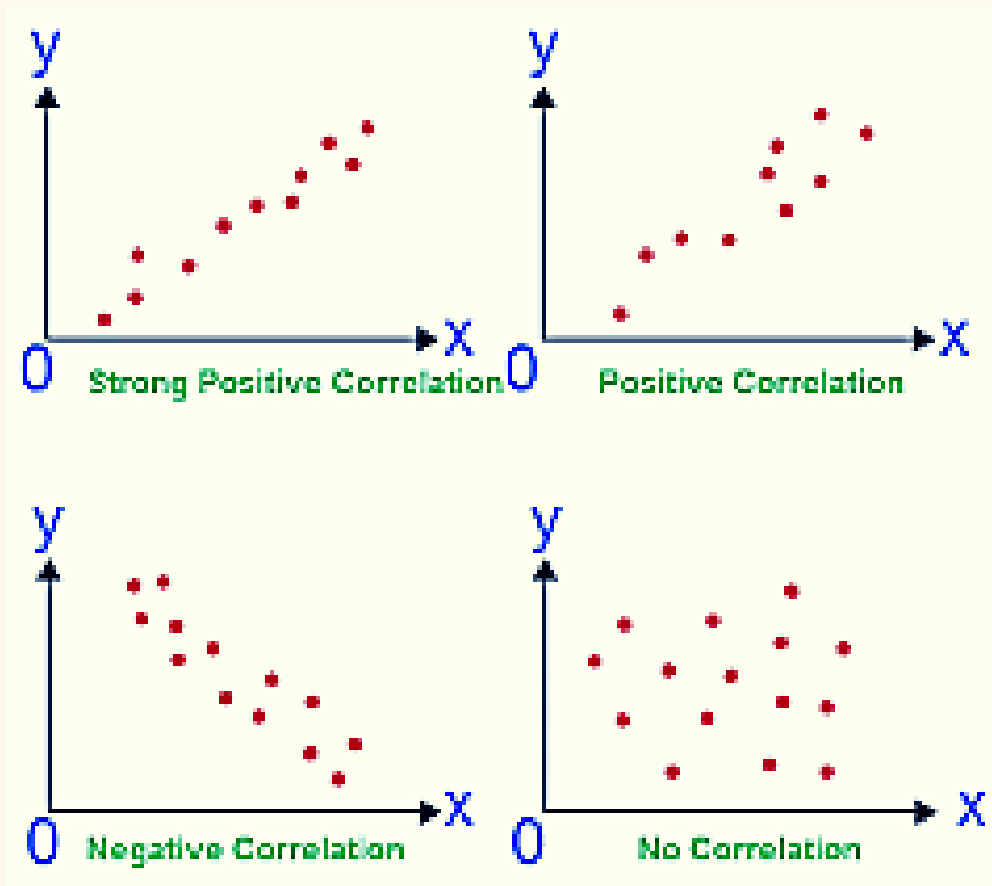


# Correlation



Correlation is a statistical technique that can show whether and how strongly pairs of variables are related.

## Def:

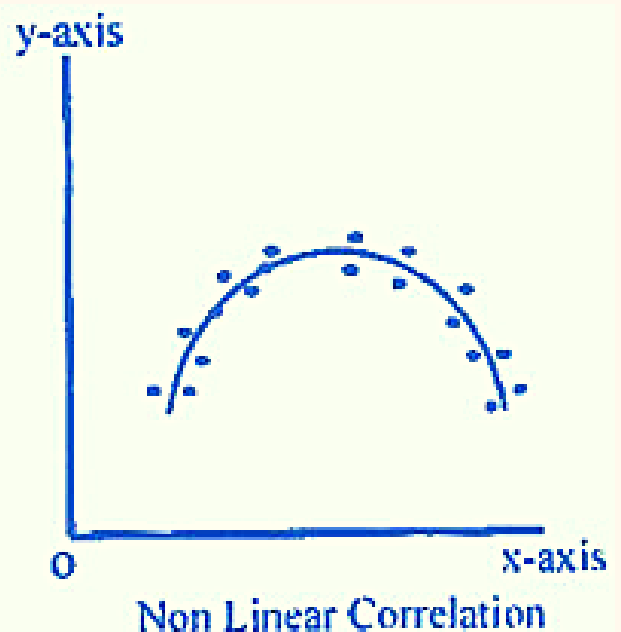
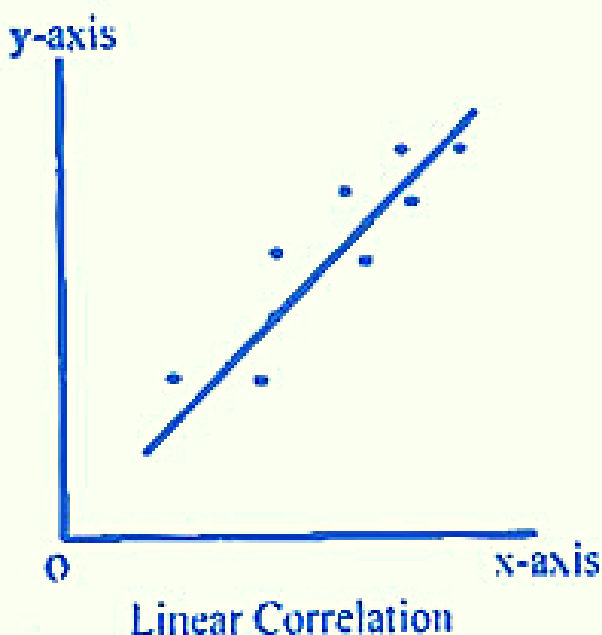
Correlation is the relationship between two sets of variables used to describe or predict information. There is an emphasis here on relationship. Sometimes we can use correlation to find causality, but not always. Remember that correlation can either be positive or negative.

### Linear Correlation

Correlation is said to be linear if the ratio of change is constant. When the amount of output in a factory is doubled by doubling the number of workers, this is an example of linear correlation.

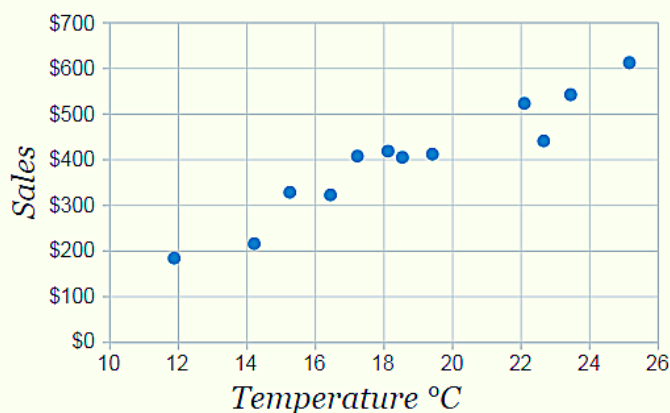
### Non Linear Correlation

Correlation is said to be non linear if the ratio of change is not constant. In other words, when all the points on the scatter diagram tend to lie near a smooth curve, the correlation is said to be non linear (curvilinear). This is shown in the figure on the right below.



# Correlation vs Causation

If there is a correlation, then sometimes we can assume that the dependent variable changes solely because the independent variables change. This is where the debate between correlation and causation occurs. However, there is a difference between cause and effect (causation) and relationship (correlation). Sometimes these areas can be confused and muddled when analyzing data.



Consider the sales of ice cream at different temperatures. We can easily see that warmer weather and higher sales go together. The relationship is good but not perfect. This is called Causation which has high positive correlation.

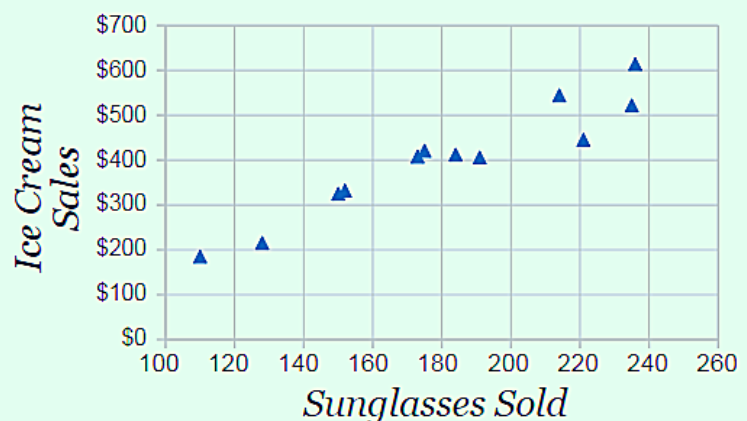
The correlation between Sunglasses and Ice Cream sales is high.



Does this mean that sunglasses make people want ice cream?



Hence Correlation doesn't always imply Causation.



# Imp points:

Correlation is Positive when the values increase together.

Correlation is Negative when one value decreases as the other increases.

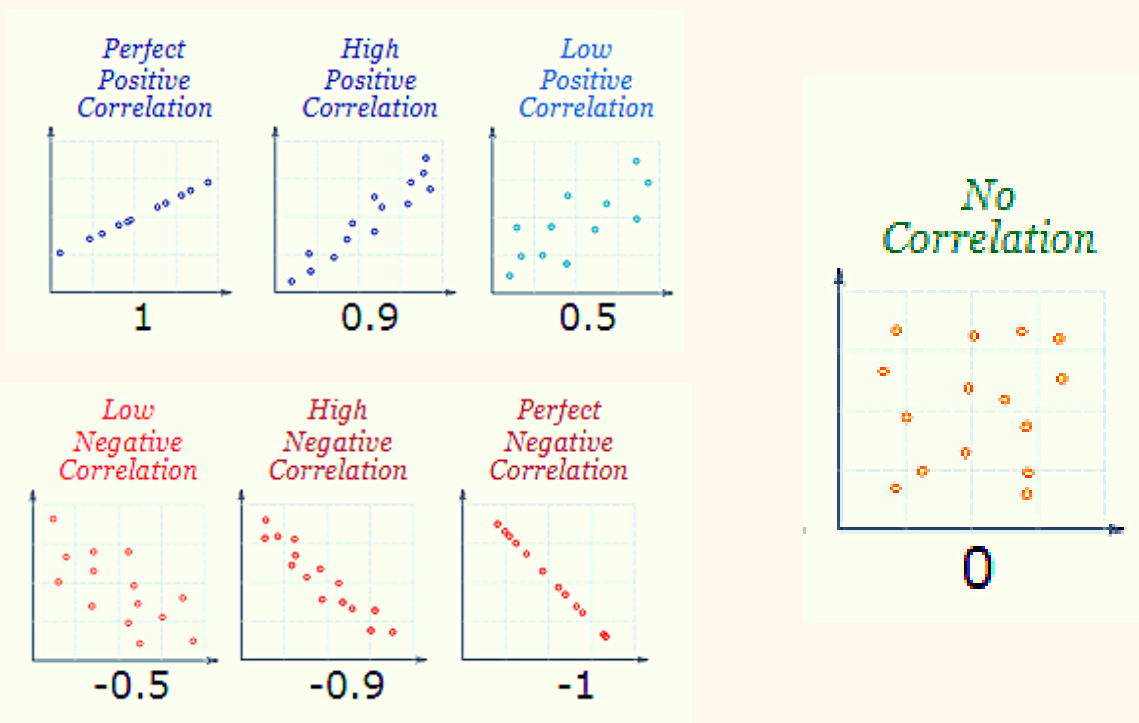
1 is a perfect positive correlation.

0 is no correlation (the values don't seem linked at all).

-1 is a perfect negative correlation.

Correlation Does not always imply Causation though it is possible in some cases.

Correlation measures how strongly two variables are related to each other



**SUBJECT** Correlation

**AUTHOR** SAMUEL