Section 6.2: Linear Regression

Duration: 2.5 hours

Concepts:

- Simple linear regression
- Multiple linear regression
- Potential problems
- Performing linear regression in R

Description: Introduction to linear regression with emphasis on interpretation of fitted models and implementation in R.

Materials: Section 6.2 slide deck; linear regression in-class exercises.

Textbook section: An Introduction to Statistical Learning, Chapter 3.

Duration	Lesson Section	Learning Objectives
30 mins	Go through the Simple Linear Regression section of the slide deck.	● Parameter estimation with least squares ● Assessing accuracy of parameter estimates ● Hypothesis testing for coefficients ● Checking quality of model fit with R^2 and RSE
15 mins	Go through the “Simple Linear Regression” section in the R Markdown file as a class.	● Use `lm()` to fit a linear model to the `Boston` data set. ● Interpret `summary()` results ● Plot model
20 mins	Go through the multiple linear regression section of the slide deck.	● Estimating and assessing fit of parameters and model ● Types of errors associated with predictions ● Prediction intervals
10 mins	Go through the “Multiple Linear Regression” section in the R Markdown file as a class.	● Use `lm()` to fit a multiple linear regression ● Interpret model results. ● Use `predict()` to find a prediction interval
15 mins	Go through the “Qualitative Predictors” section of the slide deck.	● Qualitative predictors ● Dummy variables ● Two or more levels
8 mins	Go through the “Qualitative Predictors” section in the R Markdown file as a class.	● Use `lm()` on qualitative predictor ● Interpret the coefficient estimates
10 mins	Go through the linear model extensions section of the slide deck.	● Removing the additive assumption ● Interaction effect ● Hierarchical principal

5 mins	Go through the “Interaction Term” section in the R Markdown file as a class.	● Use `lm()` with an interaction term
30 mins	Go through the remaining sections of the slide deck.	● Non-linear response-predictor relationship ● Correlation of error terms ● Non-constant variance of error terms ● Outliers ● High-leverage points
20 mins	Go through the “Helpful Plots” section in the R Markdown file as a class. Students should be given time to finish the question at the end.	● How to use plots to examine data and fit ● Plot studentized residuals ● Cumulative exercise