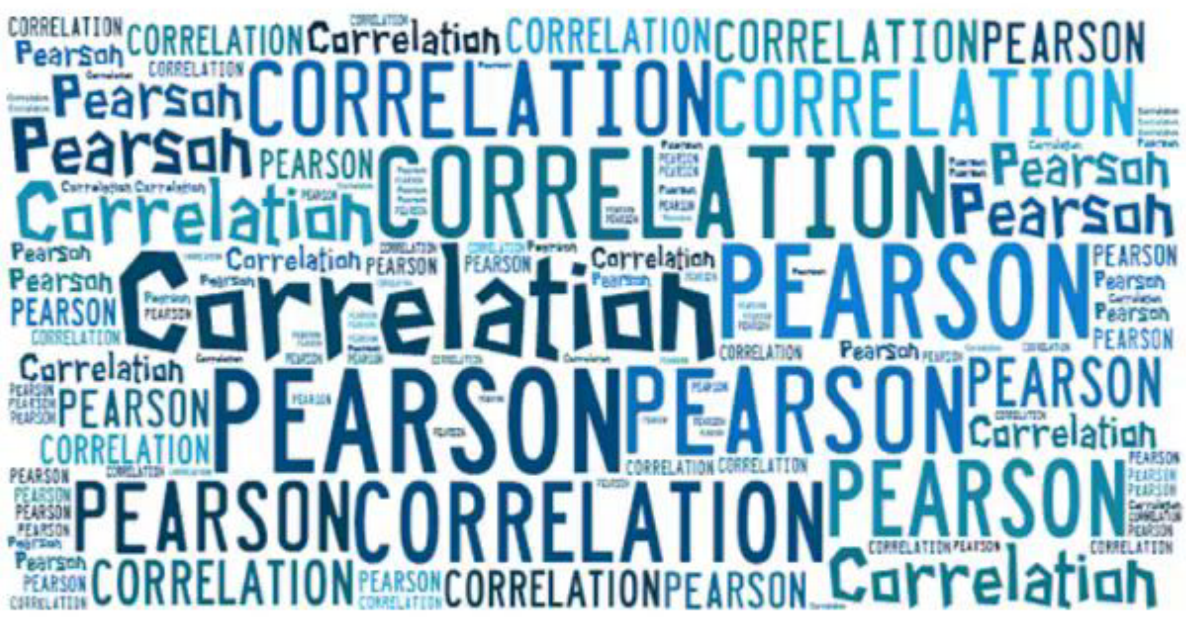
WESLEYAN
UNIVERSITY



Variable types and inferential tools

		Response	
		Categorical	Quantitative
Explanatory	Categorical	C → C	C → Q
	Quatitative	Q → C	Q → Q

Variable types and inferential tools

WESLEYAN UNIVERSITY

Chi Square Test of Independence

Analysis of Variance

		Response	
		Categorical	Quantitative
Explanatory	Categorical	C →C	C →Q
	Quatitative	Q →C	Q →Q

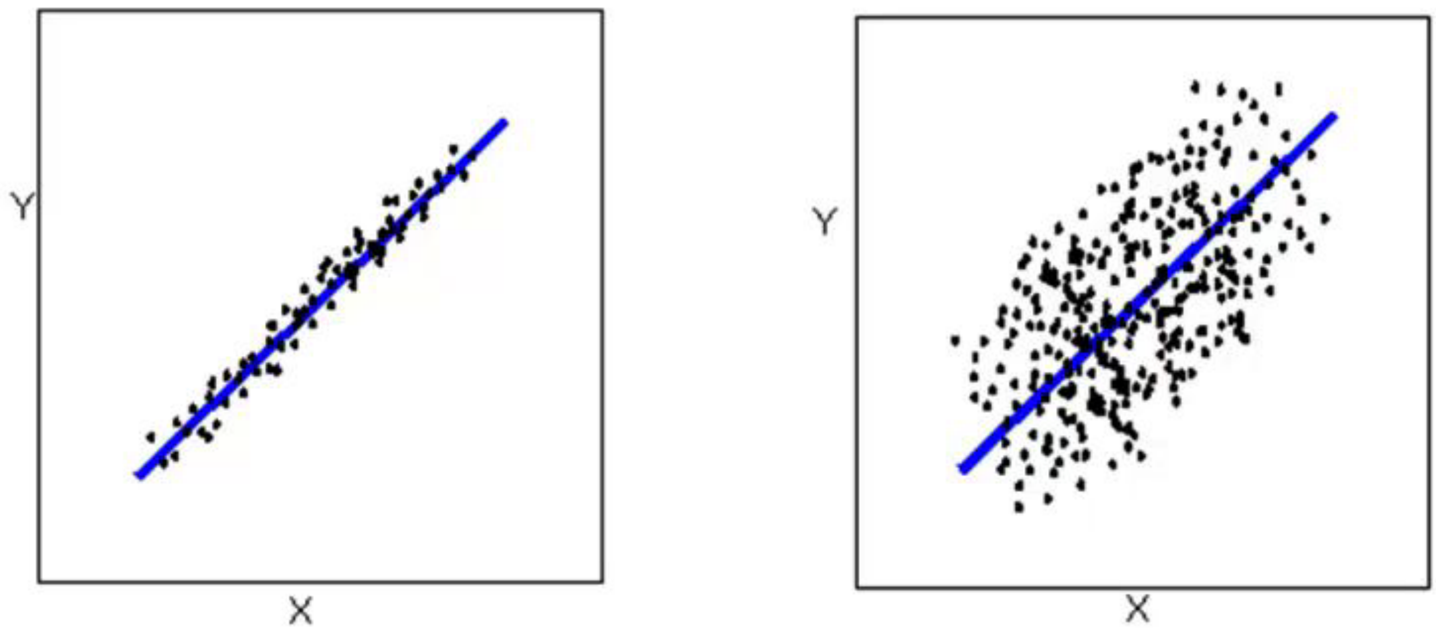
Pearson Correlation

Open Learning Initiative

Transforming higher education through the science of learning.

The Correlation Coefficient (r)

Definition: The **correlation coefficient (r)** is a numerical measure that measures the **strength** and **direction** of a linear relationship between two quantitative variables.

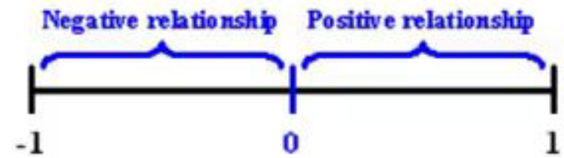


strong relationship   weaker relationship

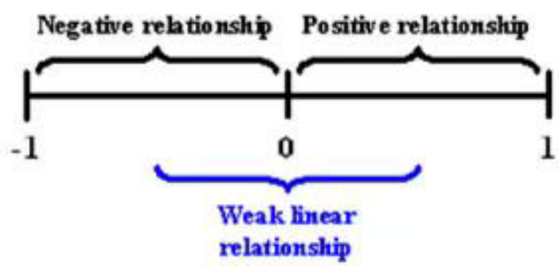
The Correlation Coefficient (r)

The value of r ranges from -1 to 1.

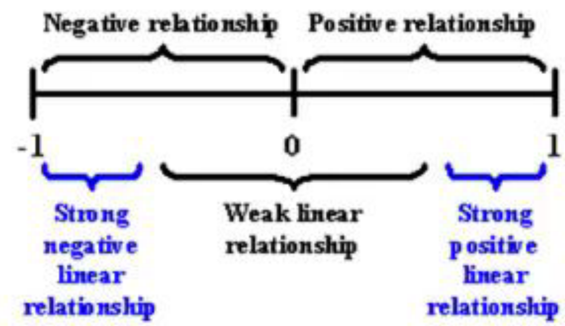
Negative values of r indicate a negative direction for a linear relationship, and positive values of r indicate a positive direction for a linear relationship.



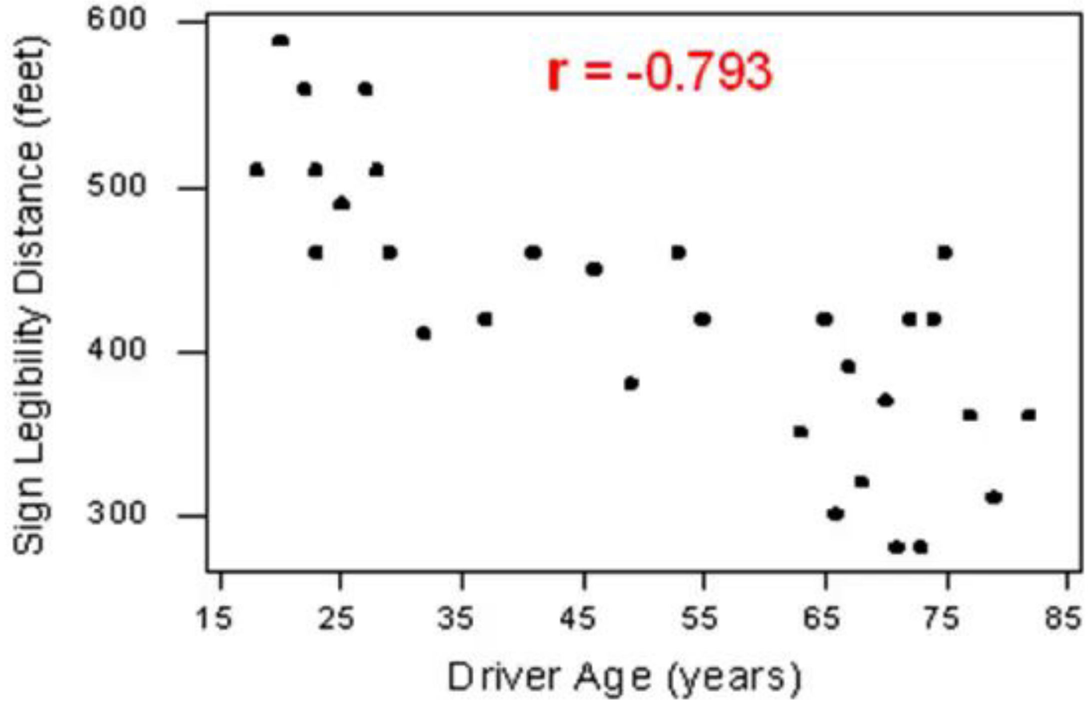
Values of r that are close to 0—either negative or positive—indicate a weak linear relationship.



Values that are close to -1 or close to 1 indicate a strong linear relationship, either negative or positive.

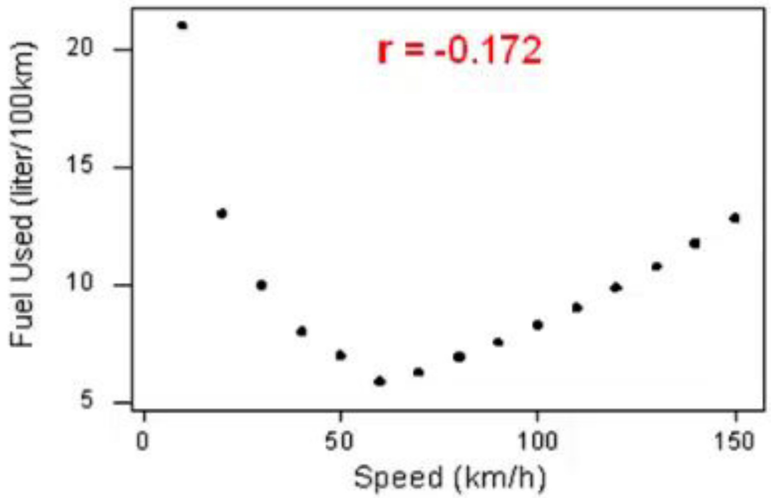


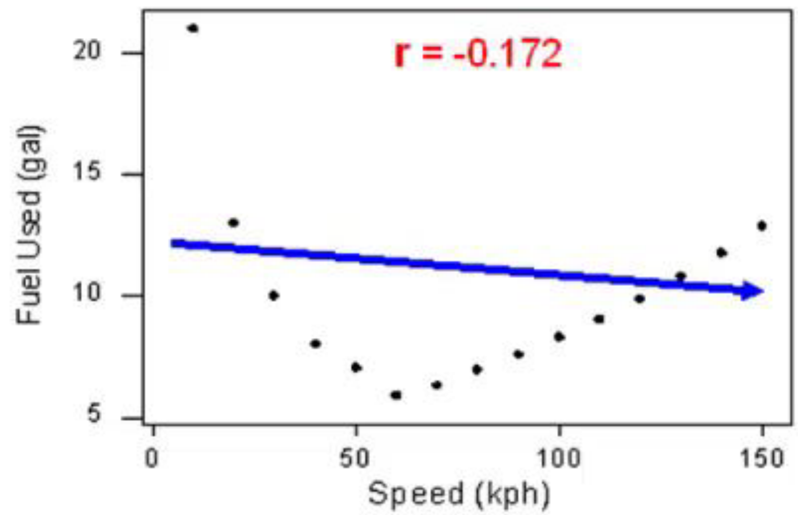
Example: Highway Sign Visibility



Utts and Heckard, Mind on Statistics (2002)

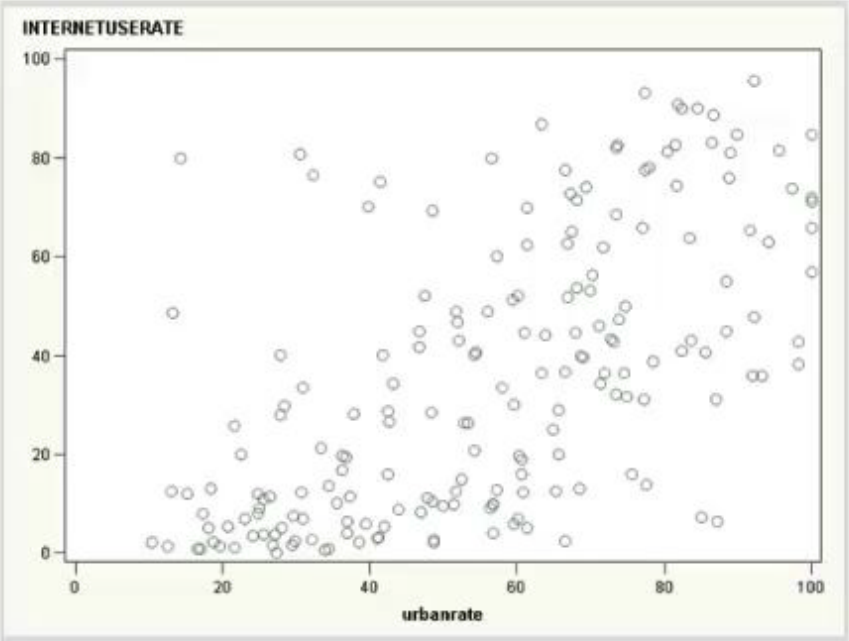
r only evaluates the linear relationship

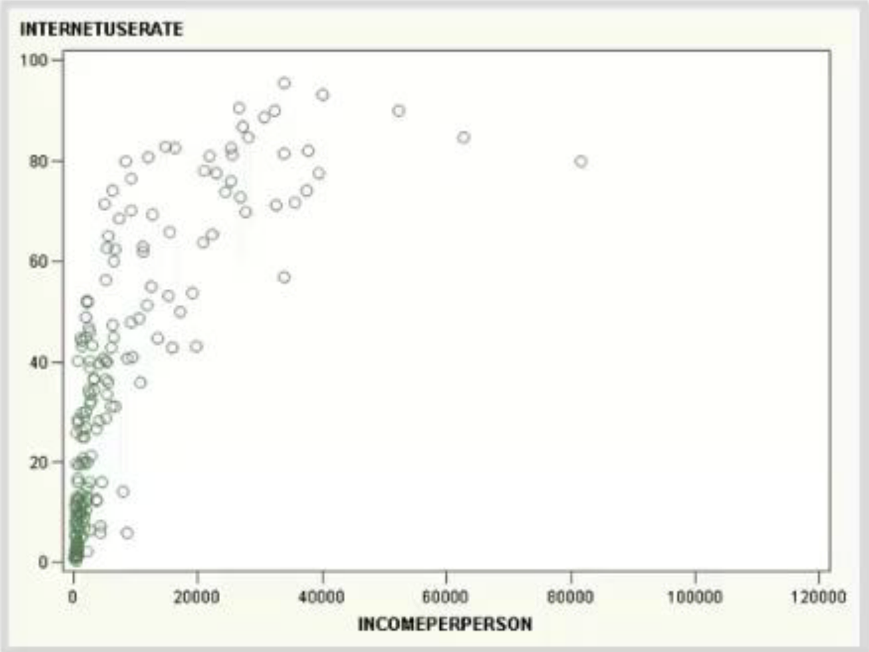




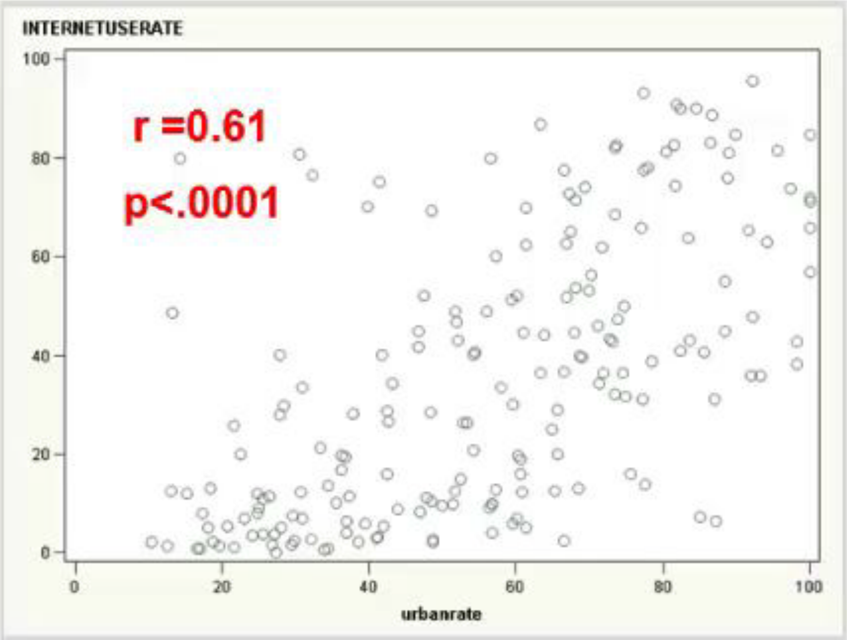
Internet Users (per 100 people) by Urban Population and Income

WESLEYAN UNIVERSITY





Internet Users (per 100 people) by Urban Population and Income

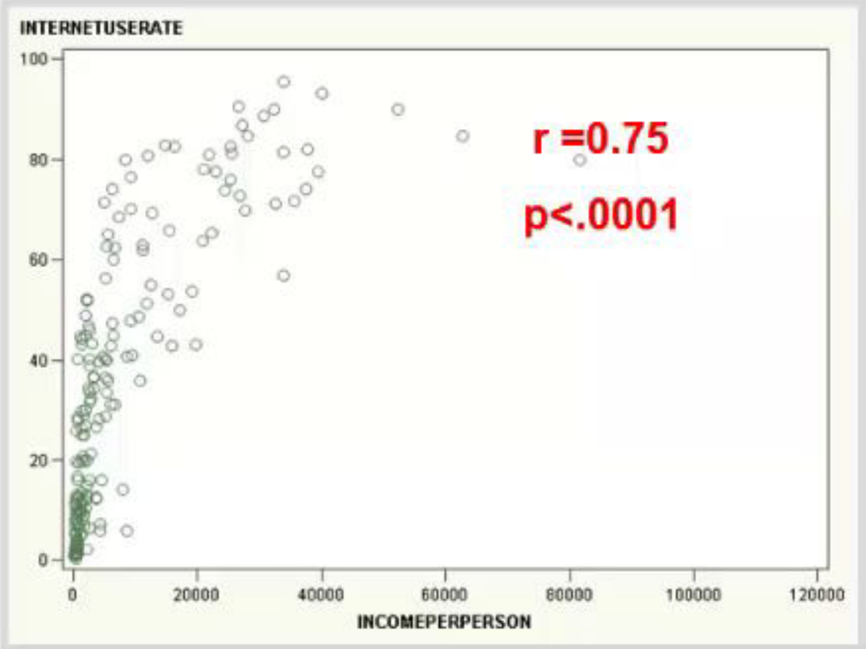


$0.61^2=0.37$

If we know the "urban rate", we can predict 37% of the variability we will see in the rate of internet use

$0.75^2=0.56$

If we know "income per person", we can predict 56% of the variability we will see in the rate of internet use



WESLEYAN

UNIVERSITY