Week 3 Suggested Readings and Exercises

Suggested readings and practice problems from OpenIntro Statistics, 3rd edition (a free online introductory statistics textbook co-authored by Dr. Cetinkaya-Rundel) for this week:

Suggested reading: Chapter 8, Section 8.1 - 8.3

Suggested exercises: (End of chapter exercises from OpenIntro Statistics)

- Regression with multiple predictors: 8.1, 8.3
- Inference for MLR: 8.5
- Model selection: 8.7, 8.9, 8.11
- Model diagnostics: 8.13

(Reminder: the solutions to the end of chapter exercises are at the end of the *OpenIntro Statistics* book)

Test yourself:

1. *How is multiple linear regression different than simple linear regression?*
2. *What does "all else held constant" mean in the interpretation of a slope coefficient in multiple linear regression?*
3. *What is collinearity? Why do we want to avoid collinearity in multiple regression models?*
4. *Explain the difference between $$R^2$$ and adjusted $$R^2$$. Which one will be higher? Which one tells us the variability in y explained by the model? Which one is a better measure of the strength of a linear regression model? Why?*
5. Define the term “parsimonious model".
6. Describe the backward-selection algorithm using adjusted R2 as the criterion for model selection.
7. If a residuals plot ($residuals$ vs. x or $residuals$ vs. $\hat{y}$) shows a fan shape, we worry about non-constant variability of residuals. What would the shape of these residuals look like if absolute value of residuals are plotted against a predictor or $\hat{y}$?