.csv(file="I:/Community Relations/Jason/ED Data/Final Data/Data-2-
11-2018.csv",
              header=TRUE)

#R functions

#f1 will use the scale "Blank"=NA, "No"=1, "Yes, somewhat"=2, "Yes,
definitely"=3
f1<-function(x){
  x<-as.character(x)
  x<-as.numeric(ifelse(x=="No","1",
                ifelse(x=="Yes, somewhat","2",
                ifelse(x=="Yes, definitely", "3", NA))))
  return(x)
}

#f2 will use the scale "Blank"=NA, "Never"=1, "Sometimes"=2, "Usually"=3,
"Always"=4
f2<-function(x){
  x<-as.character(x)
  x<-as.numeric(ifelse(x=="Never","1",
                ifelse(x=="Sometimes","2",
                ifelse(x=="Usually", "3",
                ifelse(x=="Always","4",NA)))))
  return(x)
}

#f3 will use the scale "Blank"=0, "No"=1, "Yes, somewhat"=2, "Yes,
definitely"=3

f3<-function(x){
  x<-as.character(x)
  x<-as.numeric(ifelse(x=="No","1",
                ifelse(x=="Yes, somewhat","2",
                ifelse(x=="Yes, definitely", "3", "0"))))
  return(x)
}

#Responses

#Topbox
#Topbox is when Overall score is 9 or 10.
ftable(df1$Emergency.Room...Overall.Care)
overall.score<-as.character(df1$Emergency.Room...Overall.Care)
topbox<-as.numeric(ifelse(overall.score=="10-Best possible", "1",
                          ifelse(overall.score=="9", "1",
                          ifelse(overall.score=="8", "0",
                          ifelse(overall.score=="7", "0",
                          ifelse(overall.score=="6", "0",
                          ifelse(overall.score=="5", "0",
```

```
                          ifelse(overall.score=="4", "0",
                          ifelse(overall.score=="3", "0",
                          ifelse(overall.score=="2", "0",
                          ifelse(overall.score=="1", "0",
                          ifelse(overall.score=="0-Worst possible", "0",
NA))))))))))))
ftable(topbox)

#Advocacy
#Advocacy is when patients say yes they would recomend the hospital.
ftable(df1$Patient.advocacy..likelihood.to.recommend.)
advocacy<-as.character(df1$Patient.advocacy..likelihood.to.recommend.)
advocacy<-as.numeric(ifelse(advocacy=="Definitely yes",  "1",
                 ifelse(advocacy=="Probably yes","1",
                 ifelse(advocacy=="Probably no", "0",
                 ifelse(advocacy=="Definitely no", "0",NA)))))
ftable(advocacy)

#Time before someone asked about reason for visit
ftable(df1$Time.before.someone.asked.about.reason.for.visit)
time.before.asked<-
as.character(df1$Time.before.someone.asked.about.reason.for.visit)
ftable(time.before.asked)

#Received care within 30 minutes of arriving in ER
ftable(df1$Received.care.within.30.minutes.of.arriving.in.ER)

#Told purpose of new medications
ftable(df1$Told.purpose.of.new.medications)
ftable(df1$Told.purpose.of.new.medications, df1$Meds)
purpose.new.meds<-as.character(df1$Told.purpose.of.new.medications)
purpose.new.meds<-as.numeric(f3(purpose.new.meds))
ftable(purpose.new.meds)
ftable(purpose.new.meds, df1$Meds)

#Doctors and nurses did everything to help with pain
ftable(df1$Doctors.and.nurses.did.everything.to.help.with.pain)
pain<-df1$Doctors.and.nurses.did.everything.to.help.with.pain
pain<-f3(pain)
ftable(pain)

#Given as much information as wanted test results
ftable(df1$Given.as.much.information.as.wanted.on.test.results)
ftable(df1$Given.as.much.information.as.wanted.on.test.results, df1$Test)
test<-df1$Given.as.much.information.as.wanted.on.test.results
test<-f3(test)
test[test==0 & df1$Test=='Y']<-NA
ftable(test)
ftable(test, df1$Test)

#Kept informed on who was in-charge of care
ftable(df1$Kept.informed.on.who.was.in.charge.of.care)
charge.of.care<-df1$Kept.informed.on.who.was.in.charge.of.care
charge.of.care<-f1(charge.of.care)
ftable(charge.of.care)

#Courtesy/respect of nurses
```

```
ftable(df1$Courtesy...respect.of.nurses)
courtesy.nurses<-f2(df1$Courtesy...respect.of.nurses)
ftable(courtesy.nurses)

#Nurses listen carefully
ftable(df1$Nurses.listened.carefully)
nurses.listen<-f2(df1$Nurses.listened.carefully)
ftable(nurses.listen)

#Clear communication by nurses
ftable(df1$Clear.communication.by.nurses)
nurses.communication<-f2(df1$Clear.communication.by.nurses)
ftable(nurses.communication)

#Nurses spent enough time with patients
ftable(df1$Nurses.spent.enough.time.with.patient)
nurses.time<-f1(df1$Nurses.spent.enough.time.with.patient)
ftable(nurses.time)

#Courtesy/respect of doctors
ftable(df1$Courtesy...respect.of.doctors)
courtesy.doctors<-f2(df1$Courtesy...respect.of.doctors)
ftable(courtesy.doctors)

#Doctors listened
ftable(df1$Doctors.listened.carefully)
doctors.listen<-f2(df1$Doctors.listened.carefully)
ftable(doctors.listen)

#Clear communication by doctors
ftable(df1$Clear.communication.by.doctors)
doctors.communication<-f2(df1$Clear.communication.by.doctors)
ftable(doctors.communication)

#Doctors spent enough time with patients
ftable(df1$Doctors.spent.enough.time.with.patient)
doctors.time<-f1(df1$Doctors.spent.enough.time.with.patient)
ftable(doctors.time)

df2<-data.frame(topbox, advocacy,
                time.before.asked,
care.within.30=df1$Received.care.within.30.minutes.of.arriving.in.ER,
                purpose.new.meds=purpose.new.meds, pain,
information.tests=test, charge.of.care,
                courtesy.nurses, nurses.listen, nurses.communication,
nurses.time,
                courtesy.doctors, doctors.listen, doctors.communication,
doctors.time,
                test=df1$Test, meds=df1$Meds,
coverage.area=df1$Coverage.Area,
                shift=df1$Shift, gender=df1$Gender,
class=df1$Class,age=df1$Age)

#Write to file
write.csv(df2, file="I:/Community Relations/Jason/ED Data/Final Data/Final-
Data.csv",row.name=TRUE)
write.csv(df2, file="H:/Desktop/Final-Data.csv",row.name=TRUE)
```

Logistic Code Topbox Overall Score

```
#Topbox

rm(list=ls())

#Reading in the data: Topbox

df1<-read.csv(file="I:/Community Relations/Jason/ED Data/Final Data/Final-
Data.csv", header=TRUE)

y<-df1$topbox

df1$class<-as.factor(df1$class)
df1$time.before.asked<-as.factor(df1$time.before.asked)
df2<-data.frame(y, df1$purpose.new.meds, df1$information.tests,
df1$charge.of.care,
                df1$courtesy.nurses, df1$nurses.listen,
df1$nurses.communication,
                df1$courtesy.doctors, df1$doctors.listen,
df1$doctors.communication,
                df1$test ,df1$meds, df1$coverage.area, df1$shift, df1$gender,
df1$class, df1$age,
                df1$time.before.asked, df1$care.within.30,
                df1$nurses.time, df1$doctors.time, df1$pain)
df2<-df2[complete.cases(df2),]

#Step backward selection: Topbox

#full model
m1<-glm(y~., family=binomial(link='logit'), data=df2);summary(m1)
m01<-step(m1, direction="backward"); summary(m01)
m02<-step(m1, direction="backward", k=2); summary(m02)
m03<-step(m1, direction="backward", k=3); summary(m03)
m04<-step(m1, direction="backward", k=4); summary(m04)
m05<-step(m1, direction="backward", k=5); summary(m05)
m06<-step(m1, direction="backward", k=6); summary(m06)
m07<-step(m1, direction="backward", k=7); summary(m07)
m08<-step(m1, direction="backward", k=8); summary(m08)
m09<-step(m1, direction="backward", k=9); summary(m09)
m10<-step(m1, direction="backward", k=10); summary(m10)
m11<-step(m1, direction="backward", k=11); summary(m11)
m12<-step(m1, direction="backward", k=12); summary(m12)
m13<-step(m1, direction="backward", k=13); summary(m13)
m14<-step(m1, direction="backward", k=14); summary(m14)
m15<-step(m1, direction="backward", k=15); summary(m15)
m16<-step(m1, direction="backward", k=16); summary(m16)
m17<-step(m1, direction="backward", k=17); summary(m17)
m18<-step(m1, direction="backward", k=18); summary(m18)
m19<-step(m1, direction="backward", k=19); summary(m19)
m20<-step(m1, direction="backward", k=20); summary(m20)
m25<-step(m1, direction="backward", k=25); summary(m25)
m30<-step(m1, direction="backward", k=30); summary(m30)
m40<-step(m1, direction="backward", k=40); summary(m40)

#Step forward selection: Topbox
```

```
#Null model
m<-glm(y~1, family=binomial(link='logit'), data=df2); summary(m)

#Full model
m1<-glm(y~., family=binomial(link='logit'), data=df2); summary(m1)

m02<-step(m, direction="forward", scope=list(upper=m1, lower=m));
summary(m02)
m03<-step(m, direction="forward", scope=list(upper=m1, lower=m), k=3);
summary(m03)
m04<-step(m, direction="forward", scope=list(upper=m1, lower=m), k=4);
summary(m04)
m05<-step(m, direction="forward", scope=list(upper=m1, lower=m), k=5);
summary(m05)
m06<-step(m, direction="forward", scope=list(upper=m1, lower=m), k=6);
summary(m06)
m07<-step(m, direction="forward", scope=list(upper=m1, lower=m), k=7);
summary(m07)
m08<-step(m, direction="forward", scope=list(upper=m1, lower=m), k=8);
summary(m08)
m09<-step(m, direction="forward", scope=list(upper=m1, lower=m), k=9);
summary(m09)
m10<-step(m, direction="forward", scope=list(upper=m1, lower=m), k=10);
summary(m10)
m11<-step(m, direction="forward", scope=list(upper=m1, lower=m), k=11);
summary(m11)
m12<-step(m, direction="forward", scope=list(upper=m1, lower=m), k=12);
summary(m12)
m13<-step(m, direction="forward", scope=list(upper=m1, lower=m), k=13);
summary(m13)
m14<-step(m, direction="forward", scope=list(upper=m1, lower=m), k=14);
summary(m14)
m15<-step(m, direction="forward", scope=list(upper=m1, lower=m), k=15);
summary(m15)
m16<-step(m, direction="forward", scope=list(upper=m1, lower=m), k=16);
summary(m16)
m17<-step(m, direction="forward", scope=list(upper=m1, lower=m), k=17);
summary(m17)
m18<-step(m, direction="forward", scope=list(upper=m1, lower=m), k=18);
summary(m18)
m19<-step(m, direction="forward", scope=list(upper=m1, lower=m), k=19);
summary(m19)
m20<-step(m, direction="forward", scope=list(upper=m1, lower=m), k=20);
summary(m20)
m25<-step(m, direction="forward", scope=list(upper=m1, lower=m), k=25);
summary(m25)
m30<-step(m, direction="forward", scope=list(upper=m1, lower=m), k=30);
summary(m30)
m35<-step(m, direction="forward", scope=list(upper=m1, lower=m), k=35);
summary(m35)
m40<-step(m, direction="forward", scope=list(upper=m1, lower=m), k=40);
summary(m40)

#fastbw backward selection: Topbox

#full model
```

```
m<-glm(y~., family=binomial(link='logit'), data=df2)
summary(m)

m001<-step(m1, direction="backward"); summary(m01)

m002<-lrm(y~., data=df2)

fastbw(m01, rule="p", sls=0.1, type="individual")
fastbw(m01, rule="p", sls=0.05, type="individual")
fastbw(m01, rule="p", sls=0.015, type="individual")
fastbw(m01, rule="p", sls=0.01, type="individual")
fastbw(m01, rule="p", sls=0.005, type="individual")
fastbw(m01, rule="p", sls=0.001, type="individual")

m1<-glm(y~df1.care.within.30+df1.doctors.time, family=binomial(link='logit'),
data=df2)
summary(m1)

#interaction model: Topbox

m<-glm(y~df1.nurses.communication*df1.care.within.30*df1.doctors.time,
       family=binomial(link='logit'), data=df2)

summary(m)

#full interaction model
mI<-glm(y~df1.nurses.communication*df1.care.within.30*df1.doctors.time,
family=binomial(link='logit'), data=df2)
summary(mI)

#full additive model
mA<-glm(y~df1.nurses.communication+df1.care.within.30+df1.doctors.time,
family=binomial(link='logit'), data=df2)
summary(mA)

1-pchisq(mA$deviance-mI$deviance, mA$df.residual-mI$df.residual)

#Cook distance: Topbox
m<-glm(y~df1.nurses.communication+df1.care.within.30+df1.doctors.time,
       family=binomial(link='logit'), data=df2)

c1<-cooks.distance(m)

summary(c1)

pf(max(c1), 4, 418-4)

#Hosmer-Lemeshow: Topbox
m<-glm(y~df1.nurses.communication+df1.care.within.30+df1.doctors.time,
       family=binomial(link='logit'), data=df2)

yhat<-predict.glm(m,newdata=df2, type="response", se.fit=FALSE)
dftopbox<-data.frame(y=df2$y, yhat)
dftopbox<-dftopbox[order(dftopbox$yhat),]

dftopbox<-dftopbox[!(is.na(dftopbox$y)),]
```

```
obs<-length(dftopbox$y)
obs<-c(1:obs)
dftopbox[3]<-obs

(e1<-sum(dftopbox[c(1:47),]$yhat))
(o1<-sum(dftopbox[c(1:47),]$y))

dftopbox[c(1:47),]$yhat

(e2<-sum(dftopbox[c(48:99),]$yhat))
(o2<-sum(dftopbox[c(48:99),]$y))

dftopbox[c(48:99),]$yhat

(e3<-sum(dftopbox[c(100:139),]$yhat))
(o3<-sum(dftopbox[c(100:139),]$y))

dftopbox[c(100:139),]$yhat

(e4<-sum(dftopbox[c(140:178),]$yhat))
(o4<-sum(dftopbox[c(140:178),]$y))

dftopbox[c(140:178),]$yhat

(e5<-sum(dftopbox[c(179:418),]$yhat))
(o5<-sum(dftopbox[c(179:418),]$y))

dftopbox[c(179:418),]$yhat

(x1<-(o1-e1)^2/e1+(o2-e2)^2/e2+(o3-e3)^2/e3+(o4-e4)^2/e4+(o5-e5)^2/e5)

1-pchisq(x1, 5-2)

#PseudoR2: Topbox
m<-glm(y~df1.nurses.communication+df1.care.within.30+df1.doctors.time,
       family=binomial(link='logit'), data=df2)
summary(m)
```

Logistic Code for Advocacy

```
#Advocacy

rm(list=ls())

#Reading in the data: Advocacy
#Backward selection using step for Advocacy

df1<-read.csv(file="I:/Community Relations/Jason/ED Data/Final Data/Final-Data.csv", header=TRUE)
y<-df1$advocacy
df1$class<-as.factor(df1$class)
df1$time.before.asked<-as.factor(df1$time.before.asked)
df2<-data.frame(y, df1$purpose.new.meds, df1$information.tests, df1$charge.of.care,
                df1$courtesy.nurses, df1$nurses.listen, df1$nurses.communication,
                df1$courtesy.doctors, df1$doctors.listen, df1$doctors.communication,
                df1$test ,df1$meds, df1$coverage.area, df1$shift, df1$gender, df1$class, df1$age,
                df1$time.before.asked, df1$care.within.30,
                df1$nurses.time, df1$doctors.time, df1$pain)
df2<-df2[complete.cases(df2),]

#Step backward selection: Advocacy

#full model
m1<-glm(y~., family=binomial(link='logit'), data=df2)
summary(m1)

m01<-step(m1, direction="backward"); summary(m01)
m02<-step(m1, direction="backward", k=2); summary(m02)
m03<-step(m1, direction="backward", k=3); summary(m03)
m04<-step(m1, direction="backward", k=4); summary(m04)
m05<-step(m1, direction="backward", k=5); summary(m05)
m06<-step(m1, direction="backward", k=6); summary(m06)
m07<-step(m1, direction="backward", k=7); summary(m07)
m08<-step(m1, direction="backward", k=8); summary(m08)
m09<-step(m1, direction="backward", k=9); summary(m09)
m10<-step(m1, direction="backward", k=10); summary(m10)
m11<-step(m1, direction="backward", k=11); summary(m11)
m12<-step(m1, direction="backward", k=12); summary(m12)
m13<-step(m1, direction="backward", k=13); summary(m13)
m14<-step(m1, direction="backward", k=14); summary(m14)
m15<-step(m1, direction="backward", k=15); summary(m15)
m16<-step(m1, direction="backward", k=16); summary(m16)
m17<-step(m1, direction="backward", k=17); summary(m17)
m18<-step(m1, direction="backward", k=18); summary(m18)
m19<-step(m1, direction="backward", k=19); summary(m19)
m20<-step(m1, direction="backward", k=20); summary(m20)
m25<-step(m1, direction="backward", k=25); summary(m25)
m30<-step(m1, direction="backward", k=30); summary(m30)
m40<-step(m1, direction="backward", k=40); summary(m40)
```

```
#Step forward selection: Advocacy

#Forward Selection using step for Advocacy

m<-glm(y~1, family=binomial(link='logit'), data=df2); summary(m)

#full model
m1<-glm(y~., family=binomial(link='logit'), data=df2); summary(m1)
m02<-step(m, direction="forward", scope=list(upper=m1, lower=m));
summary(m02)
m03<-step(m, direction="forward", scope=list(upper=m1, lower=m), k=3);
summary(m03)
m04<-step(m, direction="forward", scope=list(upper=m1, lower=m), k=4);
summary(m04)
m05<-step(m, direction="forward", scope=list(upper=m1, lower=m), k=5);
summary(m05)
m06<-step(m, direction="forward", scope=list(upper=m1, lower=m), k=6);
summary(m06)
m07<-step(m, direction="forward", scope=list(upper=m1, lower=m), k=7);
summary(m07)
m08<-step(m, direction="forward", scope=list(upper=m1, lower=m), k=8);
summary(m08)
m09<-step(m, direction="forward", scope=list(upper=m1, lower=m), k=9);
summary(m09)
m10<-step(m, direction="forward", scope=list(upper=m1, lower=m), k=10);
summary(m10)
m11<-step(m, direction="forward", scope=list(upper=m1, lower=m), k=11);
summary(m11)
m12<-step(m, direction="forward", scope=list(upper=m1, lower=m), k=12);
summary(m12)
m13<-step(m, direction="forward", scope=list(upper=m1, lower=m), k=13);
summary(m13)
m14<-step(m, direction="forward", scope=list(upper=m1, lower=m), k=14);
summary(m14)
m15<-step(m, direction="forward", scope=list(upper=m1, lower=m), k=15);
summary(m15)
m16<-step(m, direction="forward", scope=list(upper=m1, lower=m), k=16);
summary(m16)
m17<-step(m, direction="forward", scope=list(upper=m1, lower=m), k=17);
summary(m17)
m18<-step(m, direction="forward", scope=list(upper=m1, lower=m), k=18);
summary(m18)
m19<-step(m, direction="forward", scope=list(upper=m1, lower=m), k=19);
summary(m19)
m20<-step(m, direction="forward", scope=list(upper=m1, lower=m), k=20);
summary(m20)
m25<-step(m, direction="forward", scope=list(upper=m1, lower=m), k=25);
summary(m25)
m30<-step(m, direction="forward", scope=list(upper=m1, lower=m), k=30);
summary(m30)
m35<-step(m, direction="forward", scope=list(upper=m1, lower=m), k=35);
summary(m35)
m40<-step(m, direction="forward", scope=list(upper=m1, lower=m), k=40);
summary(m40)

#fastbw backward selection: Advocacy
library(rms)
```

```
m01<-lrm(y~., data=df2)

fastbw(m01, rule="p", sls=0.1, type="individual")
fastbw(m01, rule="p", sls=0.05, type="individual")
fastbw(m01, rule="p", sls=0.025, type="individual")
fastbw(m01, rule="p", sls=0.01, type="individual")
fastbw(m01, rule="p", sls=0.005, type="individual")
fastbw(m01, rule="p", sls=0.001, type="individual")

#full interaction model
mI<-glm(y~df1.doctors.listen*df1.nurses.time, family=binomial(link='logit'),
data=df2)
summary(mI)

#full additive model
mA<-glm(y~df1.doctors.listen+df1.nurses.time, family=binomial(link='logit'),
data=df2)
summary(mA)

1-pchisq(mA$deviance-mI$deviance, mA$df.residual-mI$df.residual)

#Cook distance: Advocacy

m1<-glm(y~df1.doctors.listen+df1.nurses.time, family=binomial(link='logit'),
data=df2)
summary(m1)

c1<-cooks.distance(m1)

summary(c1)
pf(max(c1), 3, 418-3)

#Hosmer-Lemeshow: Advocacy

m1<-glm(y~df1.doctors.listen+df1.nurses.time, family=binomial(link='logit'),
data=df2)
summary(m1)

yhat<-predict.glm(m1,newdata=df2, type="response", se.fit=FALSE)
dftopbox<-data.frame(y=df2$y, yhat)
dftopbox<-dftopbox[order(dftopbox$yhat),]

dftopbox<-dftopbox[!(is.na(dftopbox$y)),]

obs<-length(dftopbox$y)
obs<-c(1:obs)
dftopbox[3]<-obs

(e1<-sum(dftopbox[c(1:47),]$yhat))
(o1<-sum(dftopbox[c(1:47),]$y))

dftopbox[c(1:47),]$yhat


(e2<-sum(dftopbox[c(48:98),]$yhat))
(o2<-sum(dftopbox[c(48:98),]$y))
```

```
dftopbox[c(48:98),]$yhat

(e3<-sum(dftopbox[c(99:139),]$yhat))
(o3<-sum(dftopbox[c(99:139),]$y))

dftopbox[c(99:139),]$yhat

(e4<-sum(dftopbox[c(140:418),]$yhat))
(o4<-sum(dftopbox[c(140:418),]$y))

dftopbox[c(140:418),]$yhat

(x1<-(o1-e1)^2/e1+(o2-e2)^2/e2+(o3-e3)^2/e3+(o4-e4)^2/e4)

1-pchisq(x1, 4-2)
```