Executive Summary

EDA

1) The ratings are distributed normally across all the hospitals, i.e., 1425 hospitals have 3 stars, 656, 749, 117 & 110 hospitals have 2, 4, 1 & 5 stars respectively.
2) The feature that has the most correlation with Hospital overall ratings in the data set is the Patient experience at 0.59, followed up closely by Readmission & Safety of care at 0.5 & 0.42 respectively.
3) In the pivot table of state-wise distribution of ratings categorised by hospital's ownership, there are many null values, which indicates that there is no hospital of those ownership types in those particular states, like, DE, VT, DC, PR, OR, ME, MT, MN & NH.
4) PA, SD, & UT have highest average rating of government hospitals, and SD, WE & DE have highest rated Voluntary hospitals.
5) Government & Others owned hospitals generally have higher average ratings than Voluntary owned hospitals.

Predictions

Using the best model (Logistic Regression) obtained from the data, the predicted ratings for the new hospitals are as follows:

Provider ID	Ownership	Ratings
520139	Voluntary	1
520189	Government	1
370029	Government	0
370032	Others	0
370036	Government	0
370037	Voluntary	0

Where 1 means high ratings of 4 & 5, and 0 means lower ratings of 1, 2 & 3.

Suggestions:

1) The parameter that affects the hospitals ratings by the greatest extent is 'rating_group'. So, the hospitals with 0 ratings should focus on increasing their scores in this category.
2) Another parameter that influences hospitals largely, but negatively, is 'READM_30_HOSP_WIDE_Score', so hospitals with lower ratings should focus on decreasing their score in this particular category.