

Capstone_Project _Healthcare

Loading the necessary libraries to perform the analysis using Python Jupyter notebook

Received Data from CMS

deleting the unwanted columns & keeping the measure
provider id, score, hospital rating

finally merging all the datasets into final data for further analysis

Identifying the measures

performing the EDA

heatmap

splitting the data into training & testing

feature scaling by
standardscaler

Linear Regression

using statsmodel

adding constant, fitting(ols) the model with rfe, vif, adjusting manually

residual analysis on training data, got normal curve

got an accuracy of 37%

Logistic regression

got accuracy 67%

PCA & Clustering

Outlier treatment

Hopkins statistics

achieved 0.98

K-Means

Silhouette analysis

no. of cluster is 3

recommendation for Evanston Hospital

Validating the measures