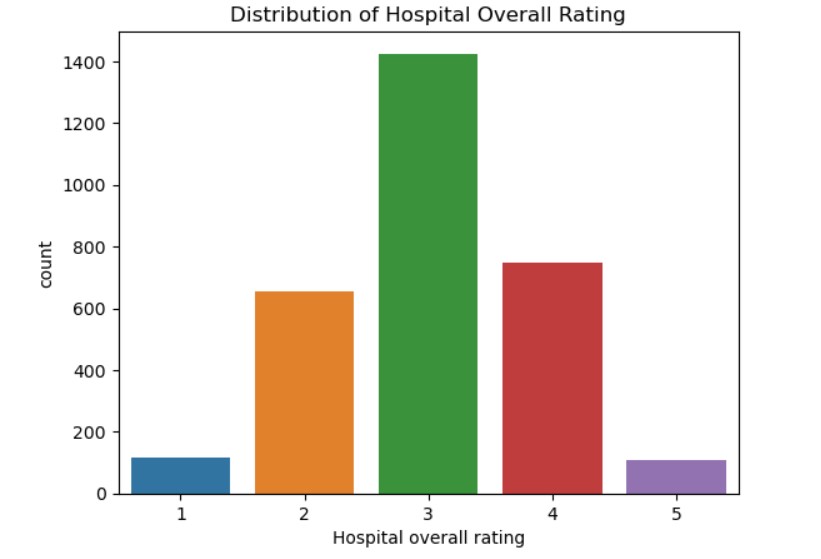
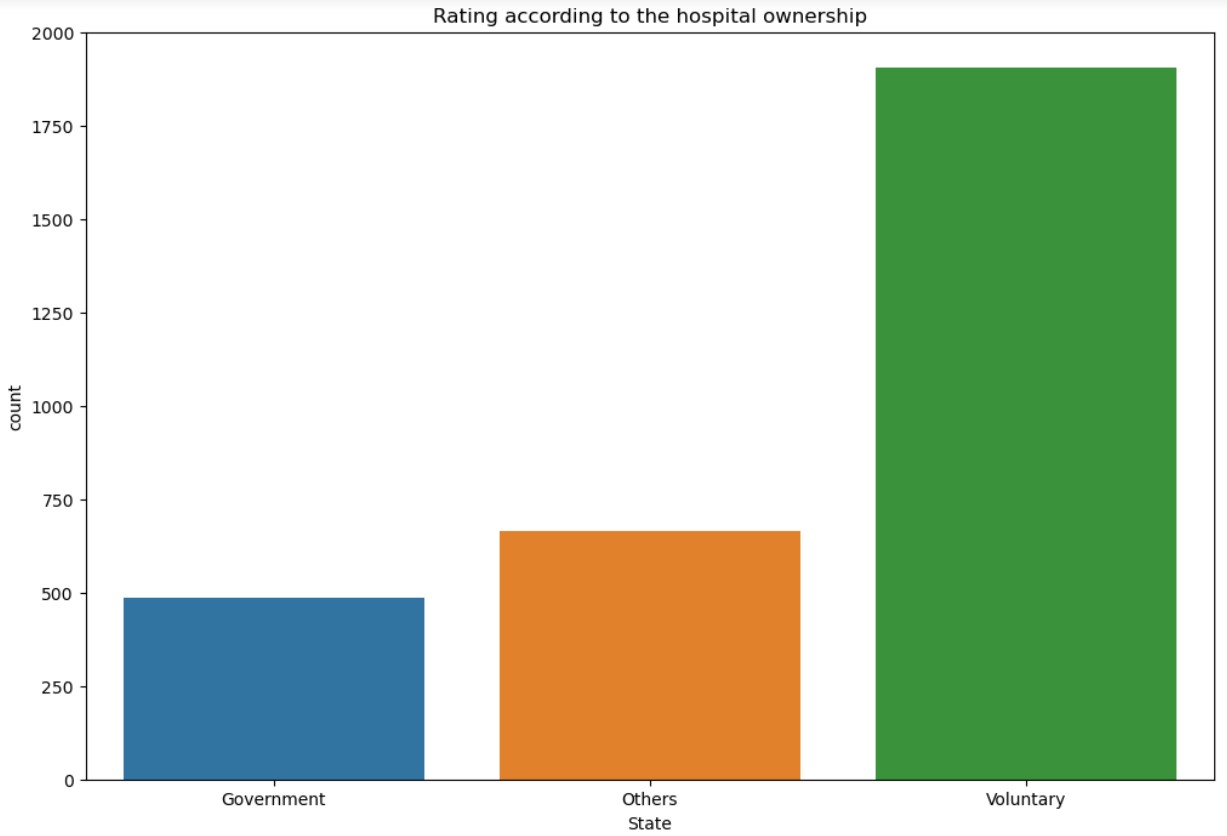
# **Executive Summary Introduction: Our project is all about the Centers for Medicare & Medicaid Services** is an agency that works in the U.S. Department of Health and Human Services (HHS) which administrates the major government health programs in the U.S. This includes insurance programs such as Medicare, Medicaid, and the Children’s Health Insurance Program (CHIP). CMS also plays a very important role in collecting and analyzing data with the goal of improving efficiency and equity in the healthcare system.

**Hospital Overall Rating Distribution: The distribution of hospital overall ratings was visualized using a count plot. This gives an overview of how ratings are distributed among hospitals.**



# **In the above count plot chart we can clearly say that the around 1400 of the hospitals overall rating is 3 and Top high rated hospital whose overall rating is 5 is somewhere around 100.**



From the above chart we can clearly say that the most of the hospitals are Voluntary.

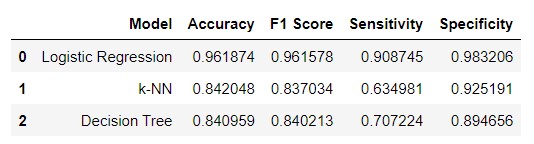
### **Building machine learning models:**

##### 1-So in the machine learning model firstly we removed the unnecessary columns from the original data set for machine learning.

##### 2- Our second task was to mapping the ratings so we use the lambda mapping function in which we did 1,2,3 are rated as 0 and 4,5 will be rated as 1.

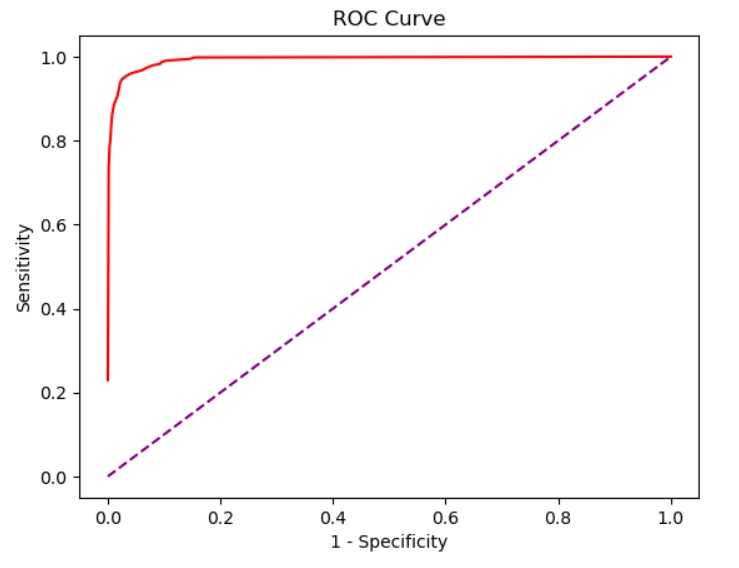
##### 3-Data preparation and Train-test-split so here our target variable is Hospital overall rating from the target variable and train test we have find the 3- models which are LOGISTIC REGRESSION , DECISION TREE and KNN model by which we can find that which model performs better and we have find the Accuracy , F-1 score , Specificity , Sensitivity.

##### From below fig. we can easily find the insights:



So, from the above figure we can easily find that the Logistic Regression F-1 Score performs better according to the other models.

We have performed the ROC / AUC curve for all the models but here I only share the ROC curve of LOGISTIC REGRESSION because this model performs better according to other models .



**Optimal-Cutoff Value for all the models** :

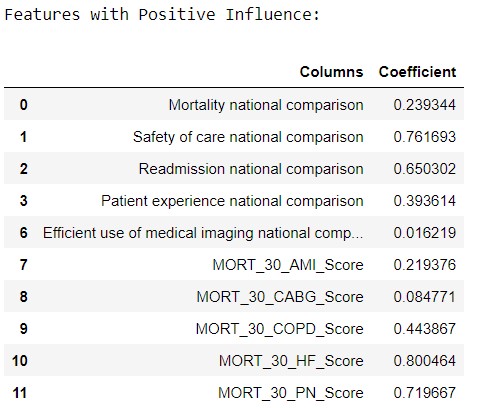
1- Optimal Cutoff value for logistic regression : 0.25

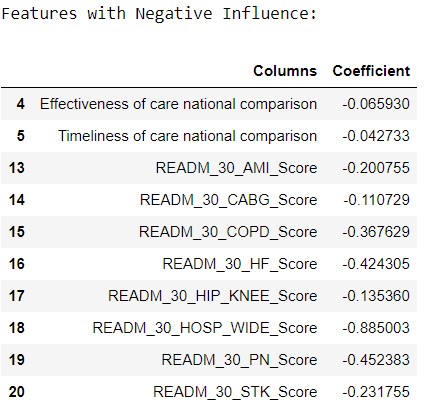
2- Optimal Cutoff value for KNN : 0.40

3- Optimal Cutoff value for Decision Tree : 1.00

### **Approach to identify areas of improvement:**

Logistic regression coefficients are used to identify that which columns are Positively influenced and which are negatively influenced with hospital overall rating.





**Conclusion:**

Hospitals with lower ratings must awareness on improving the identified influential capabilities. Hospitals should work on the cleanliness department because hospital should be very Hygienic for the patients and for their relatives. Low Rating hospital should work on for not include the external charges on the discharge bill. Most important in those hospitals their should be availability of doctors should be their if doctors are not visiting properly in hospital then the patients will go for different hospital this will also Lower down the rating.