Executive Summary

- This particular project **"Starter code for the Hospital Classification for the Capstone Project"** consisted of 2 datasets,**'hospital-info.csv'** & **'not_yet_rated.csv'** and we performed 3 machine learning models for each of them and found the best model for the same.

- First, we analyzed the **'hospital-info.csv**' dataset and obtained our 3 machine learning models particularly by 3 methods,
  1. **Linear regression**
  2. **Decision Tree**
  3. **KNN Nearest Neighbors.**

- As we calculated the **RMSE's, Predictions & Accuracies** of the following models & compared them; we came to the conclusion that; **Linear Regression model** is the **best** machine learning model among the 3 models.

- We also performed **regularization and hyperparameter tuning techniques** to improve your model performance for each of the models and again the result was same.

- Then we analyzed another dataset from this project which was **'not_yet_rated.csv'**, and selected the categorical & numerical columns accordingly.

- Again, we repeated the same steps which we had done in our first dataset of making 3 machine learning models by doing the execution with same methods.

- By calculating & comparing all the accuracies & predictions, we came to the conclusion that **Linear Regression model** is the **best** model among the 3. We also performed regularization and hyperparameter tuning techniques.

- The additional part we did here is we identified the measures which have a **positive influence** as well as **negative influence** on the **overall hospital ratings**. According to that we **checked the coefficients** and visualized the data by means of **Histogram.**

- We also identified in which of the **measures** a low-rated hospital is currently **lagging behind**, because these measures need to be **improved**. We have showed the **visualization** respectively.