

STAT 113

Simple Linear Regression II

Colin Reimer Dawson

Oberlin College

September 26, 2017

Outline

Residuals

Transformations

The Regression Effect

An Experiment

You are about to see an array of symbols. You will have 15 seconds to memorize as many as you can.

An Experiment



An Experiment

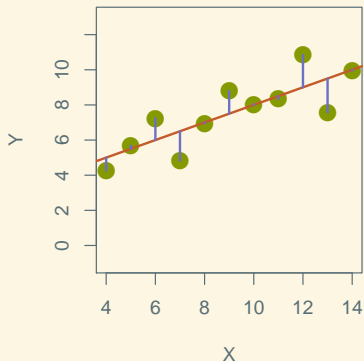
Now write down as many names as you can remember.

An Experiment

How many did you get?



Residuals



- Every line will have a different set of errors associated with it.
- These errors are called **residuals**.

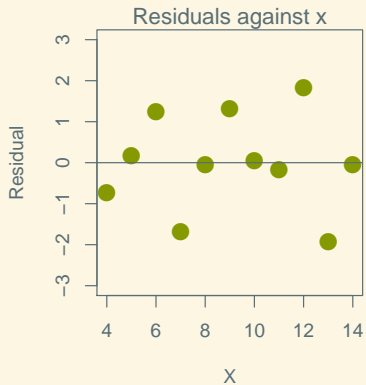
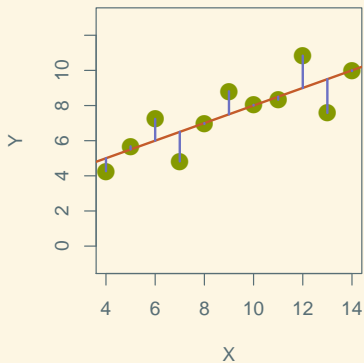
Randomness of Residuals

- If our regression model captures the relationship, we shouldn't be able to predict the residuals from anything else: they should be truly random.
- If the residuals are random, they should be unrelated to both x (explanatory variable) and \hat{y} (predicted values).

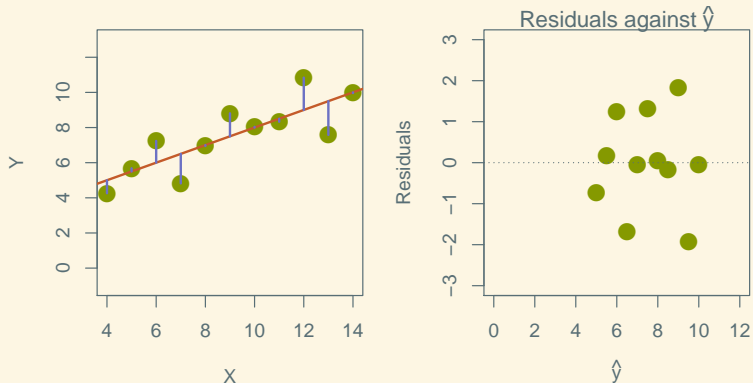
Residual Plots

- Two useful plots:
 1. Residuals against x (explanatory variable)
 2. Residuals against \hat{y} (predicted values)
- If residuals are random, these should both look like an unstructured “cloud”.

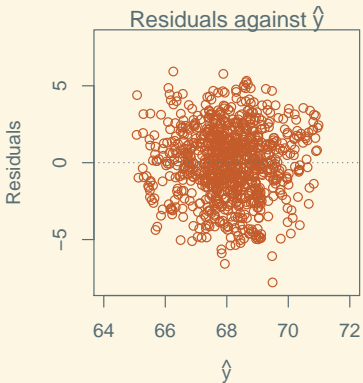
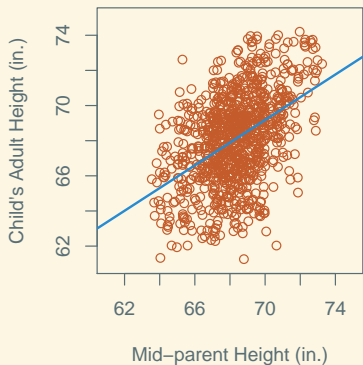
Residual Plots



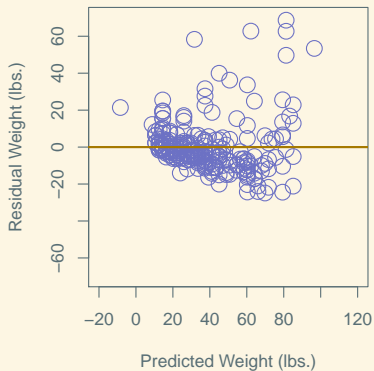
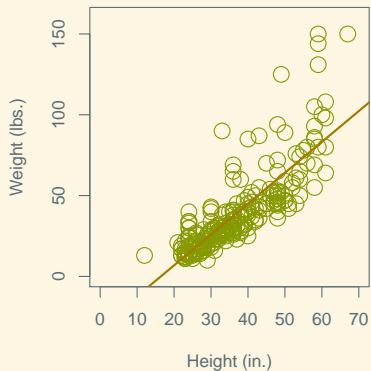
Residual Plots



Residual Plots



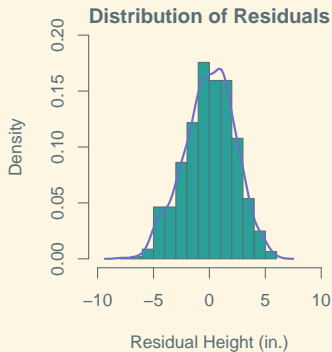
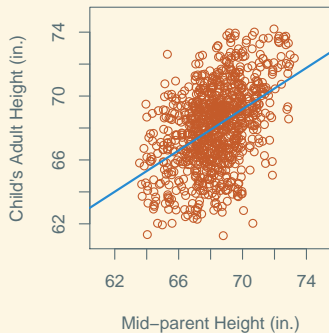
Nonlinear Residual Plots



The Residual Distribution

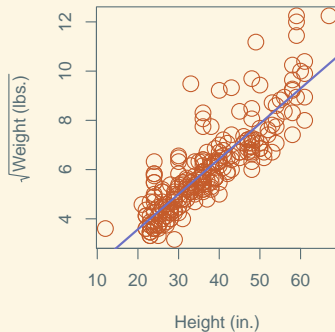
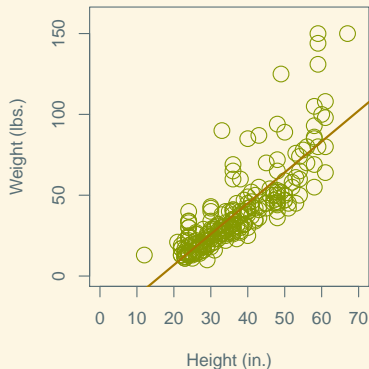
- It's also useful to look at the **distribution** of residuals.
- In many cases, if the residuals really just capture randomness, they will have a bell-shaped (Normal) distribution.
- We can apply our univariate techniques to the residual distribution to assess this.

The Residual Distribution

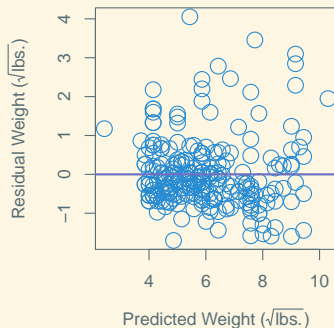
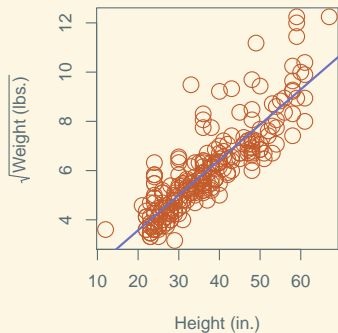


- We can use residual plots as a diagnostic tool, to see when a linear model is inadequate.
- Sometimes we need a more complex model (e.g., a higher order polynomial; one with other predictors)
- Sometimes we can create a linear relationship via a **transformation** of one or both variables.

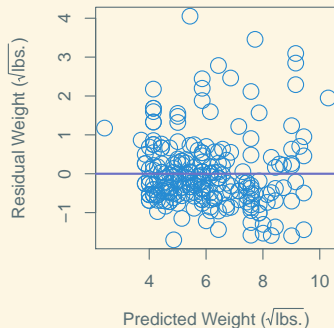
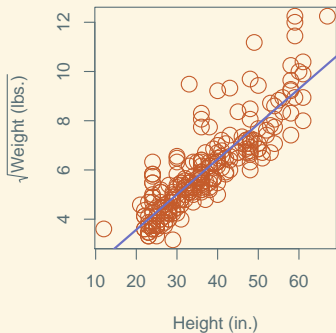
Example of a Nonlinear Relationship



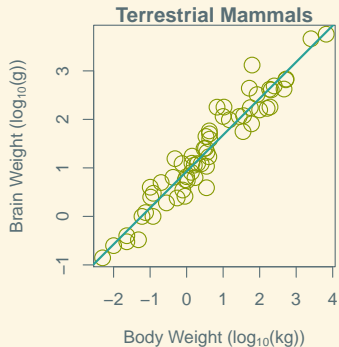
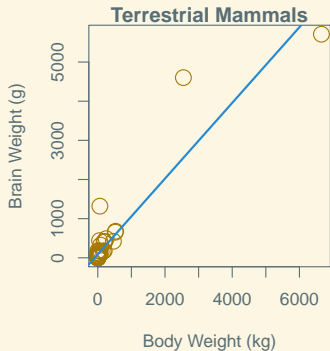
After Square Root Transformation



After Square Root Transformation



What About This?



An Experiment, Part 2

As before, focus on the pictures for 15 sec, and see how many items you can memorize.

An Experiment, Part 2



An Experiment, Part 2

Write down as many as you can

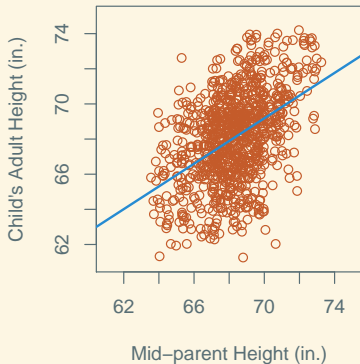
An Experiment, Part 2



Experiment Results

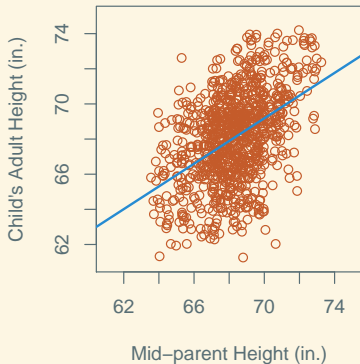
Record your results here: <https://goo.gl/gLFuac>

Regression Example



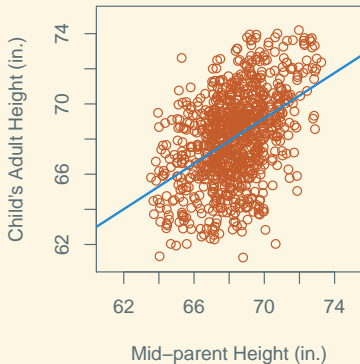
- The “father of regression”, Francis Galton, looked at parents’ and children’s heights.
- Here’s his data, with the associated regression line.

Regression Example



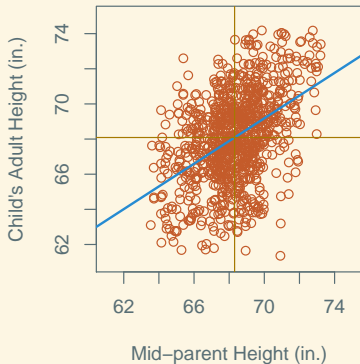
- We have
 $\hat{y} = 23.94 + 0.646x$.
- What does the 23.94 mean?
- What does the 0.646 mean?

Regression Example



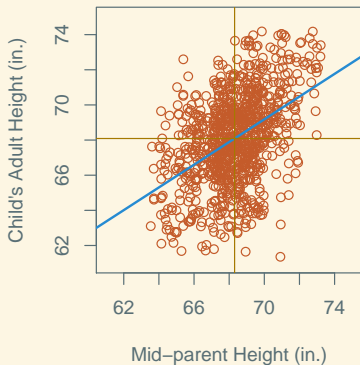
- What would you predict for a child whose parents' average height is 64 in.?
- How about if the parents' average height is 72 in.?
- What about for the average parents' height?

Regression to the Mean



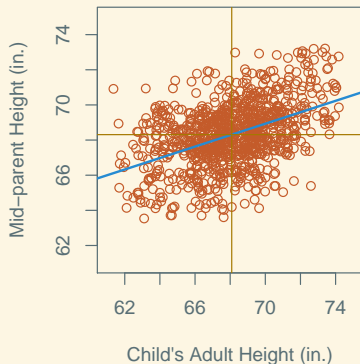
- Since the x and y have the same units, the slope less than 1 means that, on average, children are closer to average than their parents. Why?

Regression to the Mean



- Galton called this phenomenon “reversion to the mean”. Later changed to “regression to the mean”.
- This is the origin of the term “regression”.

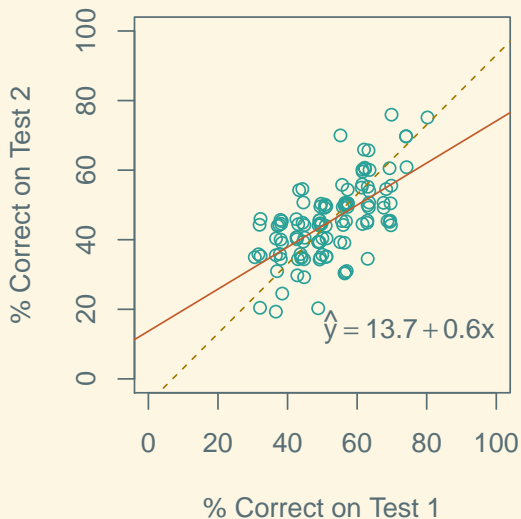
Regression to the Mean



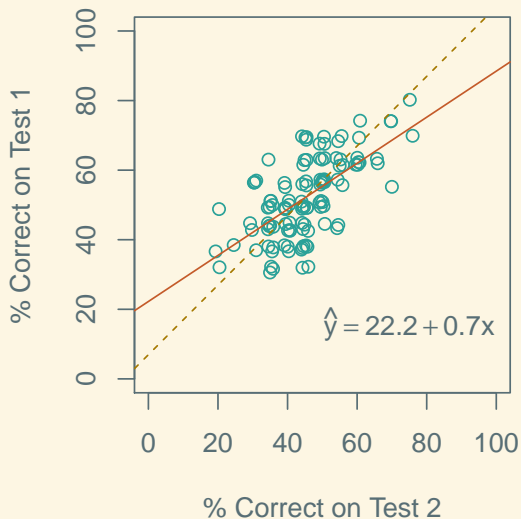
- If we try to “predict” parent from child, we get the following regression equation:

$$\hat{x}_i = 46.14 + 0.33y_i$$

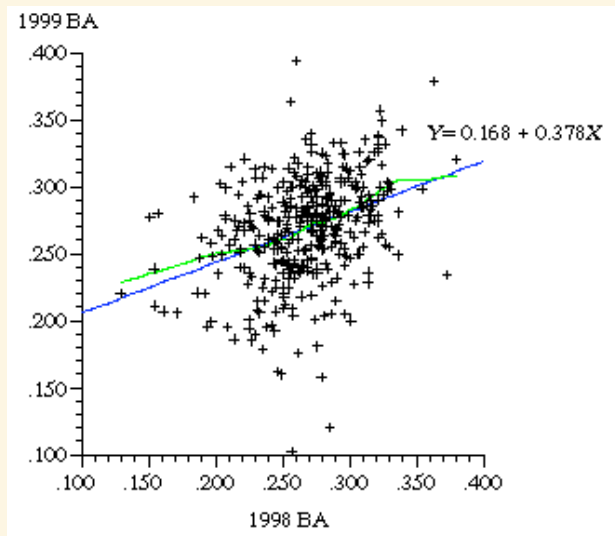
Memory Test Results





Note: points are “jittered” slightly to show multiple identical values. 35 / 40

Memory Test Results (X and Y Reversed)

Example: Batting Average in Successive Seasons



Why our children's future no longer looks so bright

  Text Size  Print  E-mail  Reprints

By Robert J. Samuelson, Published: October 16

Aspecter haunts America: downward mobility. Every generation, we believe, should live better than its predecessor. By and large, Americans still embrace that promise. A Pew survey earlier this year found that 48 percent of respondents felt that their children's living standards would exceed their own. Although that's down from 61 percent in 2002, it's on a par with the mid-1990s. But these expectations could be dashed. For young Americans, the future could be dimmer.

Along with jobs, the 2012 presidential election could be fought over this issue. "[Can the Middle Class Be Saved?](#)" worried a recent cover story in the Atlantic. Pessimism rises with schooling. In [the Pew poll](#), 54 percent of respondents with a high-school diploma or less felt their children would do better; only 35 percent of graduate school alums agreed. "A kind of depression has set in," [writes Washington Post columnist Richard Cohen](#). "We've lost our mojo, our groove."

Figure: Source:

https://www.washingtonpost.com/opinions/why-our-childrens-future-no-longer-looks-so-bright/2011/10/14/gIQAofzlpL_story.html

“Boomerang” CEOs

“This fall, Lafley will step down for the second time, and no one will be mentioning Steve Jobs’s legendary return to Apple. Lafley hasn’t been bad – he slimmed the company down, selling off parts and getting out of less profitable businesses – but there’s been no dramatic turnaround. ... In other words, he’s been just O.K.

How could someone who, according to Fortune, was known as “an all-time C.E.O. hero” end up being just O.K.? Well, if commentators had looked at the track record of returning C.E.O.s – boomerang C.E.O.s, as they’re sometimes called – that’s precisely what they’d have predicted. A 2014 study found that profitability at companies run by boomerang C.E.O.s fell slightly, and an earlier study detected no significant difference in long-term performance between firms that reappointed a former C.E.O. and ones that hired someone new.”

Israeli Air Force Pilot Training

Trainee pilots received praise following successful execution of flight maneuvers, and criticism after unsuccessful execution. Instructors found that “high praise for good execution of complex maneuvers typically results in a decrement of performance on the next try”¹.

¹Kahneman, D. and Tversky, A. (1973), “On the Psychology of Prediction,”
Psych. Rev. 80:237-251