

Validation Set Approach

Example: Autmobile Data (ISLR2)

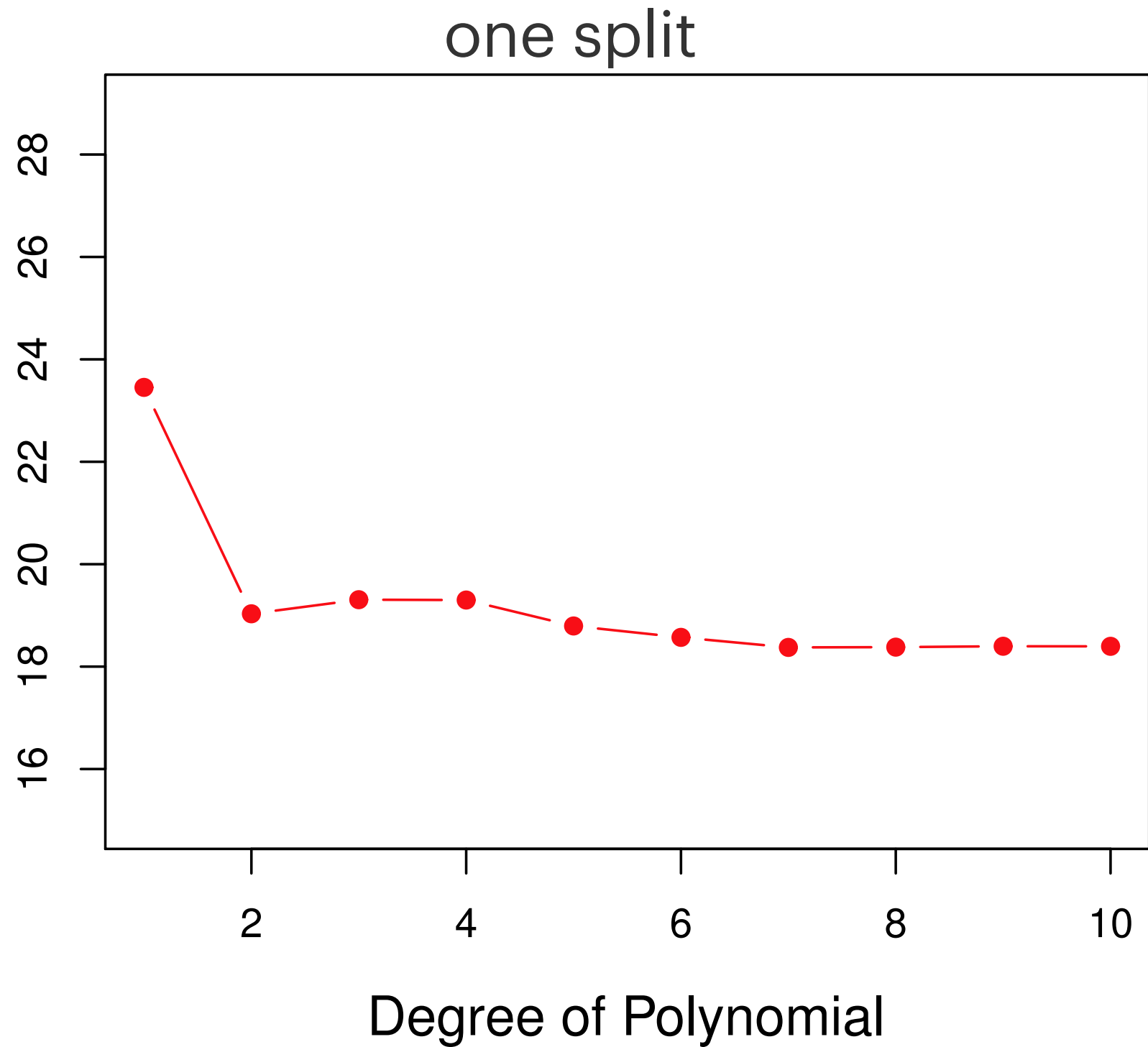
• comparing linear vs higher-order polynomial terms in a linear regression

• $y = \text{gas miles per gallon}$, $x = \text{horsepower}$

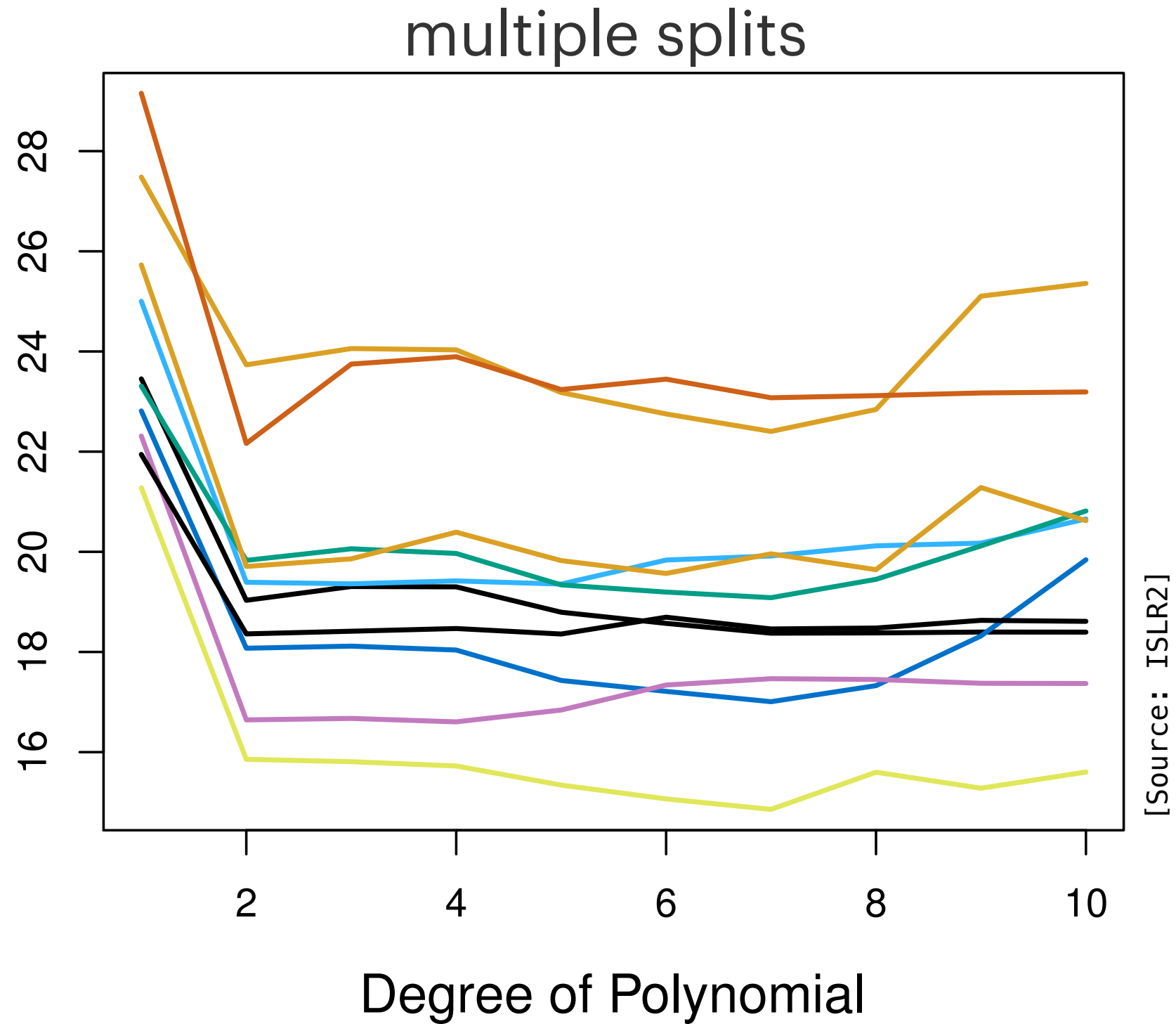
• randomly split the 392 observations into two sets:

training and validation set of 196 servers each

Mean Squared Error



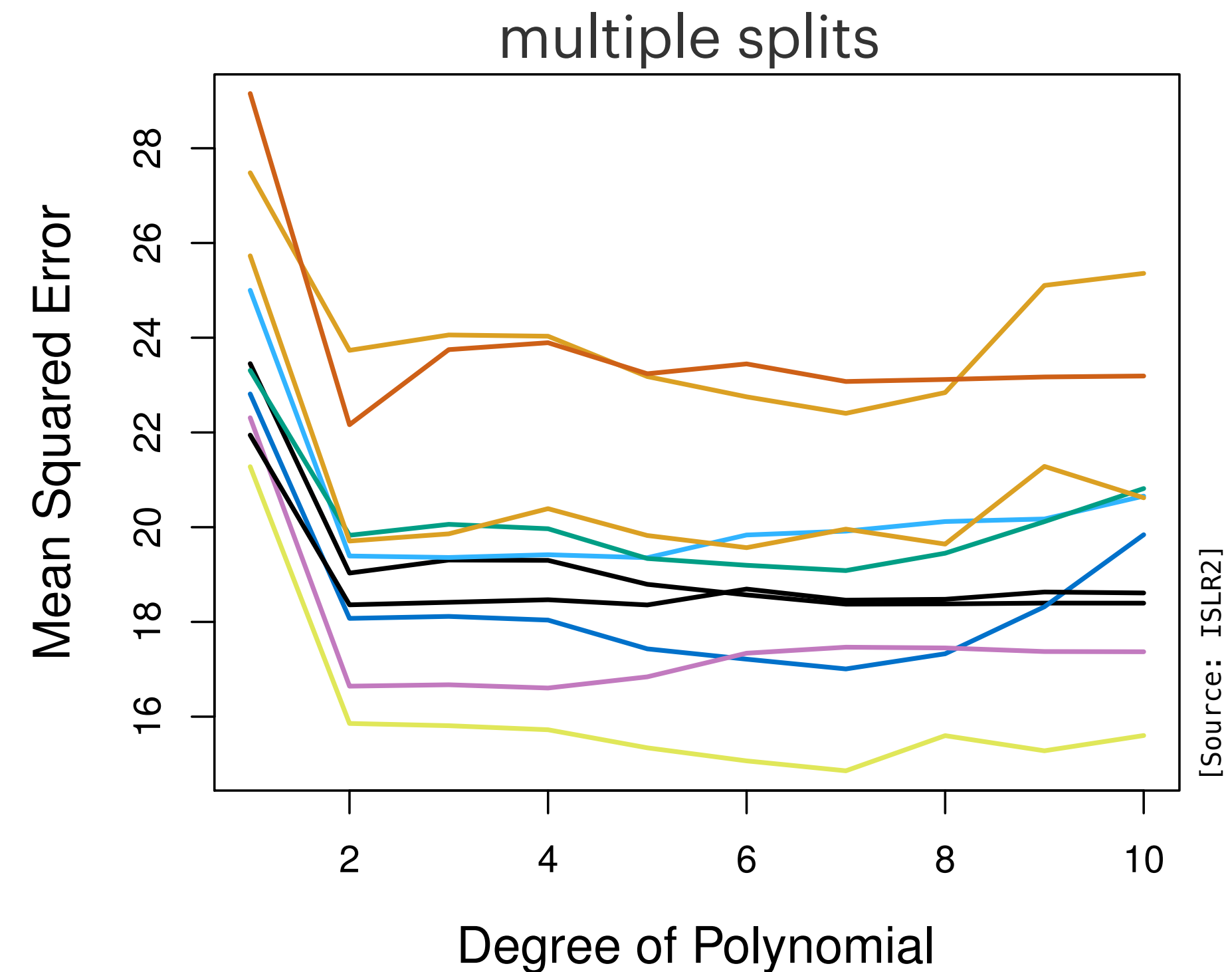
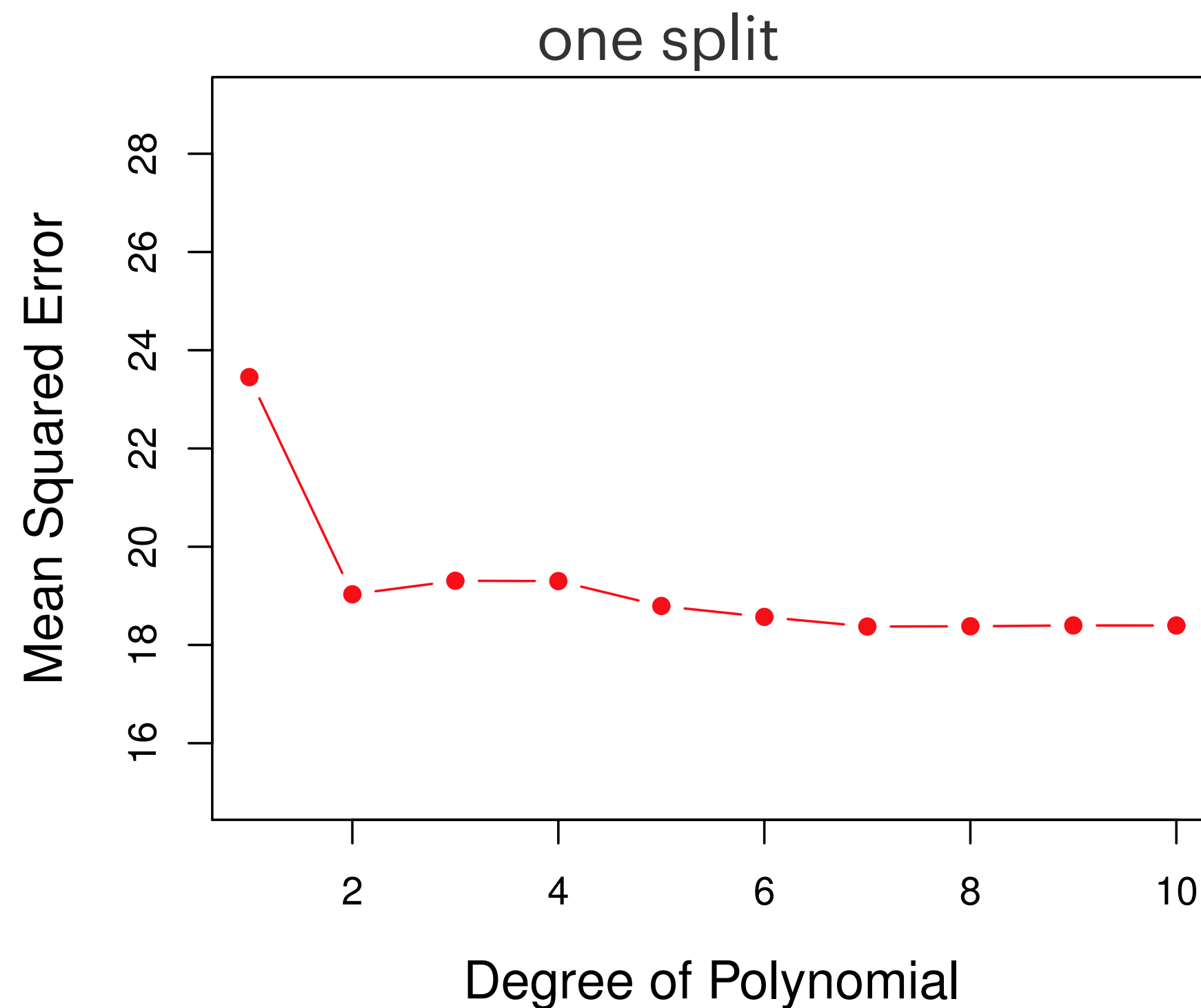
Mean Squared Error



Validation Set Approach

Example: Automobile Data (ISLR2)

- compare linear vs higher-order polynomial terms in a linear regression
- y = gas mileage in miles per gallon, x = horsepower
- randomly split the 392 observations in to two sets: training and validation set of 196 observations each



“If model selection and true error estimates are to be computed simultaneously, the data should be divided into three disjoint sets”

[Brian D. Ripley, 1996]