1) Data exploration

- Variation across categories
- Variation of numerical variables
- Quantiles, coverage intervals, z scores
- Variation across groups
- Scatter plots and correlation
- Lattice plots
- Principles of good/bad statistical graphics

2) Fitting equations to data

- Fitting straight lines by OLS
- Coefficients, fitted values, residuals
- Plug-in prediction
- Summarizing the trend
- Statistical adjustment
- Summarizing the information in a predictor
- Polynomial models
- Power laws and exponential growth/ decay

3) Predictable and unpredictable variation

- Prediction intervals
- R^2 and its interpretation

4) Grouping variables in regression

- Aggregation paradoxes and confounding
- Dummy variables
- Handling multiple grouping variables
- Main effects vs. interactions
- The analysis of variance
- Numerical and grouping variables together
- Collinearity among predictors

5) Quantifying uncertainty

Sampling distribution
Standard errors
Confidence intervals
Frequentist coverage property
Bootstrapping
Normal linear regression model

6) Multiple regression

Regression planes
Partial relationships vs. overall relationships
Grouping variables in multiple regression
Applications of multiple regression
Statistical significance from confidence intervals
ANOVA in multiple regression models

7) Hypothesis testing

Four steps of hypothesis testing
Null hypothesis, test statistic
Permutation tests: basic idea
Permutation tests in multiple regression
p values

8) Building a predictive model

Variable selection (why its hard)
MSPE/RMSPE
train/test splits
the problem with MSPE_in stepwise selection