CMS Rating Case Study

Problem Statement

- CMS rates the hospitals on a scale of 1-5 which impacts the patients choice of hospitals.
- Create supervised and unsupervised models to understand the methodology followed by CMS to star rate the hospitals and give customized recommendations to a particular hospital to improve it’s star rating in the coming cycles.

Understanding the data

- There are 7 groups and 64 measure which are used by CMS to provide ratings.
- Weightage of groups towards the overall star rating
  - 22% - Mortality, Readmission, Safety of Care, Patient Experience
  - 4% - Timeliness of care, Effectiveness of care, Medical Imaging Efficiency
- Data for these measures is spread across various file which is gathered based on the instruction provided by CMS.

Handling missing data

- Around 50% of the data consists of missing values which have been imputed by the instructions provided by CMS in the Hospital File.
- Each null value has been imputed with mean and median of that measure as instructed in the Hospital File by CMS.

Preparing Data

- Direction of each measure with the star rating has been recorded and then standardized to positive direction for all the measures and groups.
- Since the data is on different scales, it had been standardized using standard scaler.

EDA

- More than 40% of the hospitals were rated 3 by CMS.
- 2 and 4 were given moderately and 1 and 5 were very rarely give (3.5%)

CMS Rating	Count of providers
1	117
2	659
3	1426
4	749
5	110
Not Available	321

Hospital Ownership:

State and local Government and Tribal hospitals have a low average rating while non-profit hospitals have high Rating.

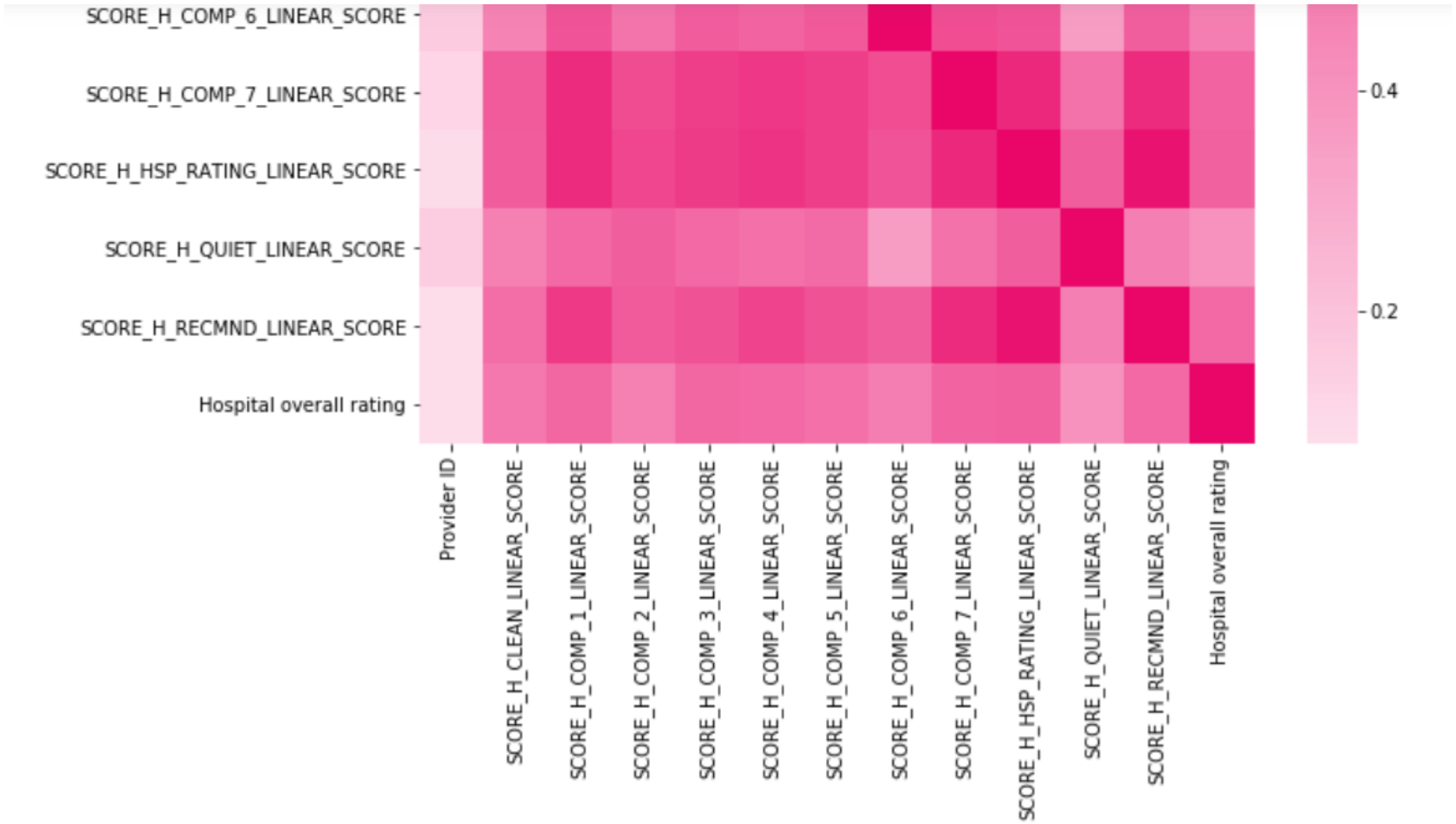
Emergency Services:

Non-Emergency Services Hospitals have a higher average rating

Importance of Groups by Impact

Patient Experience > Readmission > Safety > Mortality > Timeliness > Effectiveness > Use of Medical Imaging

•Patient Experience Group is highly important and correlated with the star rating.



Top Measures

- READM_30_HOSP_WIDE_score
- MORT_30_AMI_score
- HAI_2_SIR
- MORT_30_PN_score
- MORT_30_CABG_score
- H_COMP_1_LINEAR_SCORE_mean
- HAI_1_SIR
- MORT_30_STK_score
- ED_1b_Score

- HAI_3_SIR, READM_30_HF_score
- READM_30_PN_score
- H_COMP_7_LINEAR_SCORE_mean, HAI_4_SIR
- OP_10, H_COMP_3_LINEAR_SCORE_mean
- PSI_4_SURG_COMP_Score
- H_COMP_4_LINEAR_SCORE_mean
- READM_30_AMI_score
- H_HSP_RATING_LINEAR_SCORE_mean

Model Building

Random Forest:

- Random Forest predicts star ratings with an overall accuracy of approx. 78%.
- The error between predicted and estimated value is between -1 and +1 and never between that.
- The accuracy in predicting each class is minimum 65%.

Clustering:

- Weights for each measure was calculated using factor analysis in R and then the Group Score was computed.
- The group score was then used to calculate the Final score using group weights.
- The accuracy achieved by K-Means clustering is approx. 54%.
- Class wise minimum accuracy is 50%.

Recommendation for Hospitals

- Recommendation to improve star rating from 3 to 4 for Provider ID – 140010 –
  - Safety and Patient Experience scores are below the mean scores for rating 4.
- Measures to improve -
  - HAI_4: Surgical Site Infection from abdominal hysterectomy (SSI: Hysterectomy)
  - HAI_5: Methicillin-resistant Staphylococcus Aureus (MRSA) Blood Laboratory-identified Events (Bloodstream infections)
  - H_HSP_RATING_LINEAR_SCORE: Patients who gave their hospital a rating of 9 or 10
  - H_RECMND_LINEAR_SCORE: Patients who reported they would recommend the hospital
  - H_CLEAN_LINEAR_SCORE: Patients who reported that their room and bathroom were "Always" clean
- By improving the scores for above measure the overall star rating for Provider ID – 140010 can be improved to atleast 4.