

Major impediments to effective rural clinic management identified through Decision tree techniques, Regression and Clustering

Onkar Jadhav, Ipe Paramel

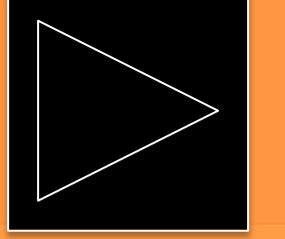
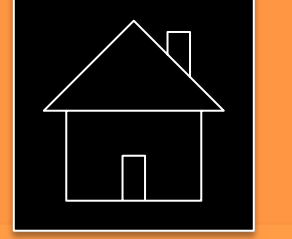
Affiliated to Oklahoma State University



; CHSI (Center for Health Systems Innovation)



; Spears School of Business



Abstract

CHSI (Center of Health Systems Innovation), OSU established in 2012 focuses on rural Oklahoma healthcare and clinical innovations. The Rural Workflow Efficiency team, part of CHSI, conducts the collection and analysis of clinical workflow data from a diverse & discrete rural population. Clinics in rural Oklahoma face several issues such as “no shows” for scheduled appointments, appointment cancellations, transportation problems, multiple interruptions and other workflow and efficiency related problems. We interviewed all clinic constituents (staff/providers/nurses) from 34 different clinics in rural Oklahoma, and recorded their responses for 29 questions. These questions were related to Electronic Medical Records, Employee Management, Patient Experience, known Patient Management Challenges and additional “question corresponding” fields. Recorded responses were numerically coded. Ratings for each clinic were developed by grouping the clinics into 5 categories from best to worst using a cluster analysis. Correlations were found among questions and between questions and ratings. In addition, we had clinic property data for the 34 clinics which had 40 metrics corresponding to each clinic i.e. number of staff, number of providers, expected number of patients per day, hours of operation etc. Descriptive analysis was performed on collected data to identify existing impediments and common problematic themes across rural Oklahoma. Correlations and linear regression were applied to detect interacting metrics and finally, a decision tree analysis was used to explore specific decision paths the data could reveal. The interaction of survey data and the clinic property data was also observed.

The Data , numeric coding and recoding

- The survey data – Responses given by clinic constituents for 29 different questions, Nature of the responses – (Strongly disagree - Disagree – Agree – Strongly agree) / (Yes-No) / (Very Frequently – Frequently – Less frequent – Never) /, etc.
Responses were numerically coded on a scale of (-2 to 2). Strongly Disagree/Never/No = -2, neutral = 0, Agree/Frequently = 1, etc. Negatively connoted responses were assigned negative values appropriately, vice versa.
Aggregation of responses for respective clinics and averaging the responses based on the number of respondents led to the development of “Clinic Ratings”, later used as a target variable for clinic evaluation.
- Clinic property data – This data had clinic parameters/classification information like number of providers, expected number of patients, actual number of patients, cancellations, hours spent by provider, etc., as numeric variables.
Examples of character variables would be services outsourced, days open, outsourced activities, etc.
- Map Layout data – This data had clinic layout parameters like clinic building shape, number of waiting room seats, exam rooms, labs and X-rays

Approach , Computations and methods

- Survey Data: Used to develop “Clinic Ratings”, later used as a target variable. Clinical ratings were developed by aggregating responses given by responders to 29 questions. Also “Question Scores” were developed to grade the questions and group them in several categories such as Electronic Med Record, Employee Management, Patient Experience and Known Patient Management Challenges. This enabled evaluation of clinics on workflow efficiency, management aspects and other facets.
- Clinic property data combined with the Map Layout Data, resulted in 44 parameters (variables). The target variable was computed from the survey data file (Clinic Ratings) and correlating variables were identified. Additionally, multi collinearity was examined and a regression model was built.
 - Decision tree analysis performed on combined clinical data and map layout data, surfaced interesting insights.
 - Clustering clinics into 5 different categories, ranging from Best (A) to Worst (E), was based on clinical ratings and conclusions from decision trees, considering the variable importance criteria from decision trees.

Fig. 2.1 Survey questions , their numeric coding , sum of responses and Average clinic ratings

	A	AA	AB	AC	AD	AE	AF	AG
1	Response ID	25_When there are many patients in the clinic, the clinic functions efficiently and effectively without disruption or chaos.	26_We have patients that return with recurrent (non-chronic) issues for which they were previously treated.	27_Patients with appointments that do not show up for their appointment affect clinic efficiency and flow?	28_The clinic accepts and works in all walk-ins.	29_Walk-ins disrupt the workflow in the clinic?	Sum of all responses	Avg. ratings for clinics
2	C1_1							
3	C1_2	1	1	1	1	1	19	
4	C1_3	1	0	0	-1	0	11	15
5	C2_1	2	0	0	2	1	25	
6	C2_2	1	-2	2	2	2	27	
7	C2_3	-1	-2	0	2	0	12	
8	C2_4	1	1	1	1	1	35	
9	C2_5	1	0	1	2	0	26	
10	C2_6	2	1	1	2	1	30	25.83
11	C3_1	-1	-1	-2	2	-2	3	
12	C3_2	1	0	1	-1	-1	14	
13	C3_3	0	-1	-1	1	-1	-6	
14	C3_4	-1	-1	-1	1	-1	2	
15	C3_5	1	0	-1	-1	-1	15	
16	C3_6	1	-2	-2	2	-2	-3	4.16
17	C4_1	2	-2	-2	2	-2	2	
18	C4_2	1	-1	-1	1	-1	7	
19	C4_3	0	0	0	0	0	0	3
20	C5_1	1	0	1	2	2	32	
21	C5_2	0	-1	1	1	1	15	

Major impediments to effective rural clinic management identified through Decision tree techniques, Regression and Clustering

Onkar Jadhav, Ipe Paramel

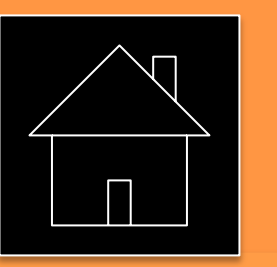
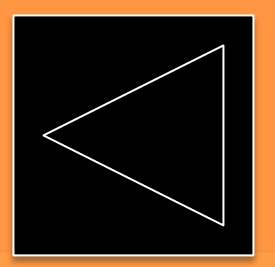
Affiliated to Oklahoma State University



; CHSI (Center for Health Systems Innovation)



; Spears School of Business



Results And Conclusions

Fig 2.2 Grouping the clinics as per the average ratings and number of respondents

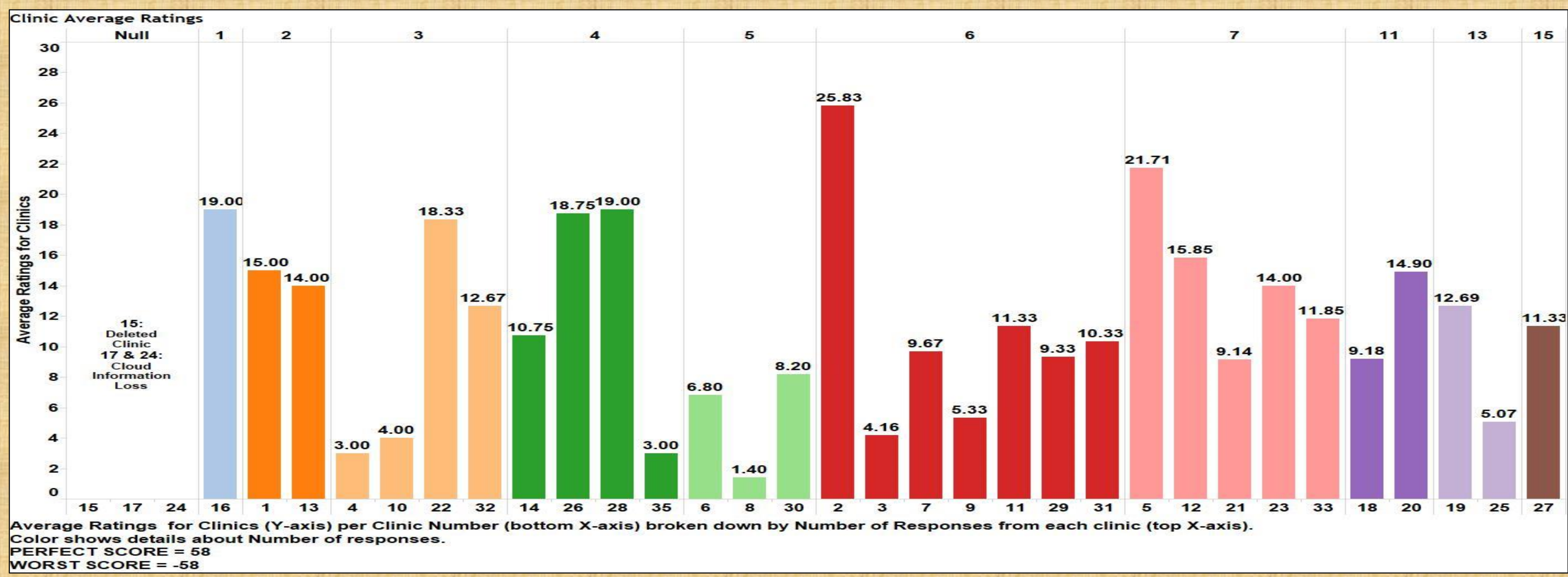
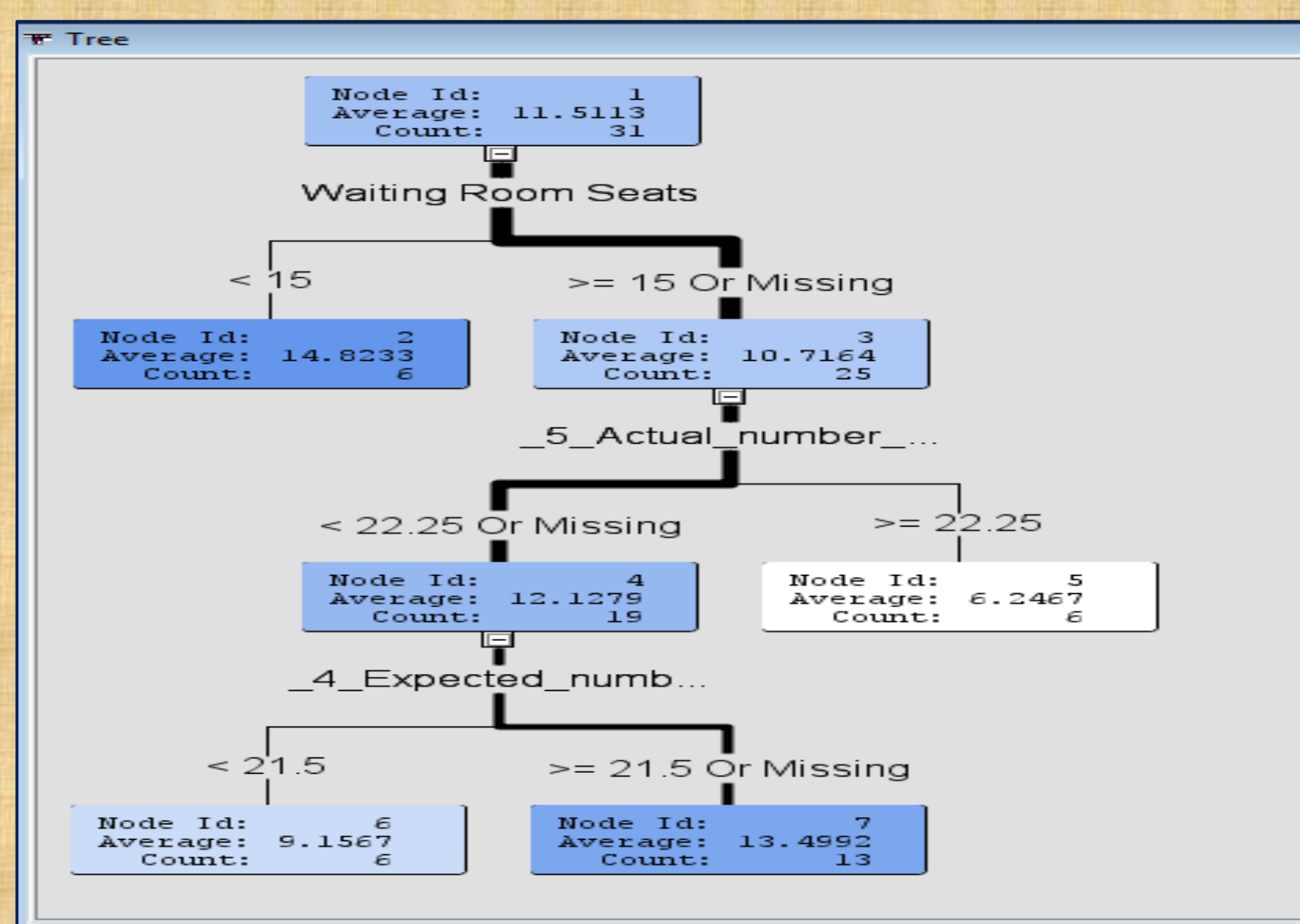
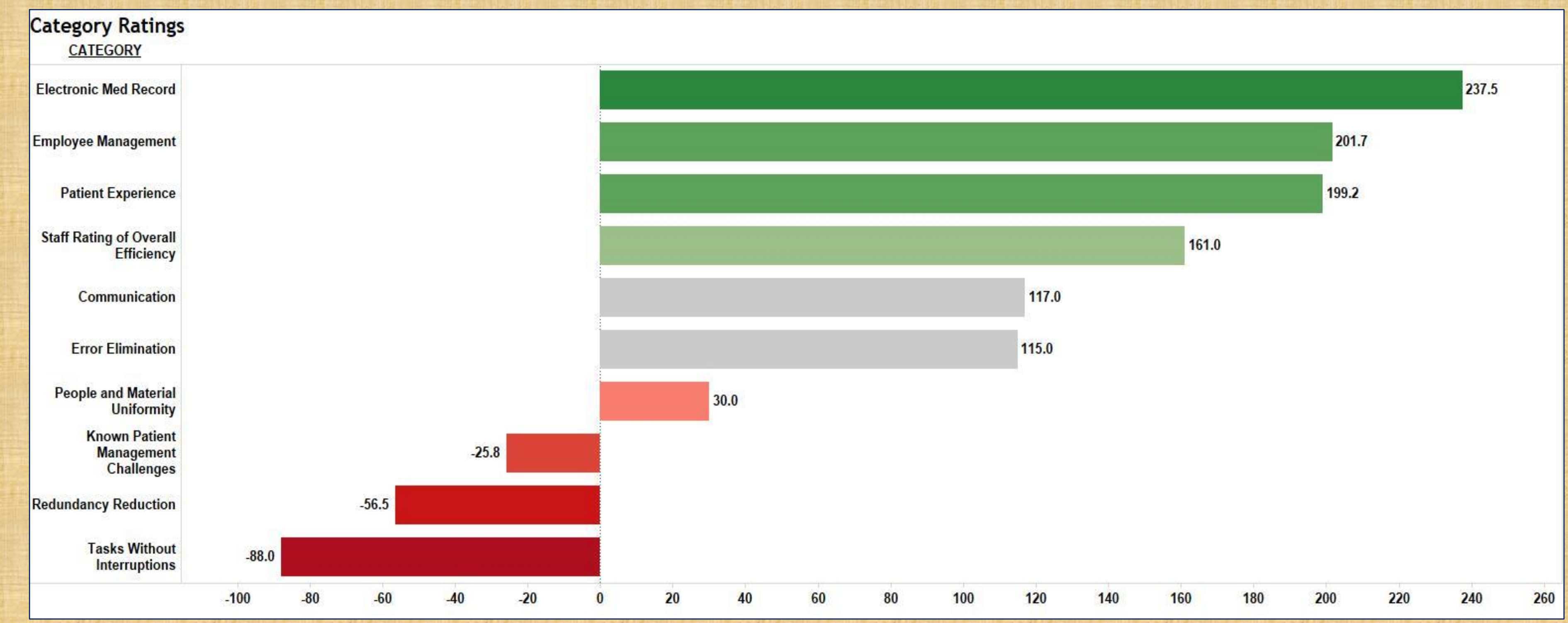


Fig. 2.3 Decision tree result snippet (Waiting Room Seats)



- Clinics 2,5, & 16 performed well on parameters listed in Fig 2.4 while clinics 4,8,10, & 35 needed immediate attention.
- If Cancellations are ≥ 4.5 , clinics are rated lower, if < 4.5 then rated higher. If the higher rated clinics at this node have Exam rooms < 7.5 , the rating (performance) is furthermore higher.
- If the number of Waiting Room Seats > 15 , clinics are rated lower, in addition if actual number of Waiting Room Seats ≥ 15 and the actual number of patients per provider per hour < 22.5 , then ratings are average (comparatively better). If the expected number of patients per provider per hour ≥ 21.5 , rating stays constant as previous level.

Fig 2.4 Categories used to measure clinic performance and scores of these categories.



- If the number of walk-ins ≥ 5.5 , clinics are rated higher, in addition to this by number of exam rooms < 7.5 , clinics are rated still higher.
- EMR, Employee management, and Patient Experience are fields where clinics fared better while there were problems associated with Interruptions in regular tasks, and Redundancy in information storing which consumed time

Results from regression analysis:

- Cancellations and Average Time Spent by provider in the room in minutes are positively correlated with coefficient $= 0.83827$

Relation : $Cancellations = -1.503 + 0.424$ (average time of provider in minutes)

Interpretation : For every 2.5 minutes extra taken by the provider there is one cancellation.
- The variables, Exam Rooms, and Number of Providers are positively correlated, correlation coefficient $= 0.674$

Relation : $exam\ rooms = 4.81 + 1.06$ (Providers)

Interpretation : For each provider currently there is one exam room, whereas ideally there should be 2.5 rooms per provider i.e. 5 rooms for 2 providers