



What is regression?

Karl W Broman

Department of Biostatistics
Johns Hopkins Bloomberg School of Public Health

<http://www.biostat.jhsph.edu/~kbroman>



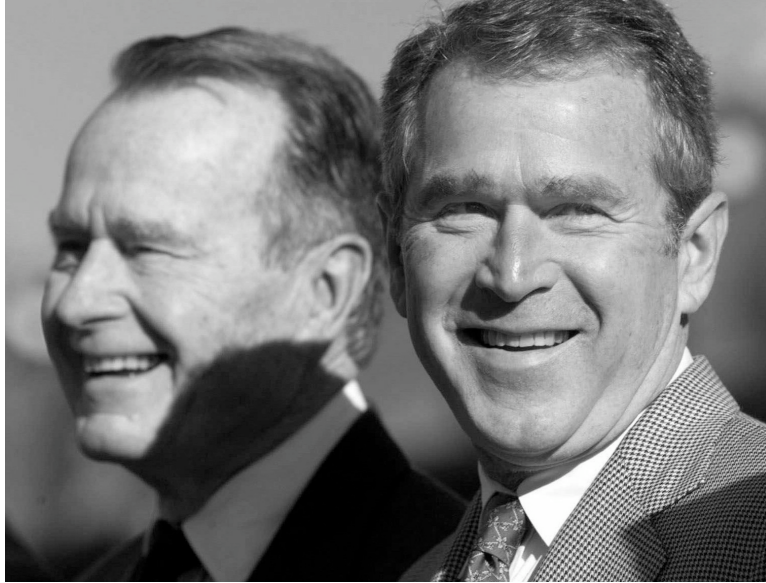
2



3



4



5



6



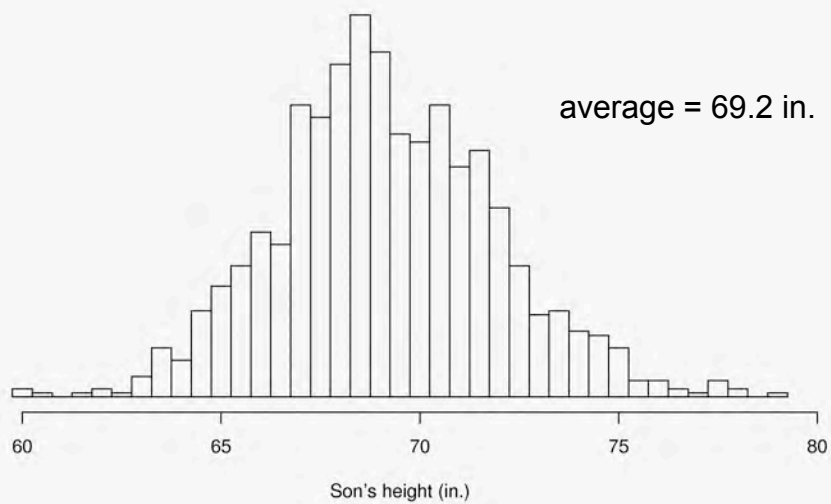
How tall will Caleb be?

7



Heights of sons

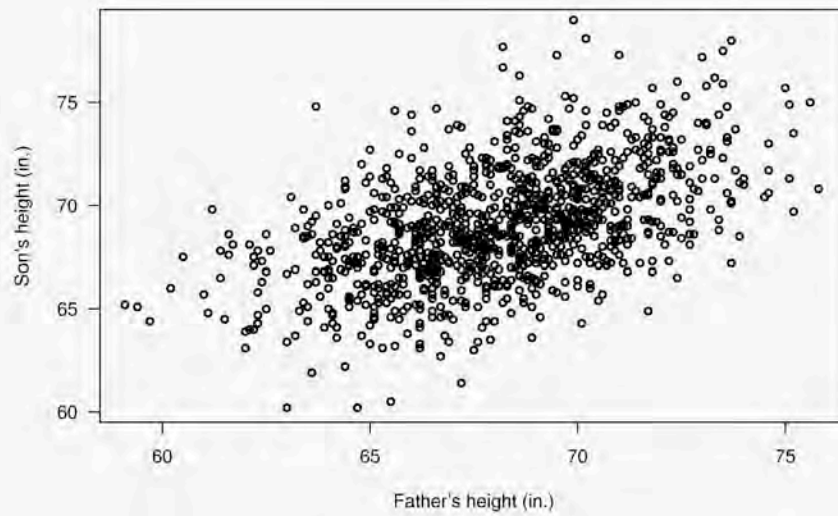
Pearson & Lee, Biometrika 2:357-462, 1906



8



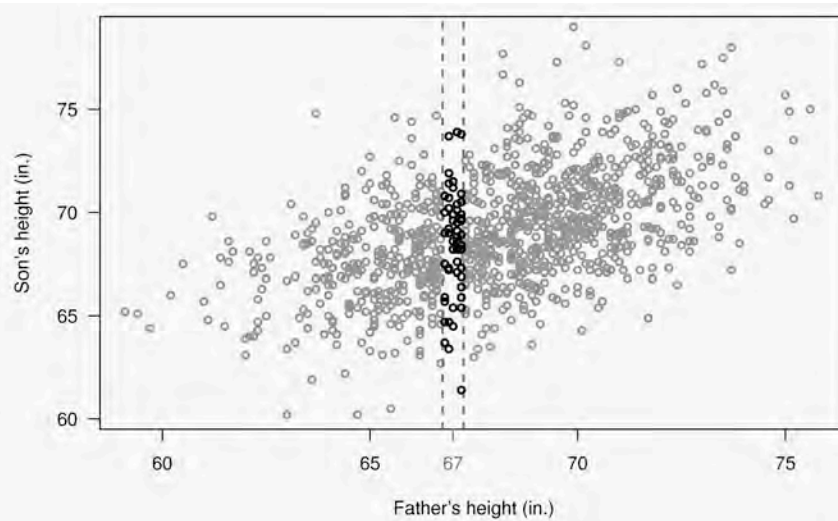
Heights of fathers and sons



9



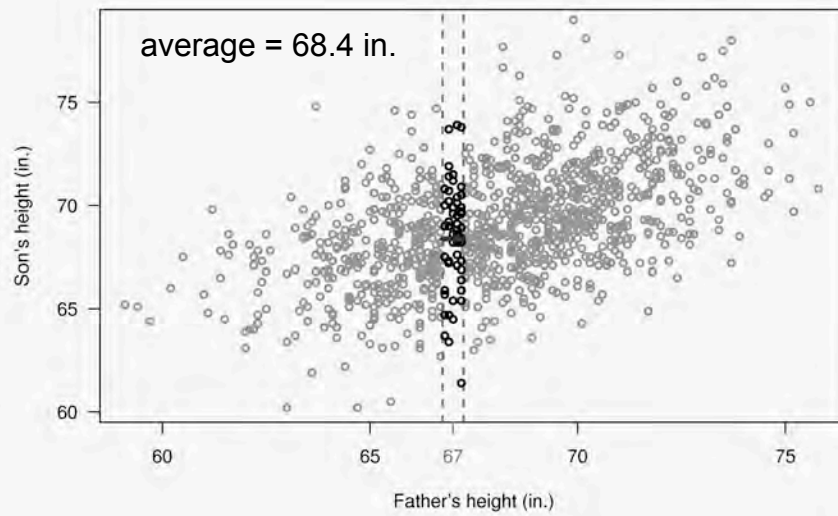
Heights of fathers and sons



10



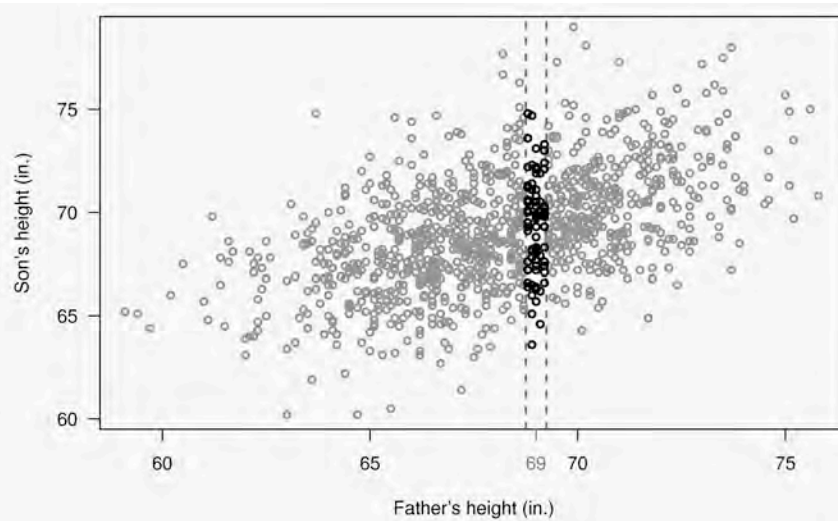
Heights of fathers and sons



11



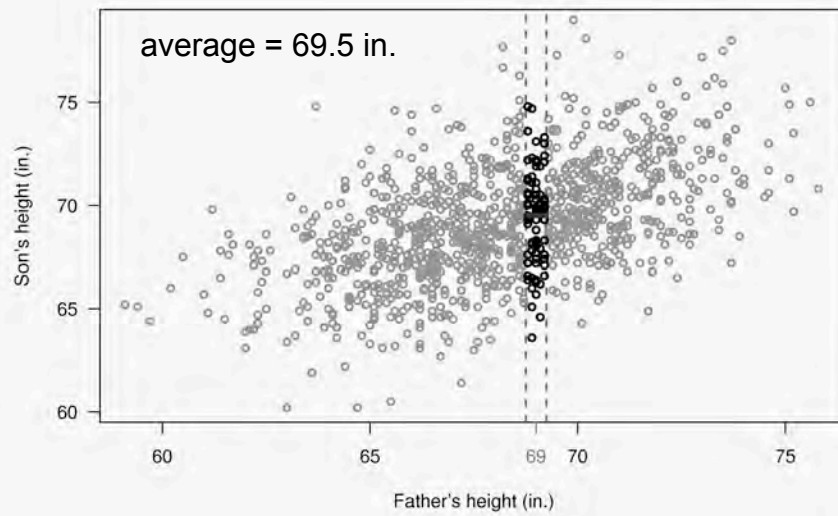
Heights of fathers and sons



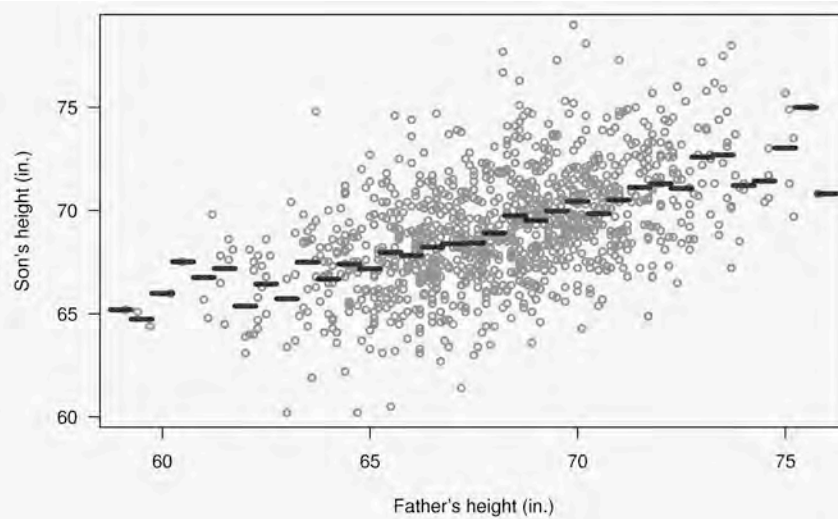
12



Heights of fathers and sons

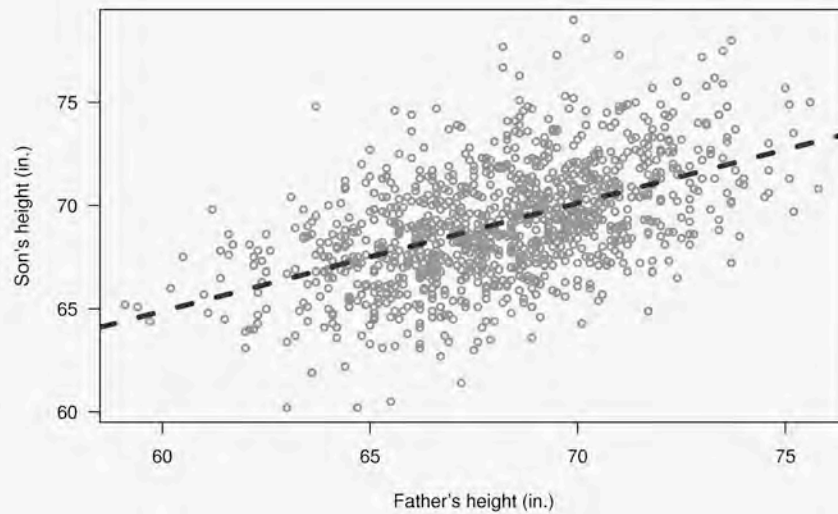


Heights of fathers and sons





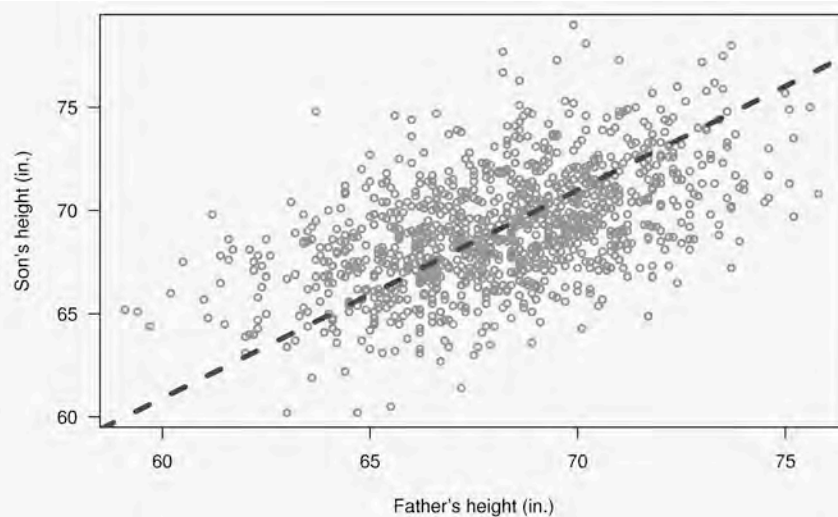
Heights of fathers and sons



15



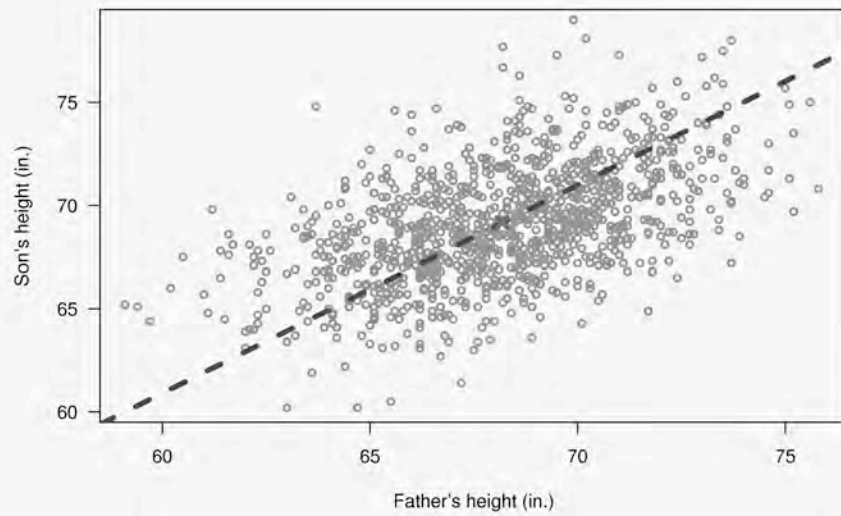
Heights of fathers and sons



16



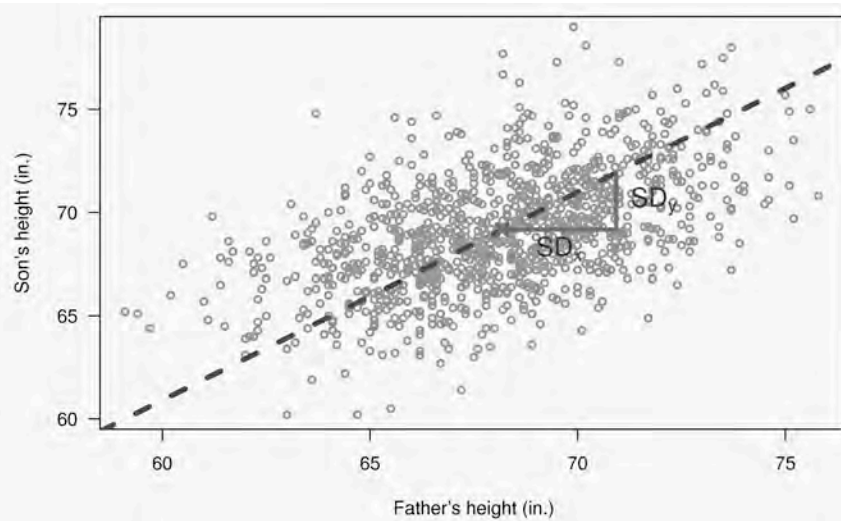
Heights of fathers and sons



17



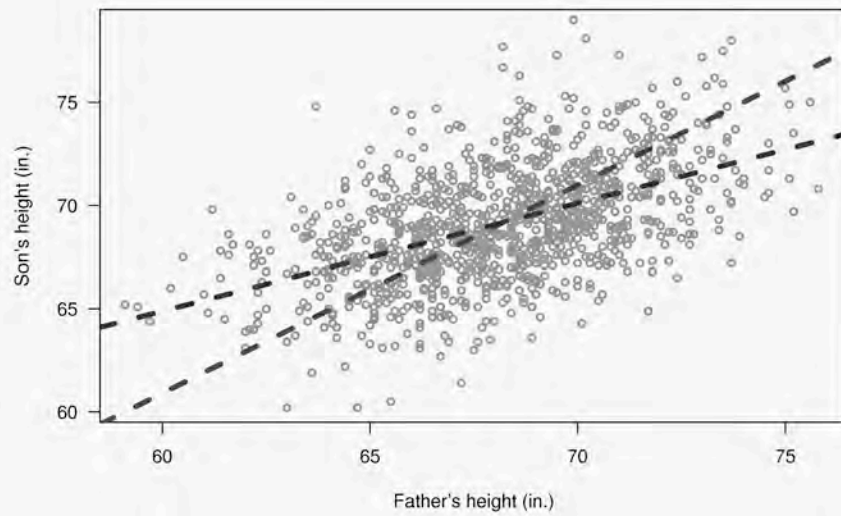
Heights of fathers and sons



18



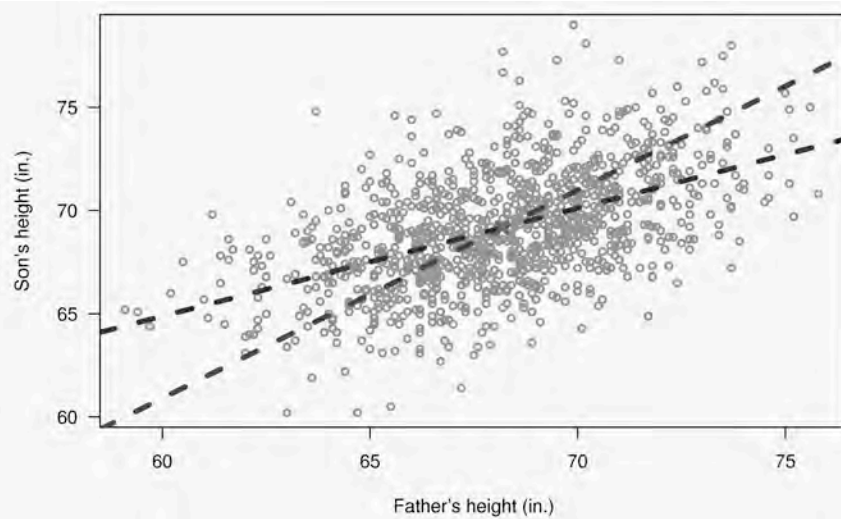
Heights of fathers and sons



19



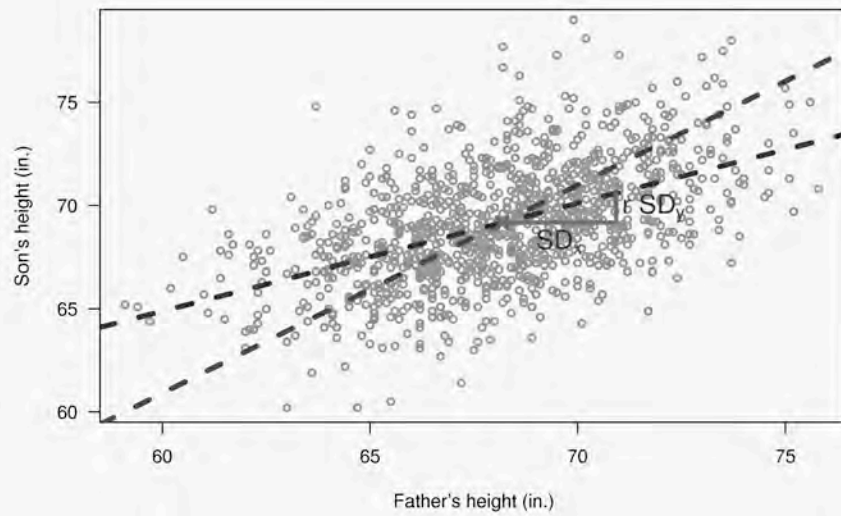
Heights of fathers and sons



20



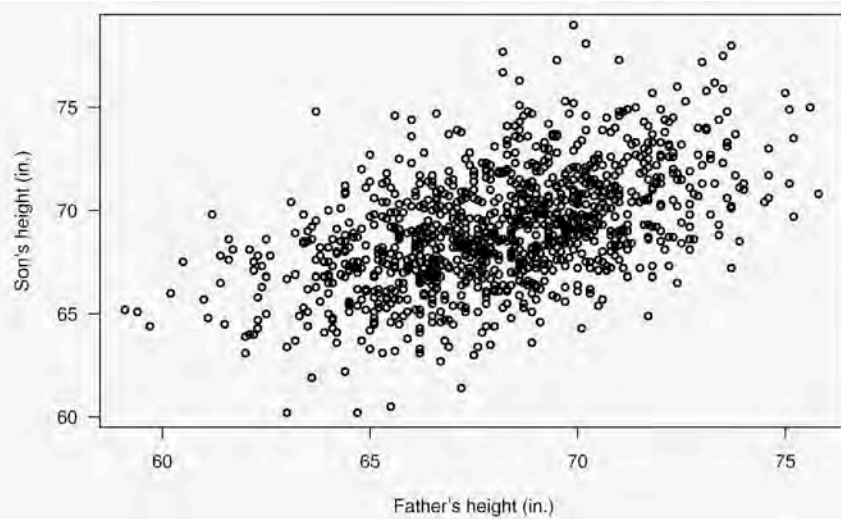
Heights of fathers and sons



21



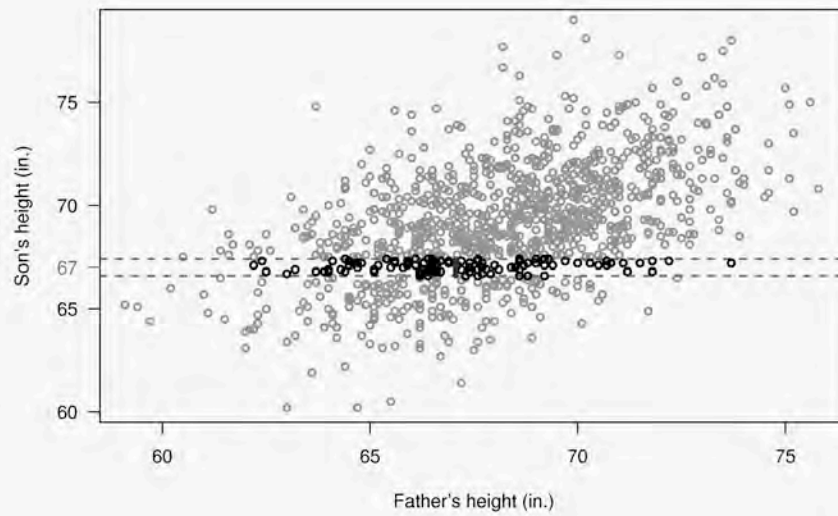
Heights of fathers and sons



22



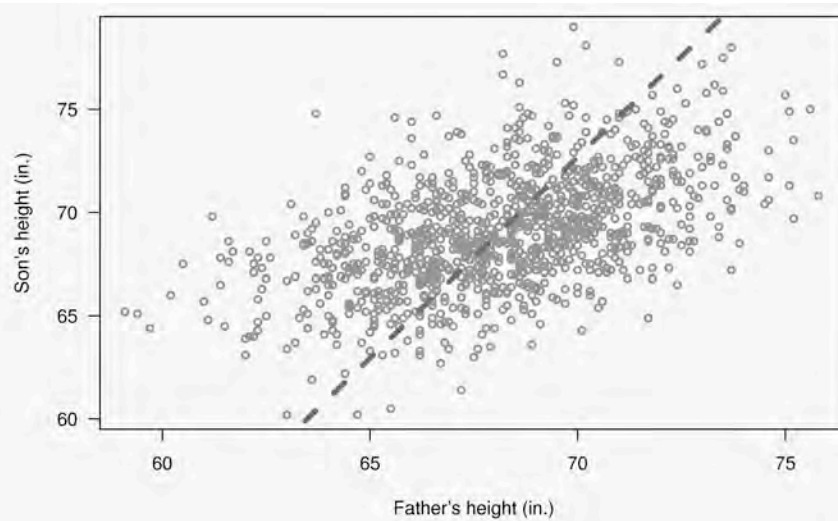
Heights of fathers and sons



23



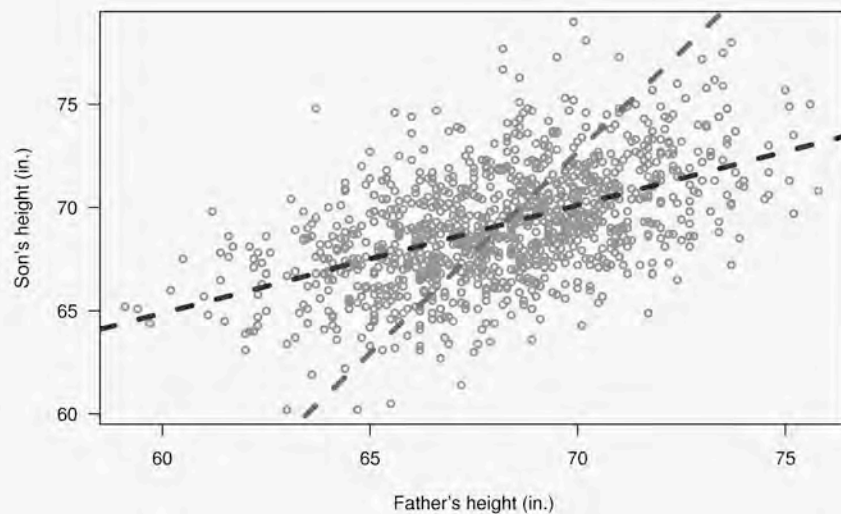
Heights of fathers and sons



24



Heights of fathers and sons



25



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.

26