Classwork - Exploring Linear Regression

Instructions: This classwork should be done independently. Use your jupyter notebook from yesterday but you do not need to turn in your notebook. Write your answers below and upload to the folder in the drive.

1. Using your notebook from yesterday - calculate the MAE, MSE and RMSE using the functions from scikit learn. Explain what the r-squared score (that you obtained yesterday) and the MAE value and the RMSE value are telling us about the model.

The r-squared score of .82 showing that 82% of the features correlate to the answer. The closer this score is to 1, the better the correlation is.

MAE = .38 , MSE = .22 , RMSE = .47

MAE is the sum of absolute errors / number of predictions.

RMSE is the square root of (the square of absolute errors / number of predicitons).

RMSE puts more weight on larger errors. Since we try to avoid large errors with linear regression – this is the preferred metric.

1. Explain what the y-intercept and coefficient values you obtained yesterday are telling us. Is the y-intercept value meaningful in this case? Why or why not?

y-intercept tells what the score would be if all the factors are zero. This is not meaningful with the happiness dataframe as it does not make sense that all the factors would be zero.

Coefficient values are the slope. They tell how much the score will increase when that specific factor changes.

Hatz