Medicaid Provider Hospital Data Analysis by Dakota Cassidy

Information and Background

Description

For this analysis, Medicaid provider hospital data from America was used that contains survey questions answered by patients. The answers are categorized by the state they were in, not the exact hospital. The main conclusion which can be gathered from this data is why someone would or would not recommend a hospital in the state based on the average percentage of patients answering the survey questions.

Question

The question the analysis is being completed to answer is: What factors may determine why a patient said they would not recommend a hospital to another person.

Explanation

This question of what factors may determine why a patient said they would not recommend a hospital to another person is important because, based on the data, the hospitals can correct the issues in the point of view of the patients. The more patients a hospital has the more money they can make. This is good from a business perspective. From a general perspective, people want to feel comfortable and well cared for in a hospital and it is important that everyone feels that way when visiting a hospital.

Methods

Data Description

The data being analyzed is from https://data.cms.gov/provider-data/?redirect=true. CMS.gov is a website that allows people to explore and find Medicare provider data.

The fields the dataset contains are strings (variable), which describe the answers the patients entered in the survey and integers as percentages, which account for the percentage of people who answered a certain way in the survey by state. Each row is an entry for the percentage of people that answered a certain way to a question in a survey in a specific state. Some slight data cleaning was performed, which included deleting a few rows that had state abbreviations AS, GU, MP, and VI, that did not exist and had no data.

Type of Problem and Analyses

This problem is a combination of descriptive and inferential. Descriptive statistics was used to describe and summarize the data. A pivot table as seen in Figure 1, and chart seen in Figure 2, were used to help summarize the data. Pivot tables are great to help develop a high understanding of the data and to accurately summarize the parts that are important to the analysis.

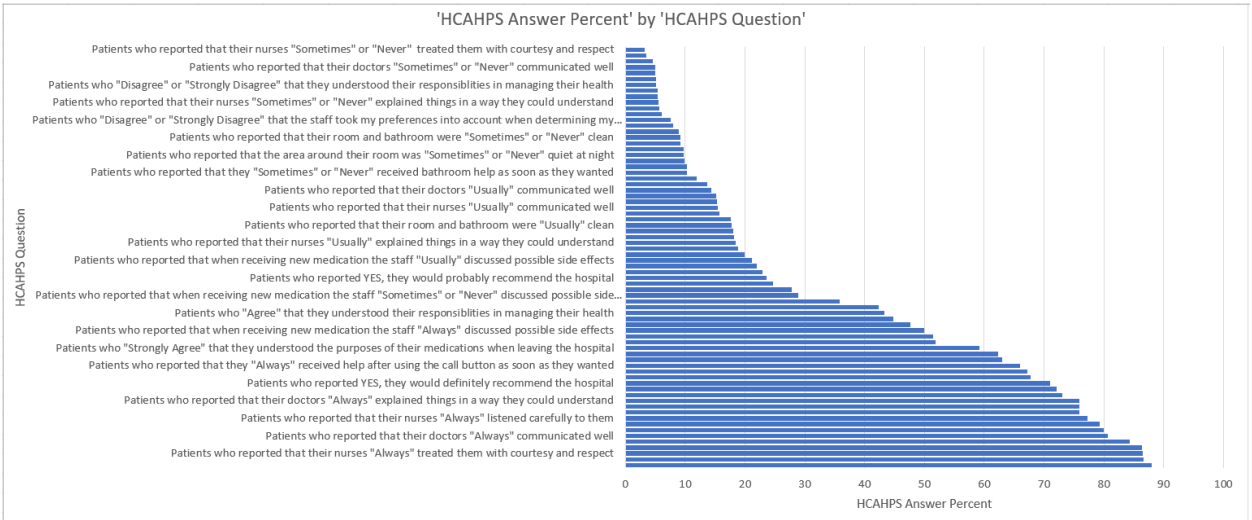
Figure 1

Aggregated data pivot table for descriptive analysis

Row Labels	Average of HCAHPS Answer Percent
"Always" quiet at night	62.34615385
"NO", patients would not recommend the hospital (they probably would not or definitely would not recommend it)	5.365384615
"Sometimes" or "never" quiet at night	9.788461538
"Usually" quiet at night	27.86538462
"YES", patients would definitely recommend the hospital	71.07692308
"YES", patients would probably recommend the hospital	23.55769231
Doctors "always" communicated well	80.71153846
Doctors "always" explained things so they could understand	75.86538462
Doctors "always" listened carefully	79.28846154
Doctors "always" treated them with courtesy and respect	86.71153846
Doctors "sometimes" or "never" communicated well	4.980769231
Doctors "sometimes" or "never" explained things so they could understand	6.115384615
Doctors "sometimes" or "never" listened carefully	5.365384615
Doctors "sometimes" or "never" treated them with courtesy and respect	3.5
Doctors "usually" treated them with courtesy and respect	9.788461538
Doctors "usually" communicated well	14.30769231
Doctors "usually" explained things so they could understand	18.01923077
Doctors "usually" listened carefully	15.34615385
No, staff "did not" give patients information about help after discharge	15.69230769
No, staff "did not" give patients information about possible symptoms	11.96153846
No, staff "did not" give patients this information	13.63461538
Nurses "always" communicated well	80.01923077
Nurses "always" explained things so they could understand	75.90384615
Nurses "always" listened carefully	77.26923077
Nurses "always" treated them with courtesy and respect	86.51923077
Nurses "sometimes" or "never" communicated well	4.538461538
Nurses "sometimes" or "never" explained things so they could understand	5.596153846
Nurses "sometimes" or "never" listened carefully	5.115384615
Nurses "sometimes" or "never" treated them with courtesy and respect	3.211538462
Nurses "usually" treated them with courtesy and respect	10.26923077
Nurses "usually" communicated well	15.44230769
Nurses "usually" explained things so they could understand	18.5
Nurses "usually" listened carefully	17.61538462
Patients "always" received bathroom help as soon as they wanted	67.78846154
Patients "always" received call button help as soon as they wanted	66.03846154
Patients "always" received help as soon as they wanted	67.17307692
Patients "sometimes" or "never" received bathroom help as soon as they wanted	10.28846154
Patients "sometimes" or "never" received call button help as soon as they wanted	9.25
Patients "sometimes" or "never" received help as soon as they wanted	9.884615385
Patients "usually" received bathroom help as soon as they wanted	21.92307692
Patients "usually" received call button help as soon as they wanted	24.71153846
Patients "usually" received help as soon as they wanted	22.94230769
Patients who "Agree" that staff took their preferences into account	47.59615385
Patients who "Agree" they understood their care when they left the hospital	42.36538462
Patients who "Agree" they understood their medications when they left the hospital	35.80769231
Patients who "Agree" they understood their responsiblities when they left the hospital	43.34615385
Patients who "Disagree" or "Strongly Disagree" that staff took their preferences into account	7.538461538
Patients who "Disagree" or "Strongly Disagree" they understood their care when they left the hospital	5.711538462
Patients who "Disagree" or "Strongly Disagree" they understood their medications when they left the hospital	4.980769231
Patients who "Disagree" or "Strongly Disagree" they understood their responsiblities when they left the hospital	5.153846154
Patients who "Strongly Agree" that staff took their preferences into account	44.86538462
Patients who "Strongly Agree" they understood their care when they left the hospital	51.92307692
Patients who "Strongly Agree" they understood their medications when they left the hospital	59.21153846
Patients who "Strongly Agree" they understood their responsiblities when they left the hospital	51.5
Patients who gave a rating of "6" or lower (low)	7.942307692
Patients who gave a rating of "7" or "8" (medium)	19.88461538
Patients who gave a rating of "9" or "10" (high)	72.17307692
Room was "always" clean	73.01923077
Room was "sometimes" or "never" clean	9.173076923
Room was "usually" clean	17.80769231
Staff "always" explained	63.05769231
Staff "always" explained new medications	75.90384615
Staff "always" explained possible side effects	49.98076923
Staff "sometimes" or "never" explained	18.82692308
Staff "sometimes" or "never" explained new medications	8.961538462
Staff "sometimes" or "never" explained possible side effects	28.90384615
Staff "usually" explained	18.11538462
Staff "usually" explained new medications	15.13461538
Staff "usually" explained possible side effects	21.11538462
Yes, staff "did" give patients information about help after discharge	84.30769231
Yes, staff "did" give patients information about possible symptoms	88.03846154
Yes, staff "did" give patients this information	86.36538462
Grand Total	**34.72222222**

Figure 2

Chart output for descriptive analysis



Using the data provided, an inferential analysis was completed about the general population to figure out which factors may have caused a patient to say they would not recommend the hospital. Regression was used because regression makes the most sense to determine the effect of different factors on the patients' answers to not recommend the hospital. This type of problem is inferential because the focus is on those who affect the input of the variables that affect the patients answer, allowing someone to be able to change those input variables to change the effect on the patients answer.

Discussion of Analyses

The descriptive and inferential analyses ran have been helpful in answering what factors may determine why a patient said they would not recommend a hospital to another person. The pivot table and chart have given an informative summary of the data by state. The simple

regression analysis has given insight into what factors affect the patients' decisions to not recommend a hospital. These factors will be useful in determining how to improve hospital ratings and recommendations. Because the data is not split into separate entries of each survey completed by a person it is harder to get an accurate cause and affect analysis. Multiple regression analysis was also run to compare the most significant independent variables against the dependent variable.

Results

Interpretation of Results of Analyses

After creating a pivot table to gather key information about the data, Regression analyses were run on each independent variable separately against the dependent variable. From each Regression output, the results were analyzed for significance. First the R-squared value was considered, being the explanation of variance of the dependent variable by the independent variable, the lowest R-squared considered was .76. Then the coefficients, and p-values were addressed. The coefficient determined the strength of the relationship between the dependent and independent variables, and the p-value determined if the independent variable was significant enough if it were lower than .05.

The most significant variables were as follows: Staff sometimes or never explained side effects with an R-squared value of .79, a coefficient of .35 and p-value less than .05, Patients who gave a rating of 6 or lower with an R-squared value of .94, a coefficient of .79, and p-value less than .05, Patients who disagree or strongly disagree they understood their responsibilities

when they left the hospital with an R-squared value of .82, a coefficient of 1.08, and a p-value less than .05, Patients who disagree or strongly disagree they understood their care when they left the hospital with an R-squared value of .9, a coefficient of 1.02 and a p-value less than .05, Patients who disagree or strongly disagree that staff took their preferences into account with an R-squared value of .92, a coefficient of .71, and a p-value less than .05, Patients who sometimes or never received help as soon as they wanted with an R-squared value of .78, a coefficient of .48 and a p-value less than .05, Patients sometimes or never received call button help as soon as they wanted with an R-squared value of .81, a coefficient of .47, and p-value less than .05, Nurses sometimes or never treated them with courtesy or respect with an R-squared value of .79, a coefficient of 1.26 and a p-value less than .05, Nurses sometimes or never listened carefully with an R-squared value of .86, a coefficient of .91, and ap-value less than .05, Nurses sometimes or never explained things so they could understand with an R-squared value of .85, a coefficient of .99, and a p-value less than .05, Nurses sometimes or never communicated well with an R-squared value of .86, a coefficient of 1.02, and a p-value less than .05, Doctors sometimes or never listed carefully or showed respect with an R-squared value of .81, a coefficient of 1.04, and a p-value less than .05, Doctor sometimes or never explained things so they could understand with an R-squared value of .76, a coefficient of .98, and p-value less than .05, lastly Doctor sometimes or never communicated well with a R-squared value of .78, a coefficient of 1.12, and a p-value less than .05.

After gathering the most significant independent variables they were run together in a Multiple Regression Analysis to compare against each other and determine which were most impactful on the dependent variable based on the coefficients. In the comparison of significant

independent variables, the Multiple Regression analysis seen in Figure 3, revealed that the most highly correlated variables were Patients who gave a rating of 6 or lower with a coefficient of .47, Nurses sometimes or never communicated well with a coefficient of .35, and Patients who disagree or strongly disagree they understood their care when they left the hospital with a coefficient of .22. The regression results for each of these variables ran separately can be seen in Figure 4, 5, and 6 respectively.

Figure 3

Multiple Regression of most significant independent variables output

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.998468115
R Square	0.996938577
Adjusted R Square	0.967885207
Standard Error	0.377413833
Observations	52

ANOVA

	df	*SS*	*MS*	*F*	*gnificance F*
Regression	16	1669.872	104.367	732.7023806	2.47E-39
Residual	36	5.127883	0.142441		
Total	52	1675			

	Coefficients	*andard Err*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*ower 95.0%*	*pper 95.0%*
Intercept	0	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
Doctors "sometimes" or "never" explained things so they could understand	0.004758922	0.109465	0.043474	0.965563667	-0.21725	0.226764	-0.21725	0.226764
Doctors "sometimes" or "never" listened carefully	0.001510805	0.129244	0.01169	0.99073784	-0.26061	0.26363	-0.26061	0.26363
Nurses "always" treated them with courtesy and respect	-0.001954467	0.005778	-0.33828	0.737118694	-0.01367	0.009763	-0.01367	0.009763
Nurses "sometimes" or "never" communicated well	0.354728677	0.188778	1.879079	0.068348734	-0.02813	0.737588	-0.02813	0.737588
Nurses "sometimes" or "never" explained things so they could understand	-0.118532544	0.179265	-0.66121	0.512685788	-0.4821	0.245035	-0.4821	0.245035
Nurses "sometimes" or "never" listened carefully	0.01273202	0.121476	0.104811	0.91710731	-0.23363	0.259096	-0.23363	0.259096
Nurses "sometimes" or "never" treated them with courtesy and respect	0.115206193	0.162035	0.710995	0.481669004	-0.21342	0.443829	-0.21342	0.443829
Patients "sometimes" or "never" received call button help as soon as they wanted	-0.006500799	0.08845	-0.0735	0.941817786	-0.18589	0.172885	-0.18589	0.172885
Patients "sometimes" or "never" received help as soon as they wanted	-0.029436864	0.089298	-0.32965	0.743576951	-0.21054	0.151668	-0.21054	0.151668
Patients who "Disagree" or "Strongly Disagree" that staff took their preferences into account	0.047254385	0.114319	0.413357	0.681797093	-0.18459	0.279103	-0.18459	0.279103
Patients who "Disagree" or "Strongly Disagree" they understood their care when they left the hospita	0.227042661	0.149292	1.5208	0.137045574	-0.07573	0.52982	-0.07573	0.52982
Patients who "Disagree" or "Strongly Disagree" they understood their responsiblities when they left	0.077948249	0.15377	0.506914	0.615305953	-0.23391	0.389808	-0.23391	0.389808
Patients who gave a rating of "6" or lower (low)	0.47105007	0.105062	4.483545	7.19253E-05	0.257975	0.684126	0.257975	0.684126
Staff "sometimes" or "never" explained	-0.053832798	0.144468	-0.37263	0.711610575	-0.34683	0.239162	-0.34683	0.239162
Staff "sometimes" or "never" explained new medications	-0.147597708	0.11506	-1.28279	0.207765074	-0.38095	0.085755	-0.38095	0.085755
Staff "sometimes" or "never" explained possible side effects	0.034843036	0.077971	0.446874	0.657642778	-0.12329	0.192975	-0.12329	0.192975

Figure 4

Regression output for inferential analysis, Patients who gave a rating of 6 or lower

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.973055
R Square	0.946837
Adjusted R	0.945774
Standard E	0.435111
Observatic	52

ANOVA

	df	*SS*	*MS*	*F*	*ignificance F*
Regressior	1	168.5916	168.5916	890.5028	1.59E-33
Residual	50	9.46609	0.189322		
Total	51	178.0577			

	Coefficients	*tandard Erı*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*ower 95.0%*	*pper 95.0%*
Intercept	-0.94782	0.219996	-4.30835	7.69E-05	-1.38969	-0.50595	-1.38969	-0.50595
Patients w	0.794883	0.026637	29.84129	1.59E-33	0.741381	0.848385	0.741381	0.848385

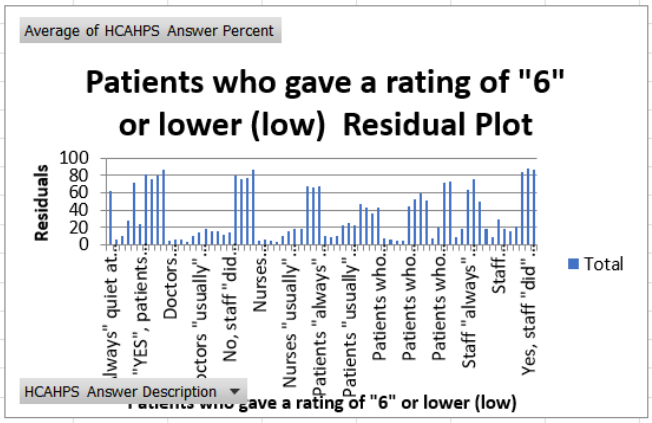


Figure 5

Regression output for inferential analysis, Nurses never communicated well

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.932153
R Square	0.86891
Adjusted R	0.866288
Standard E	0.683251
Observatic	52

ANOVA

	df	*SS*	*MS*	*F*	*ignificance F*
Regressior	1	154.7161	154.7161	331.4167	1.04E-23
Residual	50	23.34162	0.466832		
Total	51	178.0577			

	Coefficients	*tandard Erı*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*ower 95.0%*	*pper 95.0%*
Intercept	0.708115	0.272808	2.595652	0.012361	0.160164	1.256067	0.160164	1.256067
Nurses "so	1.026178	0.056368	18.20485	1.04E-23	0.912959	1.139397	0.912959	1.139397

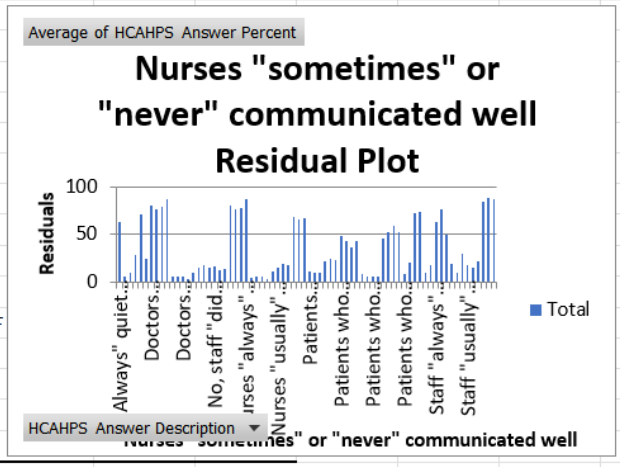


Figure 6

Regression output for inferential analysis, Patients disagree they understand their care

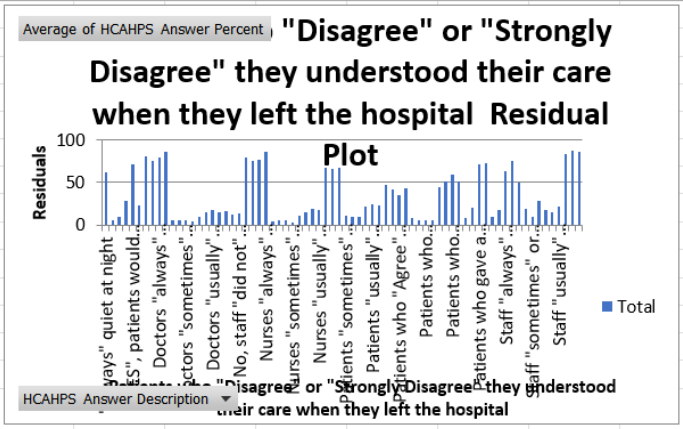
P

Regression Statistics	
Multiple R	0.949072
R Square	0.900739
Adjusted R	0.898753
Standard E	0.594546
Observatic	52

ANOVA

	df	*SS*	*MS*	*F*	*ignificance F*
Regressior	1	160.3834	160.3834	453.7202	9.81E-27
Residual	50	17.67427	0.353485		
Total	51	178.0577			

	Coefficients	*tandard Erı*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*ower 95.0%*	*pper 95.0%*
Intercept	-0.4886	0.286927	-1.70287	0.094801	-1.06491	0.087709	-1.06491	0.087709
Patients w	1.02494	0.048118	21.30071	9.81E-27	0.928293	1.121587	0.928293	1.121587



Implications and Conclusion

Discussion of the Implications of the Results

The significant independent variables that impacted the dependent variable the most were ones that made sense after reading over them. People wish to be treated with care when they are in bad health and feel like they made the right choice going to a certain hospital. Looking at the significant variables, most of them had to do with the behavior of Nurses towards patients. Even though most patients feel more trusting of a Doctor than a Nurse, the Nurses are caring for them the entire time of their visit while they see the doctors for only a few moments. In those few moments of interaction, the Doctors make a bigger impact, but Nurses have more time spent with the patients to negatively affect their opinion of the

hospital. The most obvious connection between variables was that patients who rated the hospital lower than a 6 were closely related to the patients who did not recommend the hospital. The Multiple Regression revealed that communication with Nurses is most related to a patient's decision to not recommend a hospital. How well the communication is will ultimately affect the entire stay at the hospital for a patient, and it will also affect whether they feel they understand their care when leaving the hospital, which was another of the highest correlated variables. Hospitals should work on communication skills for their staff, to improve their ratings and decrease the chance that someone would choose to not recommend their hospital. This will help more patients feel safe and cared for, and it will also tremendously increase business for the hospital.

Conclusion

The Medicaid provider hospital data from America that was used containing survey questions answered by patients from CMS.gov was able to answer the question: What factors may determine why a patient said they would not recommend a hospital to another person? Answering this question is important to help a hospital make more money, and to make patients feel safer and more comfortable. By determining what factors make a patient not recommend the hospital, they can correct and change those factors, to ensure a higher recommendation rate. The factors that need to be changed are proven to revolve around the way staff (doctors and nurses) communicate with and treat their patients. Training should be implemented to improve this issue, and hospitals will be able to yield a higher return, and happier patients.

References

Patient Survey (HCAHPS) – State. (2022, July 7). The Centers for Medicare & Medicaid Services. Retrieved July 7, 2022, from https://data.cms.gov/provider-data/dataset/84jm-wiui