

# Linear Regression evaluation metrics

## $R^2$

- Corresponds to explained variance.  $R^2 = 0.6$  means that the model accounts for 60% of the variance
- low  $R^2 \neq$  bad model  $\rightarrow$  it's domain dependent
- $R^2 = \text{residuals from model} / \text{residuals from mean}$
- literally a measure of how much better the model is at predicting results compared to predicting the average for every datapoint.
- adjusted  $R^2$  - accounts for adding independent variables
  - $R^2$  will always increase when you add feature variables

## MSE

- average of the square of the residuals
- sensitive to outliers

## RMSE

- square root of MSE  $\rightarrow$  easier to interpret b/c it will have the same units as the dependent variable

## MAE

- absolute average (manhattan) distance between prediction and results
- less sensitive to outliers compared to MSE

# References

- <https://corporatefinanceinstitute.com/resources/data-science/r-squared/>
- <https://statisticsbyjim.com/regression/interpret-r-squared-regression/>
- <https://medium.com/analytics-vidhya/mae-mse-rmse-coefficient-of-determination-adjusted-r-squared-which-metric-is-better-cd0326a5697e>