What is regression?

Karl W Broman

Department of Biostatistics
Johns Hopkins Bloomberg School of Public Health

http://www.biostat.jhsph.edu/~kbroman
How tall will Caleb be?

Heights of sons


average = 69.2 in.
Heights of fathers and sons

average = 68.4 in.

Heights of fathers and sons
Heights of fathers and sons

average = 69.5 in.
Heights of fathers and sons

Heights of fathers and sons
Heights of fathers and sons

Heights of fathers and sons
Heights of fathers and sons

Heights of fathers and sons

Father's height (in.)
Son's height (in.)
Heights of fathers and sons

Heights of fathers and sons
Heights of fathers and sons

Heights of fathers and sons
Heights of fathers and sons

Summary

- Regression concerns predicting $Y$ from $X$.
- There are two regression lines.
- The regression effect:
  - Tall fathers, on average, have sons who are not so tall.
  - Short fathers, on ave., have sons who are not so short.
- The regression fallacy: assigning some deeper (causal) meaning to the regression effect.