Variance:

\_\_basically, tells you how much the function is capable of adjusting to the change in the dataset.

\_\_A few reasons for High Variance:

* Over-fitting
* Low error on training data, but high error on test data
* Complex model and spares training data

Bias:

\_\_Bias of the estimated function tells you the capacity of the underlying model to predict the values.

\_\_For example, simpler models, in general, fail to capture the complexity of high dimensional data and hence they have higher bias.

\_\_A few reasons for high bias:

* Under fitting
* High error on both the training and test data
* Simplified model

Bias variance graphic:

\_\_Uses a bulls-eye diagram to illustrate four possible outcomes of high, low bias and high, low variance.

\_\_A hit at the center of the target means that the method perfectly predicts the correct value.

\_\_Each hit represents an individual realization of the method, given the chance variability in the training data

\_\_ Sometimes we will get a good distribution of training data so we predict very well and we are close to the bulls-eye, while sometimes our training data might be full of outliers or non-standard values resulting in poorer predictions.

\_\_The High bias low variance target shows a underfitting situation

\_\_On the other hand, Low bias, but high variance shows a overfitting situation

Trade-off

\_\_For example, add more parameters to the function, the complexity of the function increase as well as the variance

Example:

\_\_Polynomial of degree 1 is too simple to capture the sine curve, which might achieve higher bias.

\_\_However, polynomial of degree 11 is complex enough that it even captures the noise

\_\_It can be seen that Polynomial of degree 7 give us the best fit to our sine curve in the presence of noise.

Last slide:

\_\_Find a good balance between bias and variance such that it minimizes the total error

\_\_An optimal balance of bias and variance would never overfit or underfit the model