

AP Stats – Chap 8

Regression Wisdom

Chapter 7 briefly touched on What Can Go Wrong? with regression. Now we examine those issues more closely.

- THINK before rushing into regression.
- Plot the data and look for trouble.
- Design a plan to handle any problems.
- Divide data into subsets if there are multiple trends.
- Discuss the model in context.
- Don't extrapolate much – if at all.
- Always plot the residuals. What do you see?

a point with high **leverage** – its x-value is far from the mean of the other x-values

a point that is **influential** – omitting it from the data gives a very different model (equation)

Three Flavors of Unusual Points

1. points that confirm the pattern, but are extreme. They have **high leverage** but **small residuals**. (High leverage, but small influence.) Slope and intercept of model may not be affected much, but R^2 will increase...misleading reader to think that the association is stronger than it actually is.
2. outliers that don't have **high leverage**, but do have **large residuals**. The slope of the model may not change much, but the intercept usually does.
3. influential points are the **most troublesome!** They have **high leverage** AND are **not consistent** with the trend of the data. They are the hardest to detect because they have **small residuals**.