Modeling X

Cross Validation Assignment

- Instructions
  - Download Analysis 4 Zip Folder
  - Unzip Folder
  - Open Analysis 4 Rmd File
  - Knit to HTML
  - Read Introduction
- Three Part Assignment
  - Part 1: Build Many Models
  - Part 2: Identify Top Models
  - Part 3: Visualize Best Model
- Remember
  - Do Assignment in Order
  - Change “eval=F” to “eval=T”
  - Frequently Knit

Part 1: CV for RMSE

- Goals
  - Build Many Polynomial Models
  - Use 10-fold CV to Get RMSE
  - Build Matrix to Save Info
- Things to Consider
  - Look at Tutorial from Link
  - Most Coding Has Been Done
  - Minor Edits Must Be Made
  - Everything Else Builds Off This

Part 2: The Top 5

- Goals
  - Find the 5 Best Models Based on RMSE
  - Plot Predictions Under These 5 Top Models on Same Graph Over the Raw Data
  - Plot the Effect That Model Complexity has on RMSE
- Things to Consider
  - Different Models Lead to Different Prediction Accuracy
  - We Want the “Smallest” Model That Adequately Predicts

Part 3: The GOAT

- Goals
  - Function to Output Best Model
  - Plot Predicted vs Actual
  - Plot Residuals vs Input Variable
- Things to Consider
  - Function Should Output a Numeric Vector
  - This Vector Can Be Saved and Used to Obtain Predictions and Residuals
  - Carefully Read Instructions

Closing

Disperse and Make Reasonable Decisions