Topic 6: Correlation

STOR 155: Introduction to Data Models and Inference

Dr. Teressa Bergland

Fall 2025

Announcements

Course PSAs:

- Homework 4 due **TODAY, Tuesday 2/03 on WebAssign**
- Homework 5 due **TODAY, Tuesday 2/03 on WebAssign**
- Homework 6 due **Thursday 2/05**
- **First Midterm Exam: Thursday 2/12**

Warm-Up:

- Linear association recap!

Correlation

The correlation coefficient r:

$$r = \frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{x_i - \bar{x}}{s_x}\right)\left(\frac{y_i - \bar{y}}{s_y}\right)$$

	x	y
Point 1	3	4
Point 2	5	5
Point 3	7	3

Let's practice! For this data, $r = -\frac{1}{2}$

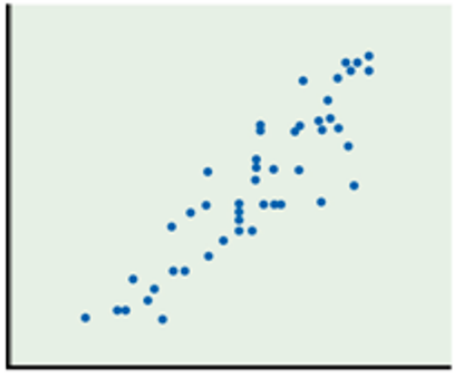
Properties of Correlation

- Possible values
- Sign meaning
- r does not change even if...
- A note on outliers

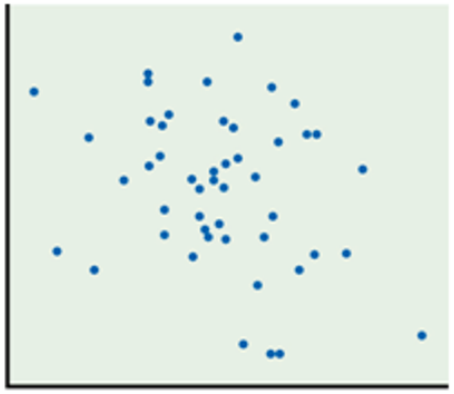
Some Examples

Correlation $r = 0$

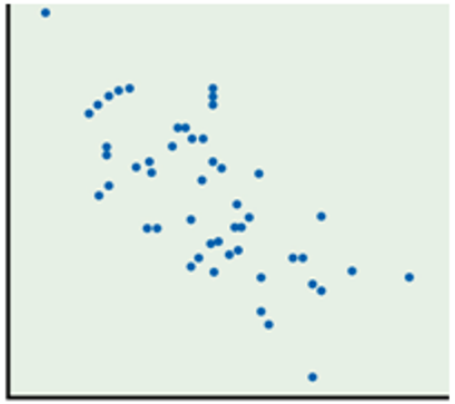
Correlation $r = 0.5$



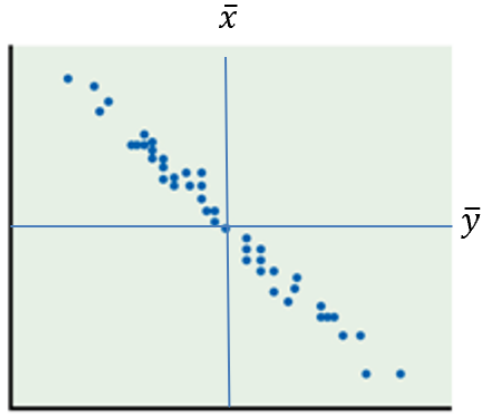
Correlation $r = 0.9$



Correlation $r = -0.3$



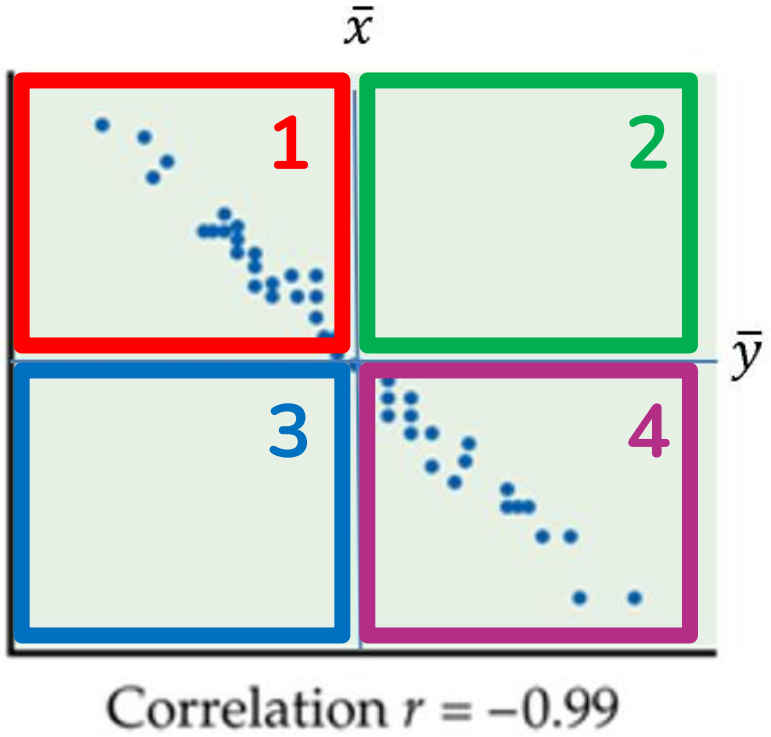
Correlation $r = -0.7$



Correlation $r = -0.99$

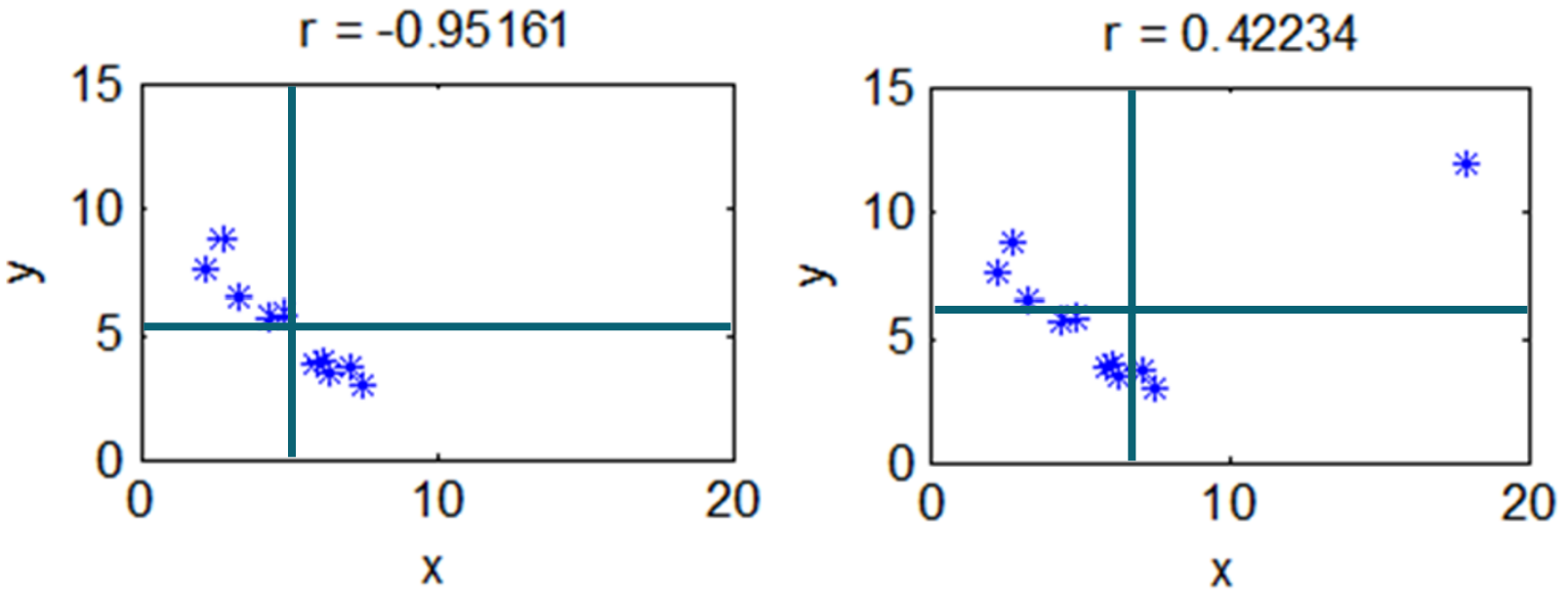
Contributions to Correlation

$$r = \frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{x_i - \bar{x}}{s_x}\right)\left(\frac{y_i - \bar{y}}{s_y}\right)$$



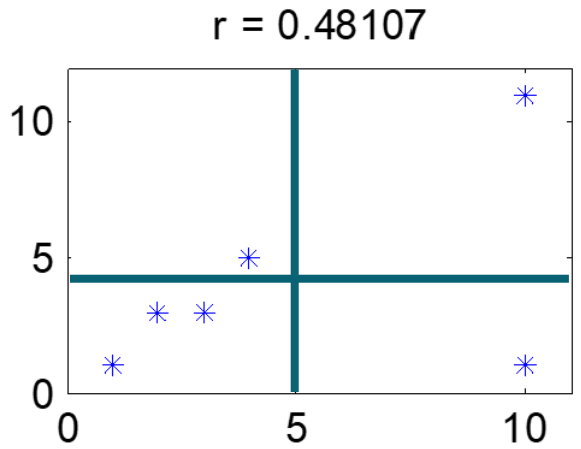
Region	$\frac{x_i - \bar{x}}{s_x}$	$\frac{y_i - \bar{y}}{s_y}$	Product
1			
2			
3			
4			

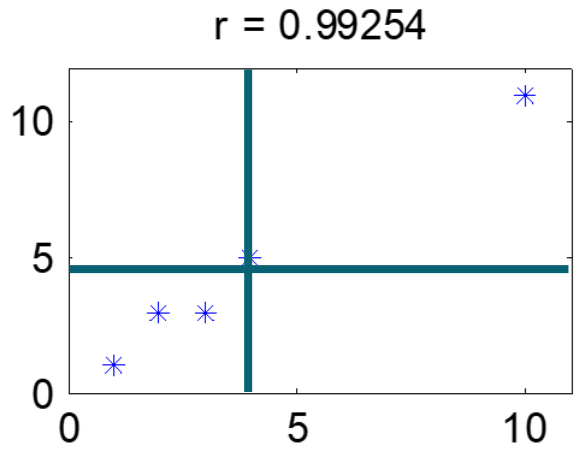
Outlier Distortion

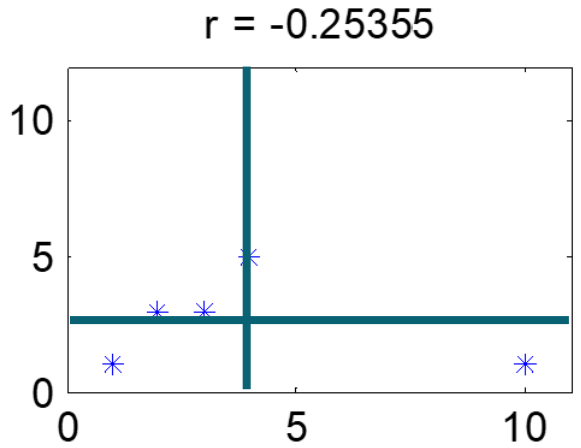


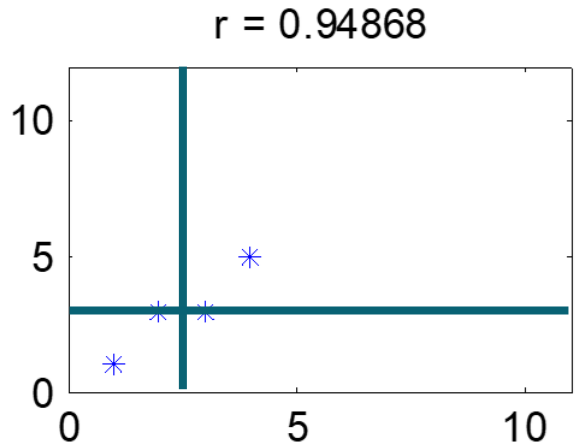
Outlier Distortion

x	y
1	1
2	3
3	3
4	5
10	1
10	11

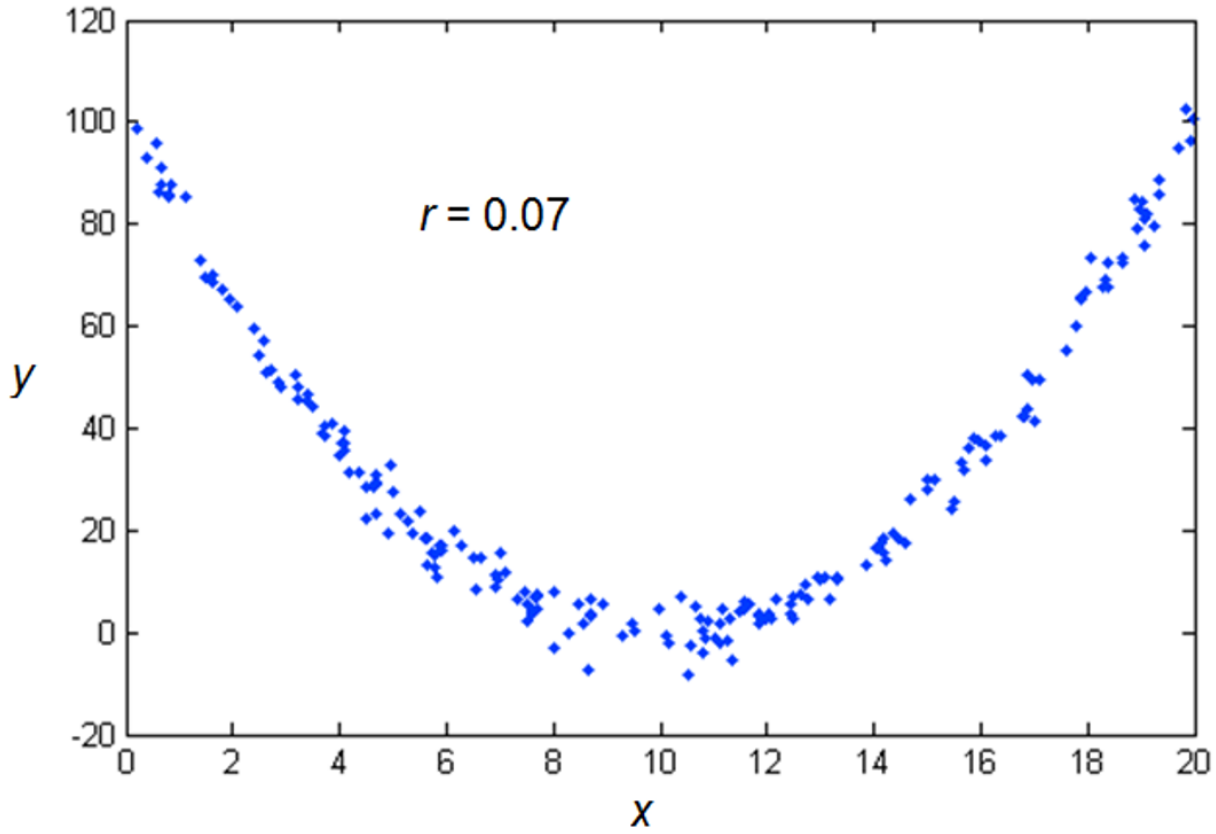








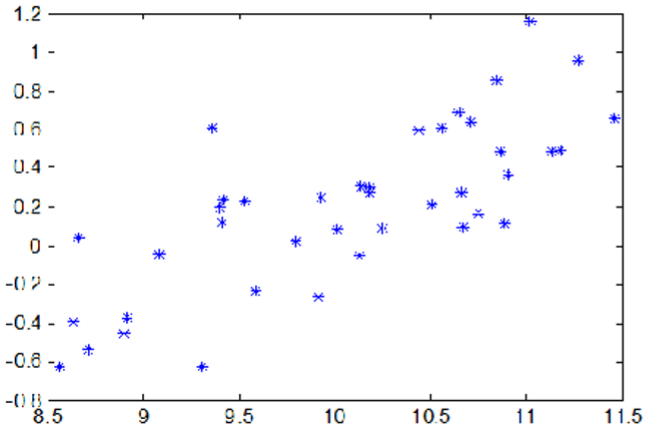
Correlation and Nonlinear Association



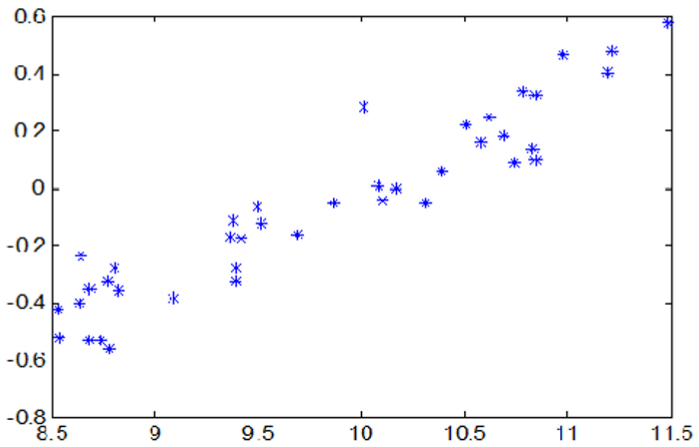
Sanity Check!

Which of these has the correlation **closest to +1?**

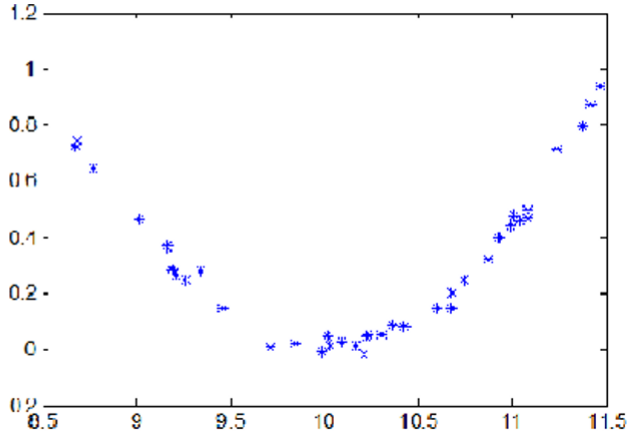
A



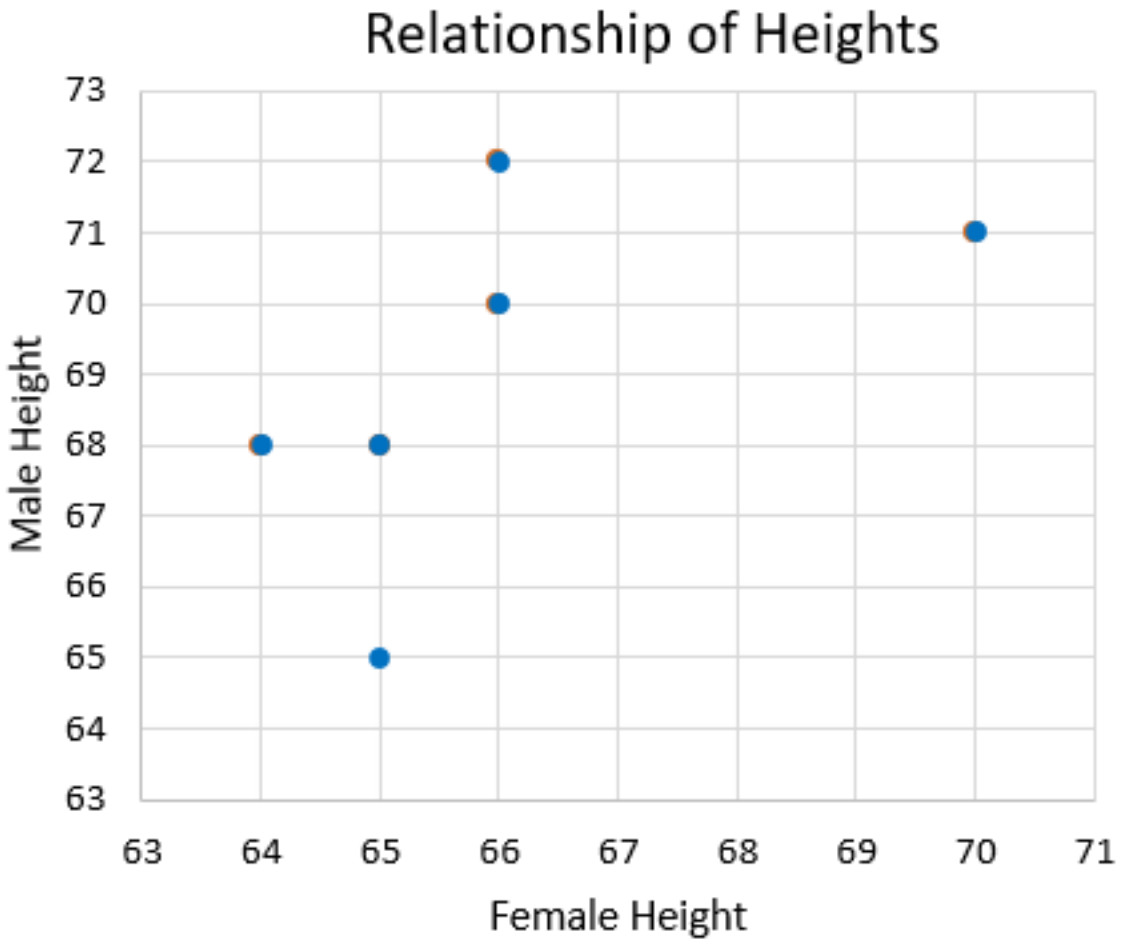
B



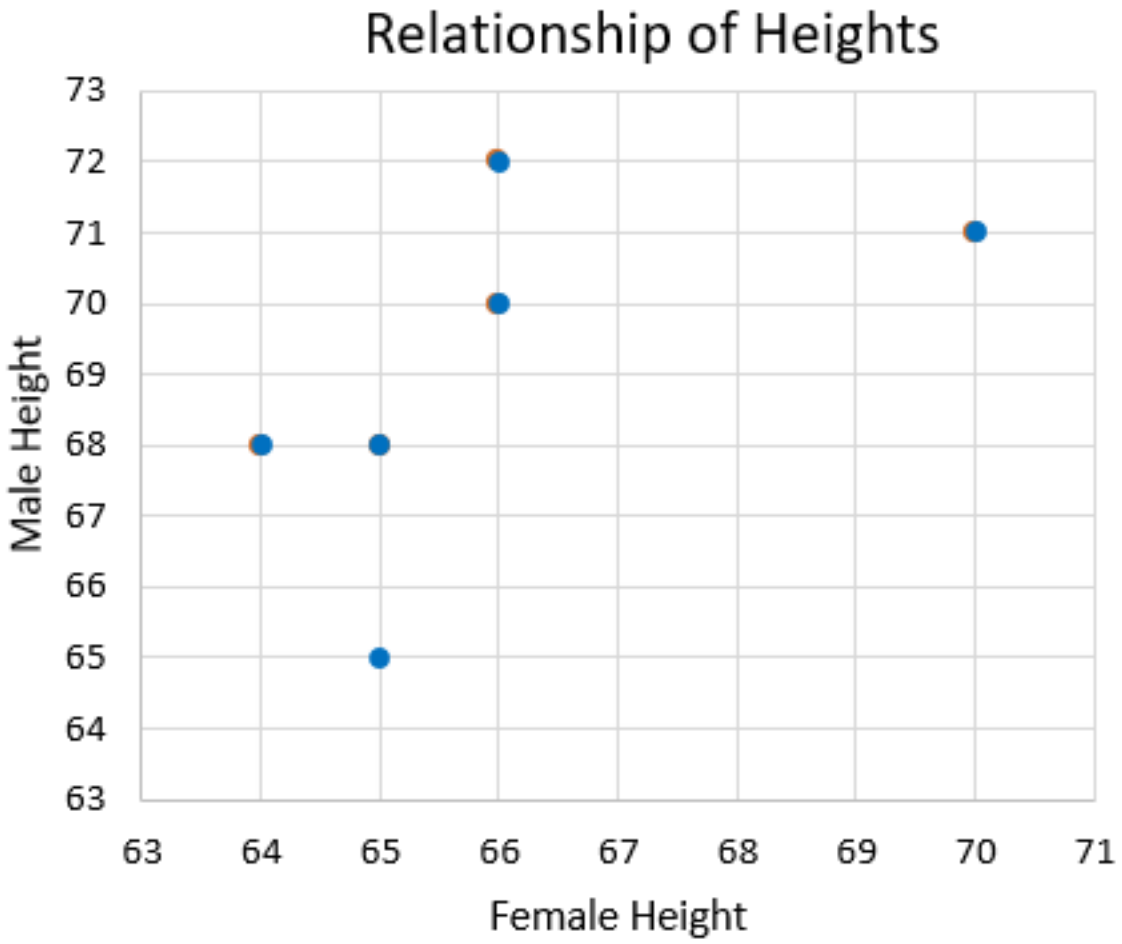
C



Let's Practice! Heights of Couples



Let's Practice! Heights of Couples



1. Effect on correlation if all men were 6 inches shorter? Conclusions about dating tendencies?

2. Effect on correlation if heights were in centimeters?

3. Correlation if each woman dated a man **exactly** 3 inches taller than she is?