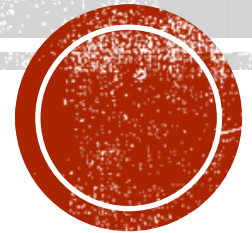


HOSPITAL OVERALL RATING

PRESENTED BY:

SEEMA S B



PROBLEM STATEMENT

- Evanston Hospital is a acute-care hospital in Illinois, US which offers services and surgical specialties', in addition to having high-end lab capabilities.
- Though hospital is equipped, it's CMS rating is 3 due to which there is decline in revenue for the hospital.
- Since the rating influence the patients choice of hospital. So as a consulting company, we need to identify factors that impact the hospital rating and suggestive measures that need to be taken.



APPROACH TO THE SOLUTION

The approach followed to get to the solution is as follows:

1. Data Understanding
2. Data Cleaning
3. Data Visualization
4. Data Preparation
5. Model Building
6. Evaluation



ABOUT THE DATASET

- 4818 rows, 28 columns
- 25 categorical columns and 3 numeric columns

```
In [2]: genInfo = pd.read_csv('Hospital General Information.csv', encoding='cp1252')
genInfo.head()
```

Out[2]:

Provider ID	Hospital Name	Address	City	State	ZIP Code	County Name	Phone Number	Hospital Type	Hospital Ownership	Emergency Services	Meets criteria for meaningful use of EHRs	Hospital overall rating	Hospital overall rating footnote	co
10001	SOUTHEAST ALABAMA MEDICAL CENTER	1108 ROSS CLARK CIRCLE	DOTHAN	AL	36301	HOUSTON	3347938701	Acute Care Hospitals	Government - Hospital District or Authority	Yes	Y	3	NaN	Sa
10005	MARSHALL MEDICAL CENTER SOUTH	2505 U S HIGHWAY 431 NORTH	BOAZ	AL	35957	MARSHALL	2565938310	Acute Care Hospitals	Government - Hospital District or Authority	Yes	Y	3	NaN	
10006	ELIZA COFFEE MEMORIAL HOSPITAL	205 MARENGO STREET	FLORENCE	AL	35631	LAUDERDALE	2567688400	Acute Care Hospitals	Government - Hospital District or Authority	Yes	Y	2	NaN	
10007	MIZELL MEMORIAL HOSPITAL	702 N MAIN ST	OPP	AL	36467	COVINGTON	3344933541	Acute Care Hospitals	Voluntary non-profit - Private	Yes	Y	3	NaN	Sa
10008	CRENSHAW COMMUNITY HOSPITAL	101 HOSPITAL CIRCLE	LUVERNE	AL	36049	CRENSHAW	3343353374	Acute Care Hospitals	Proprietary	Yes	Y	3	NaN	Sa



DATA CLEANING

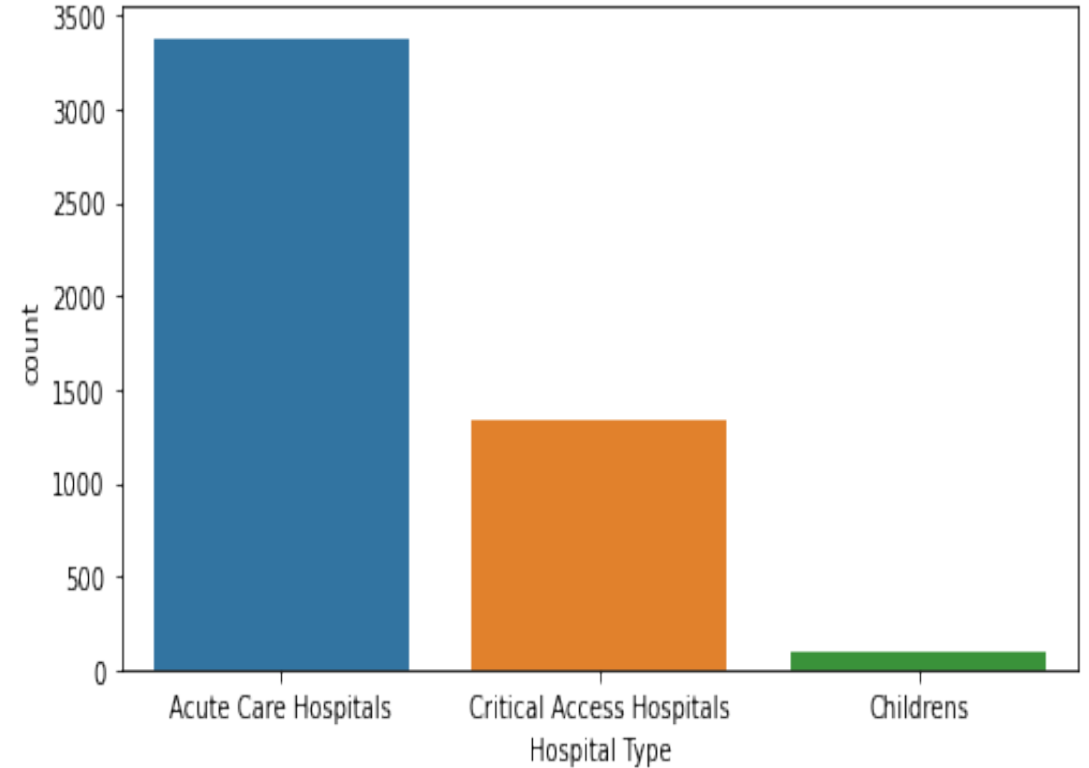
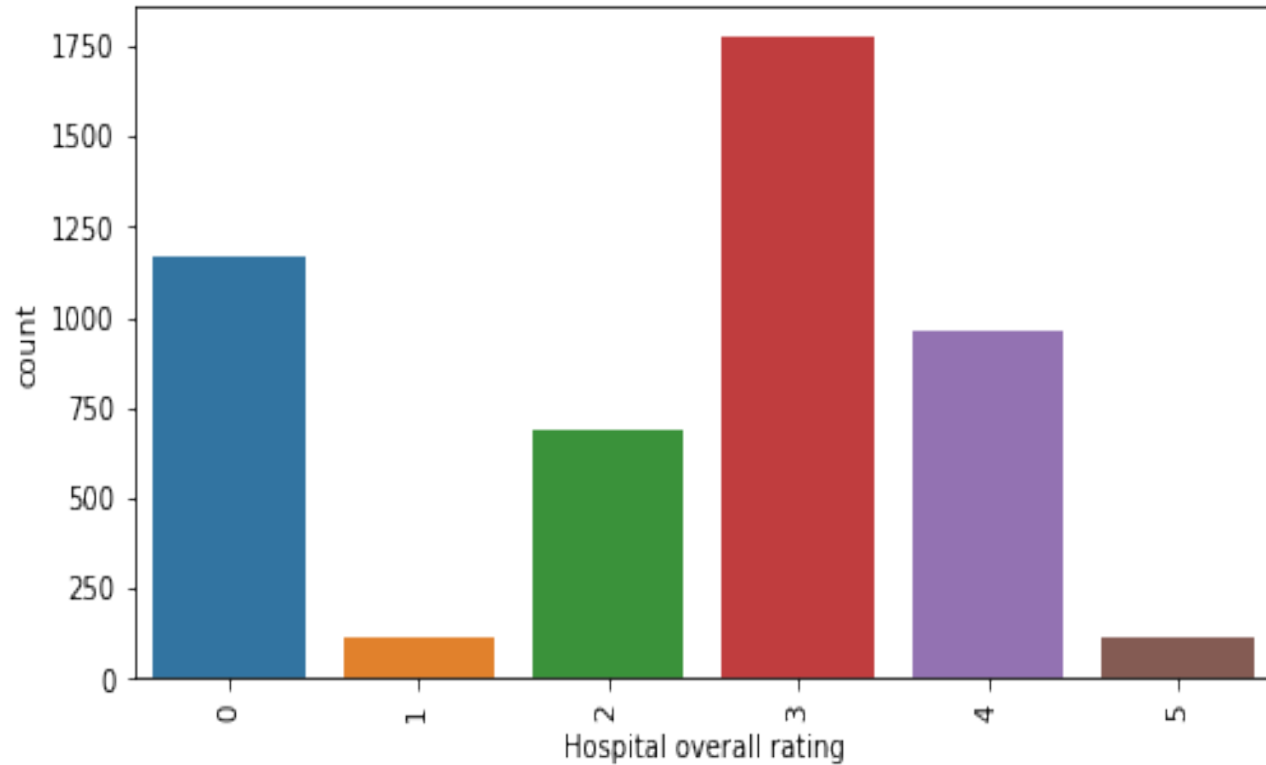
- Columns with 50% null values and more, i.e.,

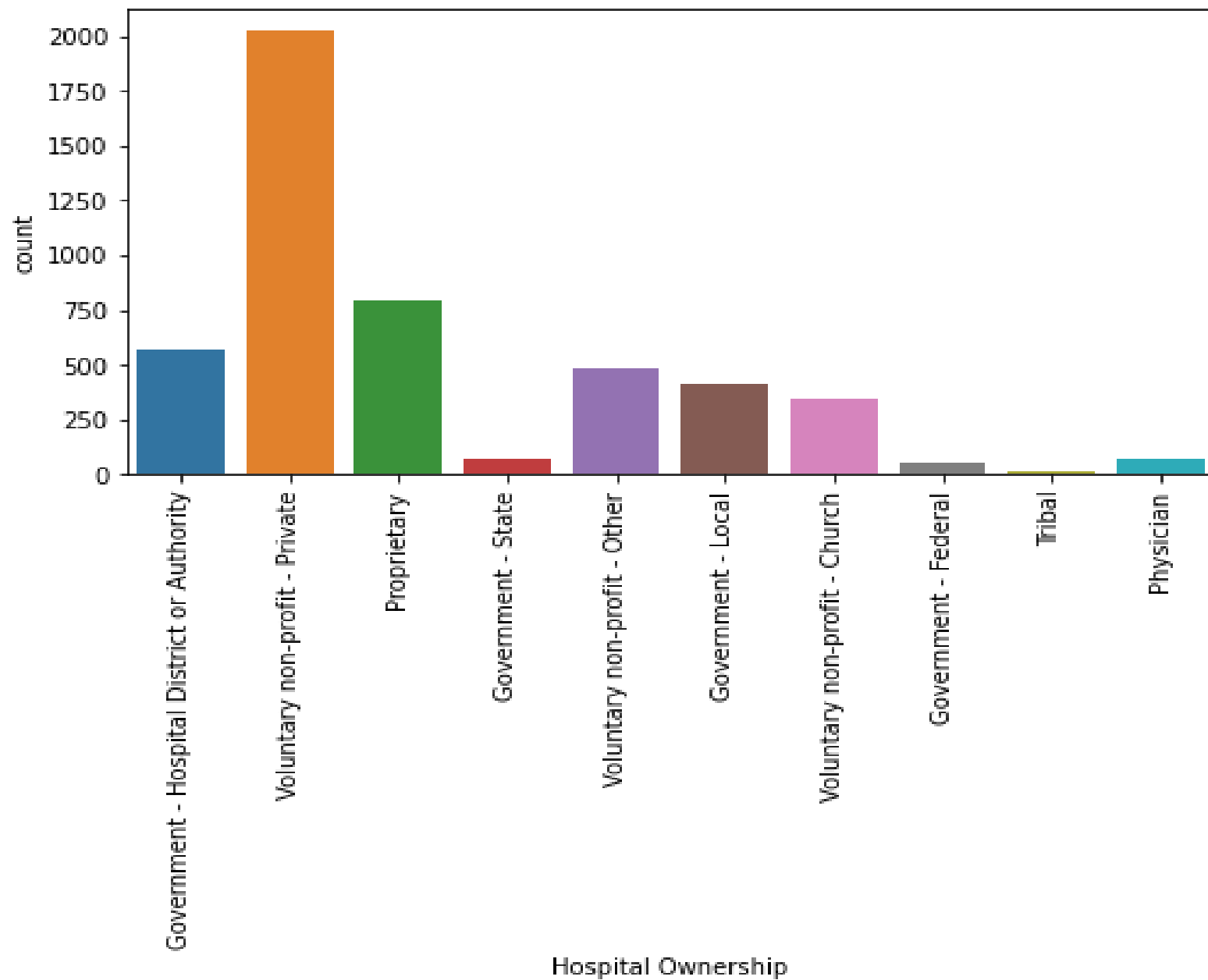
Hospital overall rating footnote, Safety of care national comparison footnote, Readmission national comparison footnote, Patient experience national comparison footnote, etc.

- Emergency Services, Meets criteria for meaningful use of EHRs columns are imputed with 0s and 1s in place of no and yes.
- The target column 'hospital overall rating' changed to 'int'.



DATA VISUALIZATION





DATA PREPARATION

- The dummies are created for the columns like

Hospital Type, Hospital Ownership, Mortality national comparison, Safety of care national comparison, Readmission national comparison, etc.

- Mentioned original columns are dropped.
- Also City, State, Country, Provider ID, Phone Number etc. are dropped.



MODEL BUILDING

- As it is a classification problem, performed Logistic Regression with help of RFE.
- Since accuracy and recall were low, performed Decision tree with hyper parameter tuning to get best model.
- Also performed Random forest with hyper parameter tuning to have better a have better predictive model.

Below is the table of summarized results:

	Method	Accuracy
0	Logistic Regression Train	0.60
0	Logistic Regression Test	0.59
0	Decision Tree Train	0.59
0	Decision Tree Test	0.57
0	Decision Tree Hyperparameter Tuning Train	0.74
0	Decision Tree Hyperparameter Tuning Test	0.69
0	Random Forest Train	0.58
0	Random Forest Test	0.56
0	Random Forest Hyperparameter Tuning Train	0.69
0	Random Forest Hyperparameter Tuning Test	0.65



EVALUATION

- Based on the table below, Decision Tree with hyperparameter tuning is the Best Model for rating prediction of past hospital data.
- Even from the confusion matrix, it can be seen that the probability of Actual rating predicted correctly by the model is good when compared to others which is important for a Hospital.

Method	Accuracy	Precision	Recall
Logistic Regression Train	0.6	0.54	0.49
Logistic Regression Test	0.59	0.53	0.47
Decision Tree Train	0.59	0.5	0.66
Decision Tree Test	0.57	0.5	0.61
Decision Tree Hyperparameter Tuning Train	0.74	0.64	0.82
Decision Tree Hyperparameter Tuning Test	0.69	0.57	0.7
Random Forest Train	0.58	0.48	0.64
Random Forest Test	0.56	0.46	0.61
Random Forest Hyperparameter Tuning Train	0.69	0.59	0.77
Random Forest Hyperparameter Tuning Test	0.65	0.55	0.7



RESULT AND CONCLUSION

The Decision tree with hyper parameter tuning is the final model chosen for prediction.

The Important features that impact the Hospital overall rating are as follows,

1. Patient experience
2. Readmission
3. Effectiveness of care
4. Safety of care
5. Mortality

