

EPSY 5261 : Introductory Statistical Methods

Day 25
Introduction to Correlation

Learning Goals

- At the end of this lesson, you should be able to...
 - Interpret scatterplots
 - Explain when to use correlation to explain a relationship between variables
 - Interpret a correlation

Scatterplots

Explanatory/Response variables

- Two quantitative variables needed for correlation
- Response (Y): the variable you want to predict
- Explanatory (X): the variable you want to use to predict Y.

Example

Coffee sales at an outdoor stadium



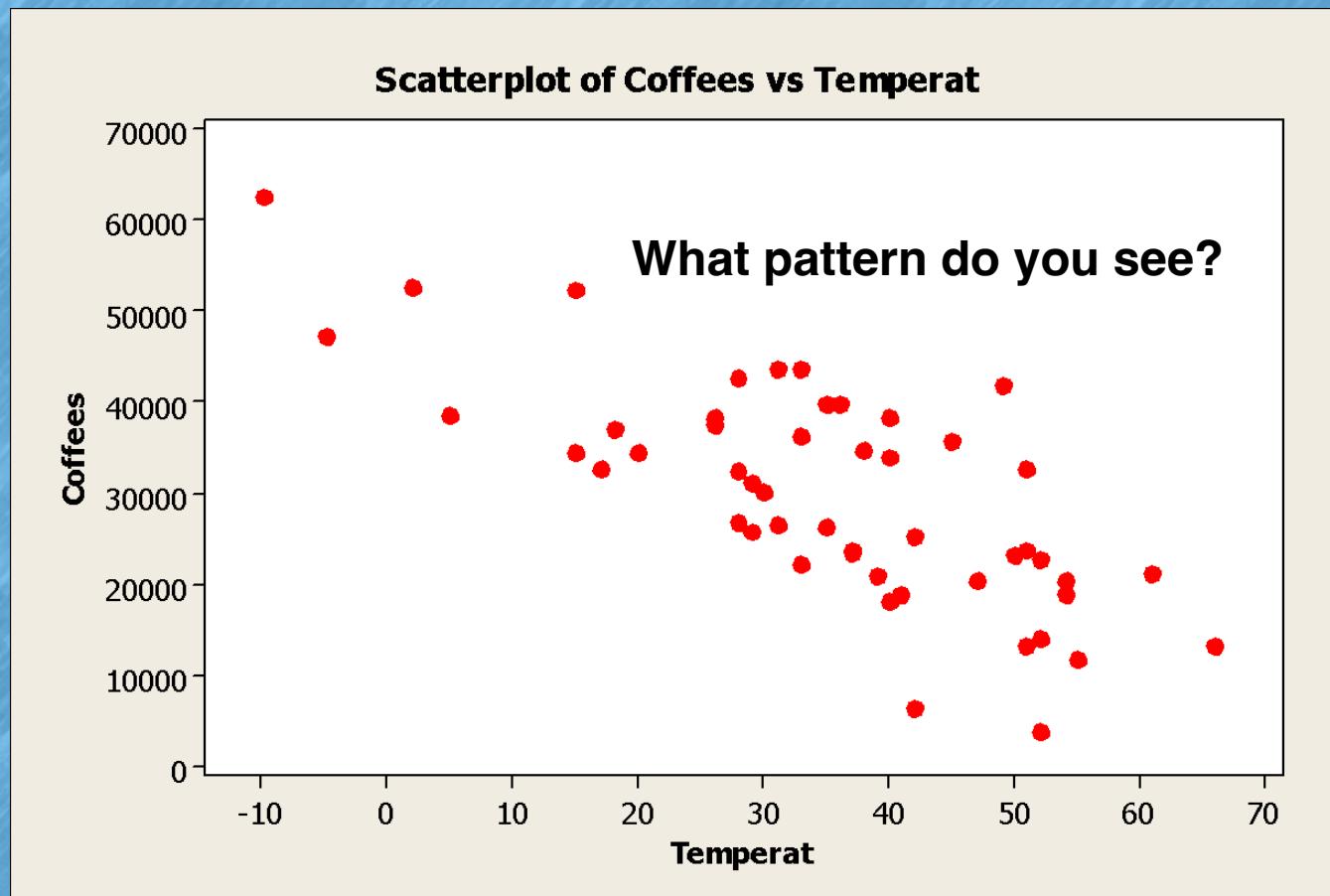
X (use to predict coffee sales)

Temperature outside

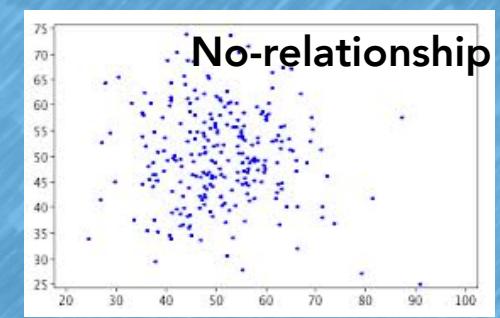
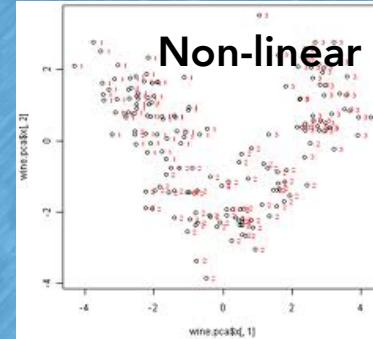
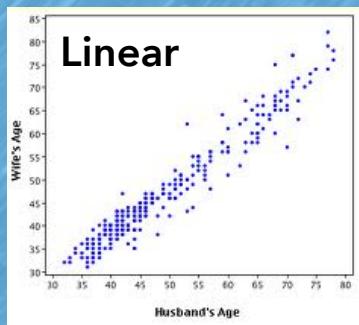


Y (want to predict)

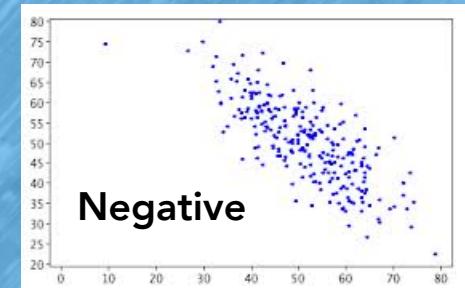
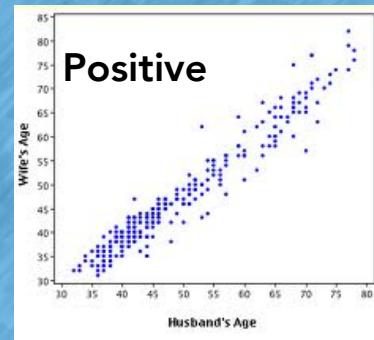
Scatterplot: temperature vs. coffees sold



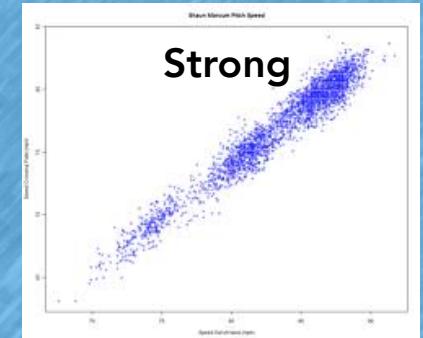
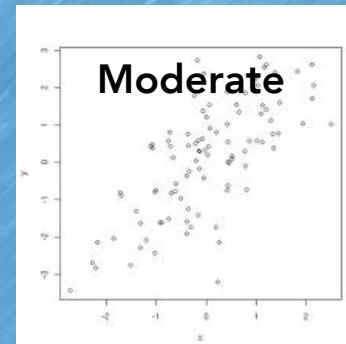
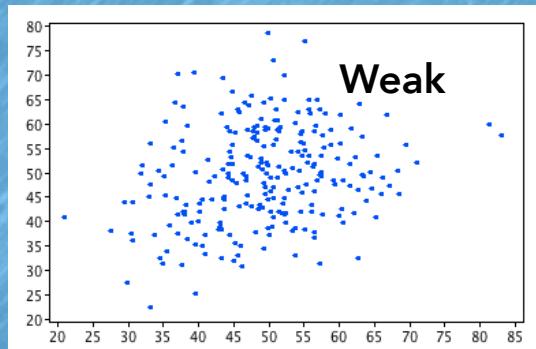
**FORM/
TREND:**



DIRECTION:



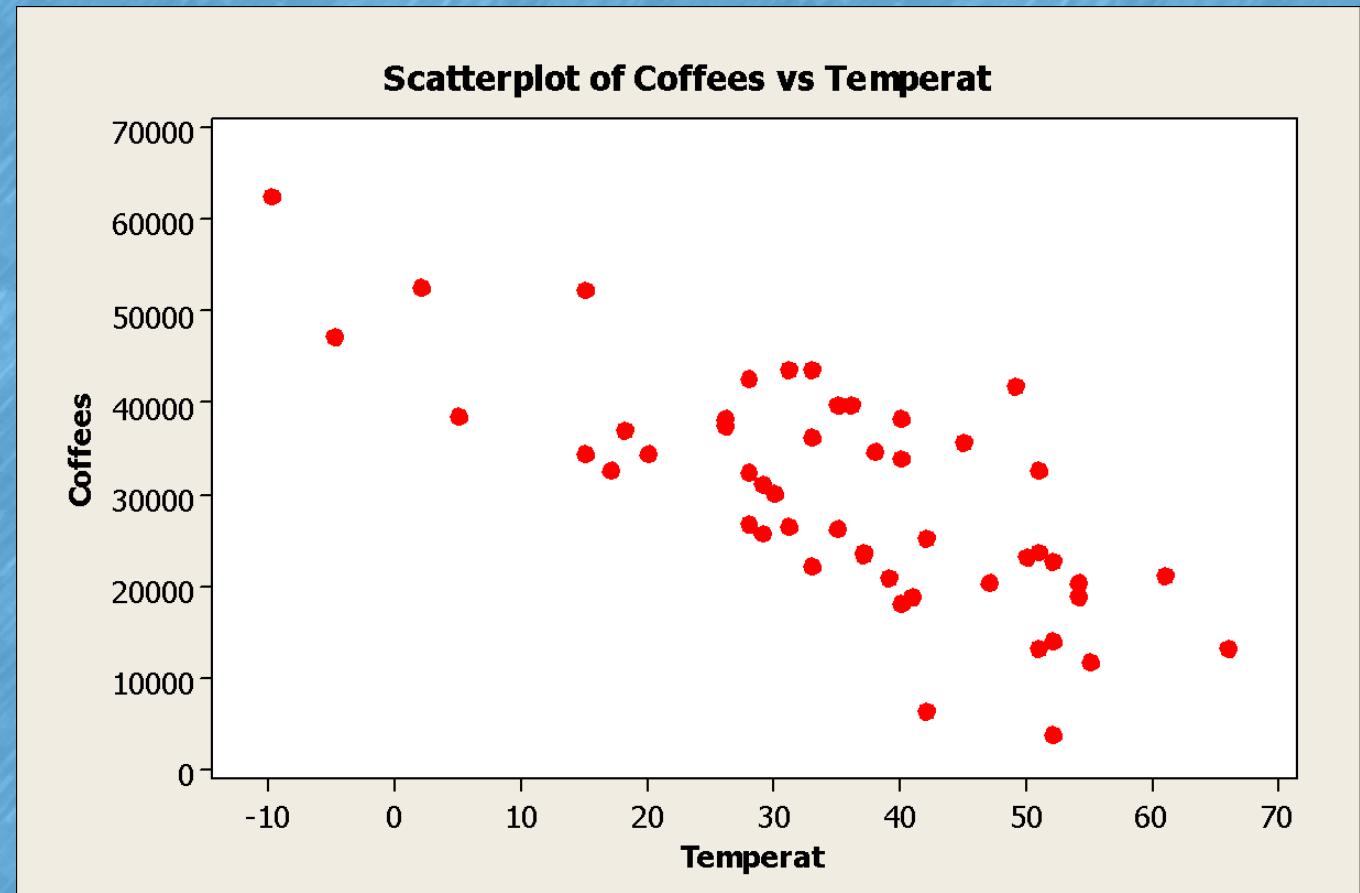
STRENGTH:

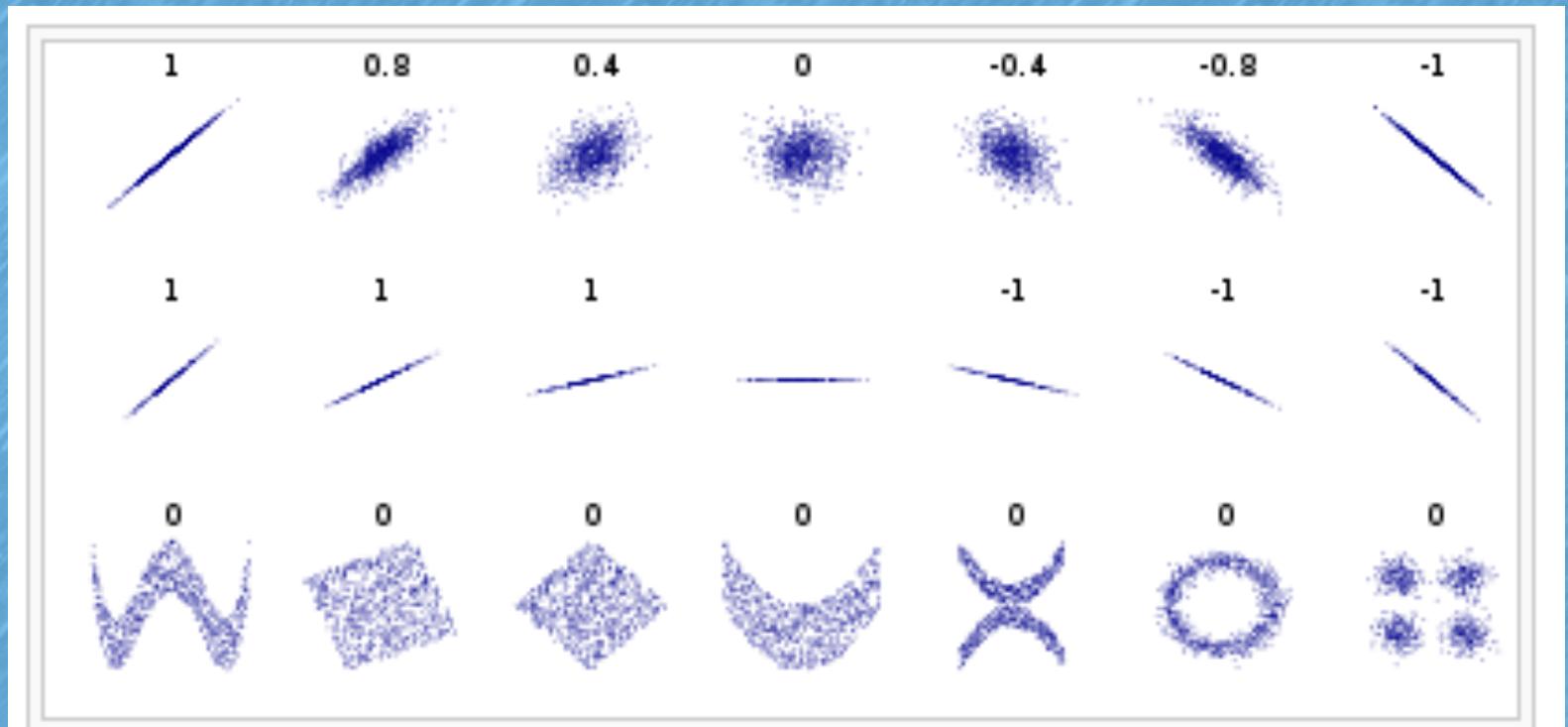


Trend?

Direction?

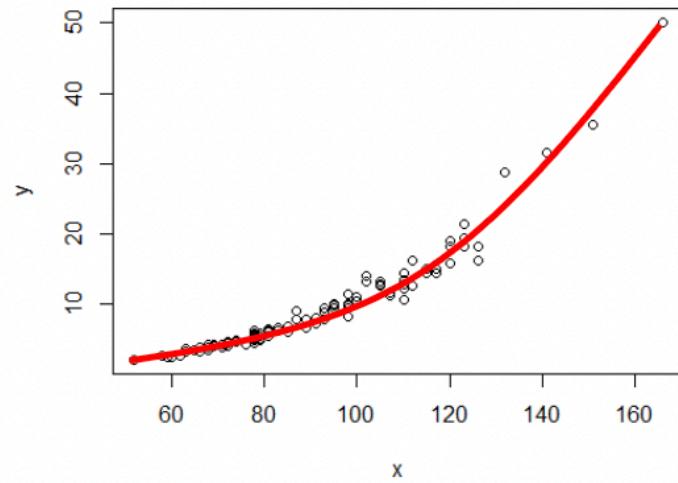
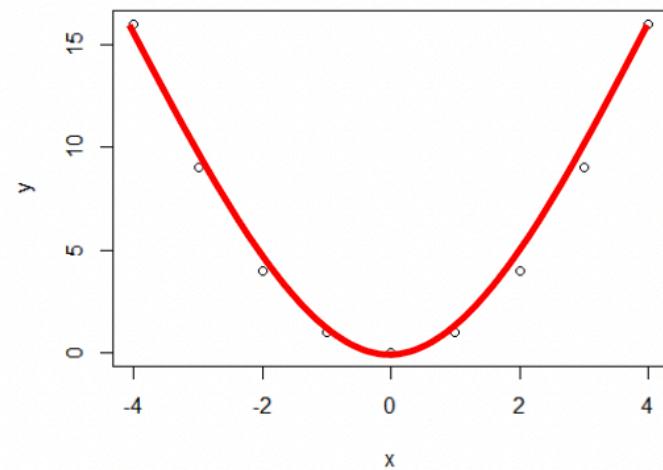
Strength?





BUT, the correlation tells us nothing about the FORM of the relationship.

Correlation Caution: always plot your data!



Interpreting a Correlation

- Direction: positive or negative?
 - If positive: As X increases, Y tends to increase
 - If negative: As X increases, Y tends to decrease
- How strong is the linear relationship?
 - Weak (closer to 0)
 - Moderate (somewhere in the middle – maybe $\pm .4$ to $\pm .7$)
 - Strong (closer to ± 1)

Example: Temperature and Coffee

- Correlation between
- Temperature and Coffees = -0.741
- How would you interpret this correlation?

Negative, moderate/strong. As temperature goes up,
of coffees sold tends to go down.

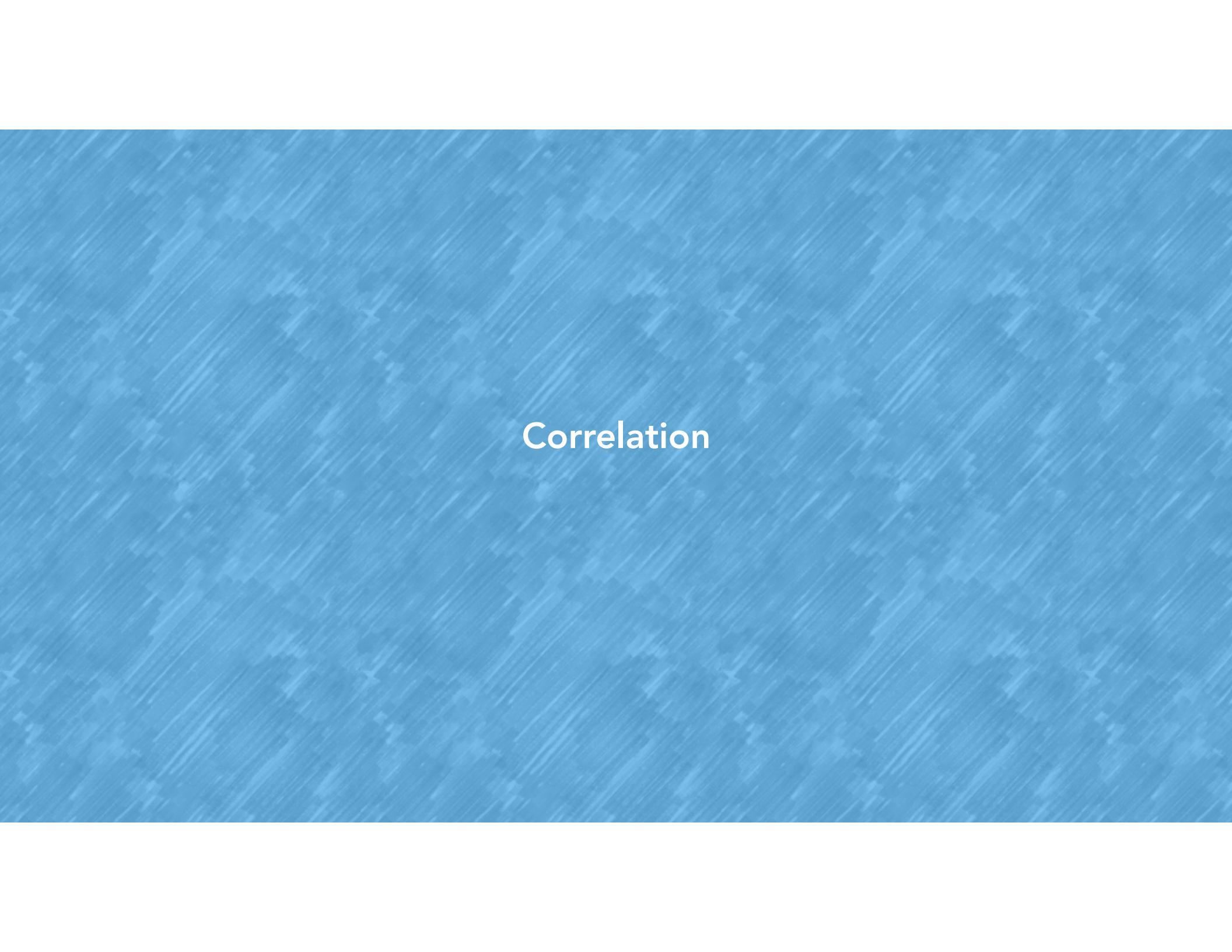
Caution: Correlation doesn't imply CAUSATION!



There may be a confounding variable -
in this case temperature

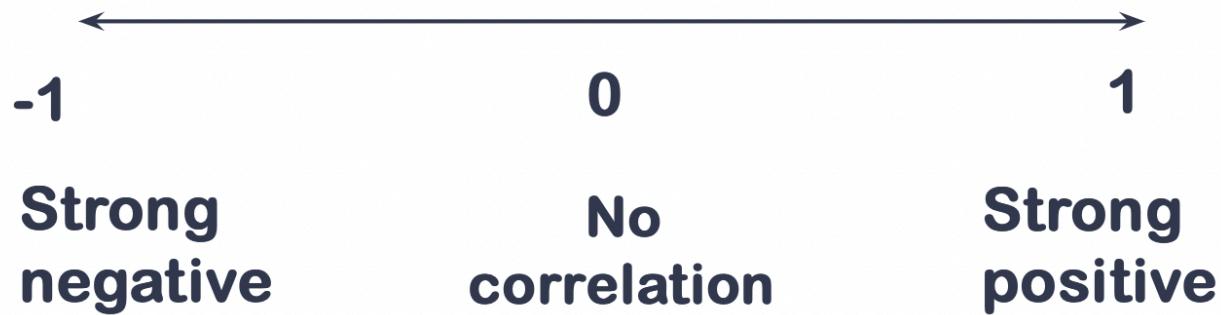
Other Correlation Facts

- Correlation between X and Y = Correlation between Y and X (symmetric).
- Correlation has no units
- If units change (e.g. kg to lbs), correlation stays the same!

The background of the slide is a solid blue color. It features a subtle, faint pattern of small white stars scattered across the surface. A single, thin, dark diagonal line runs from the top-left corner towards the bottom-right corner.

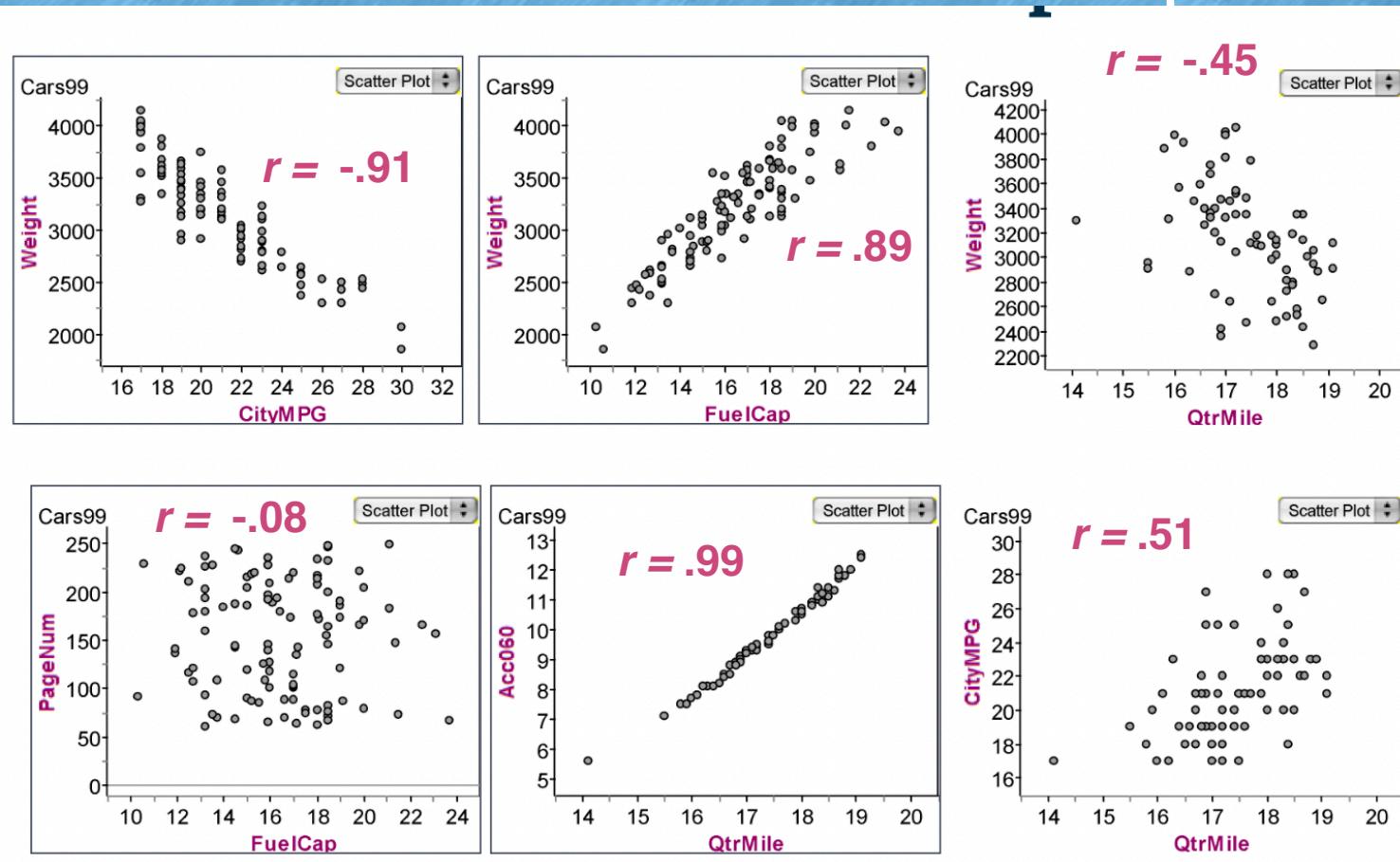
Correlation

Correlation Coefficient (r)



Tells us about strength and direction of linear relationship

Car Correlation Examples



Correlation and House Prices Practice Activity

Summary

- We can use a correlation to describe a linear relationship between two quantitative variables
- A negative correlation implies a negative/indirect relationship
- A positive correlation implies a positive/direct relationship
- Correlation can only be between -1 and 1
- Correlations close to -1 and 1 are strong
- Correlations close to 0 are weak/non existent