2: Statistical Learning

2 main purposes of statistical learning: Inference, and/or prediction

Reducible errors exist in the models we wish to fit the data, irreducible errors lie in variance associated with the error terms

- Inference
  - Which predictors associated with response?
  - what is the relationship between response and each predictor?
  - can the relationship between Y and each predictor be adequately using a particular model? or is it more complex
- Parametric Methods
  - make assumption of functional form
  - use training data to fit/train the model (OLS)
- Non-Parametric Methods
  - do not make explicit assumptions about functional form
  - estimate f that gets as close to the data points without being too rough or wiggly
  - usually require large amounts of data

Generally want model to minimize test MSE to determine if our model is accurate

- Bias-Variance Trade-Off
  - variance refers to amount by which f would change if estimated using different training set
  - bias refers to error introduced by approximated the problem to a particular model