

REBECA PRACTICE: DATA SCIENTIST SOLUTIONS

SOLUTION TO TASK 5

```
In [ ]: # We create a Dataframe for the prediction
stakeholder_movie_features = pd.DataFrame([[.5e8, 6.8, 3e4]], columns = X.columns)
stakeholder_movie_features

In [ ]: stakeholder_movie_prediction = reg_all.predict(stakeholder_movie_features)
print(stakeholder_movie_prediction)

In [ ]: # To check the effect of a 10% increase in the features,
# We create a new DataFrame
stakeholder_movie_feature_changes = pd.DataFrame([[.5e8, 6.8, 3e4],
                                                    [.5e8, 1.1*6.8, 3e4],
                                                    [.5e8, 6.8, 1.1*3e4],
                                                    [.5e8*1.1, 6.8, 3e4],
                                                    [.5e8*1.1, 6.8*1.1, 3e4*1.1]],
                                                    columns = X.columns)

stakeholder_movie_feature_changes

In [ ]: # Calculate prediction for 10% increase in Rating, Votes, Budget and all three at once
stakeholder_movie_prediction_changes = reg_all.predict(stakeholder_movie_feature_changes)
print('relative increases in revenue in %: ',
      100*(stakeholder_movie_prediction_changes[1:]/stakeholder_movie_prediction-1))
```