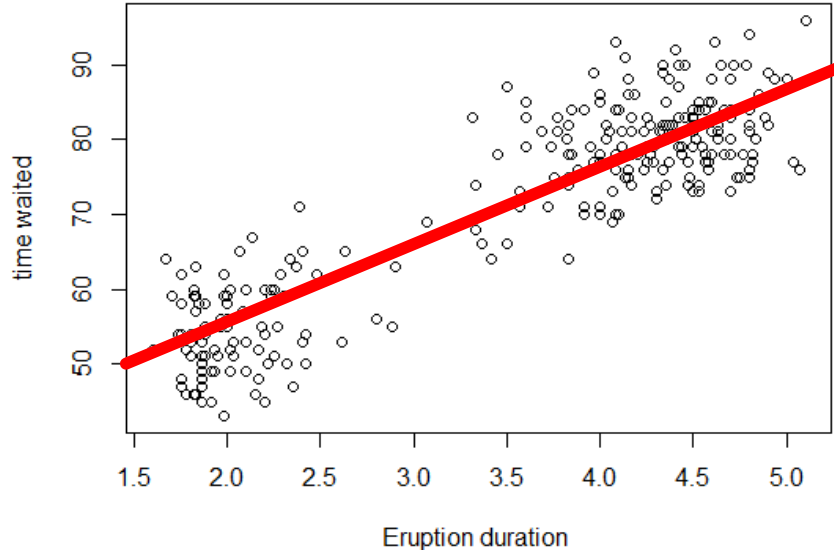# Scatter plot

A scatter plot pairs up values of two quantitative variables in a data set and display them as geometric points inside a Cartesian diagram.
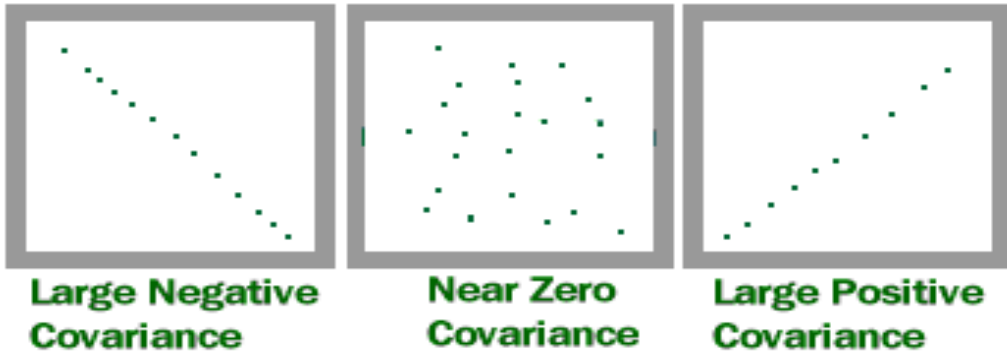

$$plot(x,y)$$

datDuration   datwait

abline(lm(datWait ~datDuration))

```
> datDuration=faithful$eruptions
>
> datWait=faithful$waiting
>
> plot(datDuration,datWait,xlab="Eruption duration ", ylab= "time waited")
>
```

Covariance

COVARIANCE

Large Negative Covariance   Near Zero Covariance   Large Positive Covariance

```
>
> cov(datDuration,datWait)
[1] 13.97781
>
>
> cov(datWait,datDuration)
[1] 13.97781
>
```

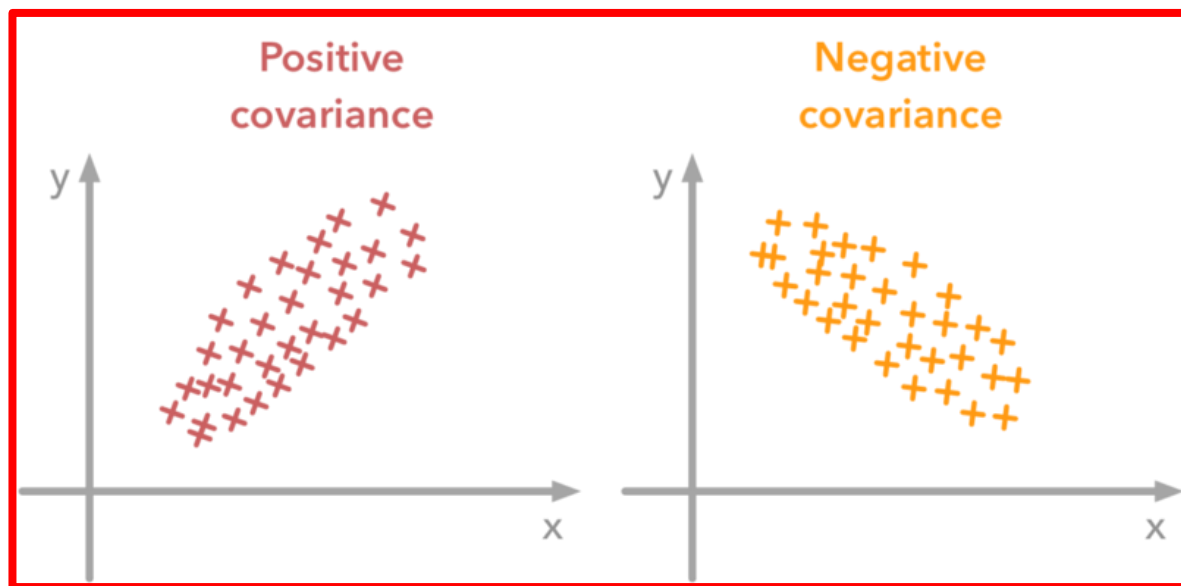$$\text{Cov}(X, Y) = \frac{\sum(X_i - \bar{X})(Y_j - \bar{Y})}{n}$$

the covariance can define three types of relationship

1) Relationship with positive trend
2) Relationship with negative trend
3) When there is no relationship because there is no trend in data

The order does not affect the result

↑X → ↑y

↓X → ↓y

↑X → ↓y
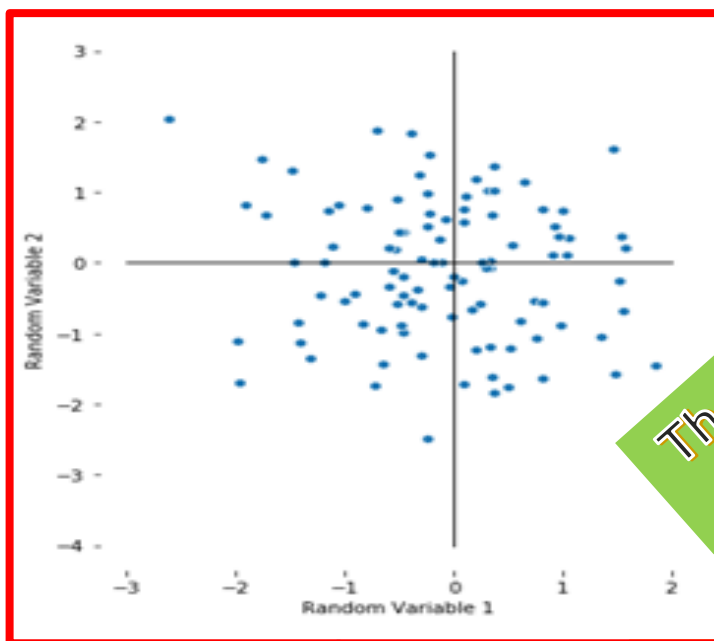
↓X → ↑y

Positive covariance

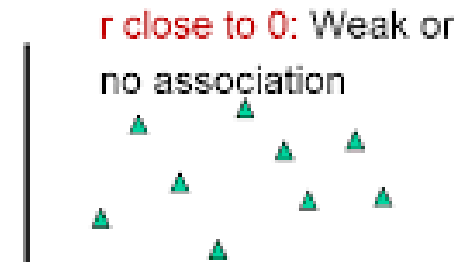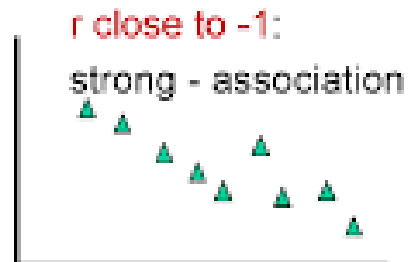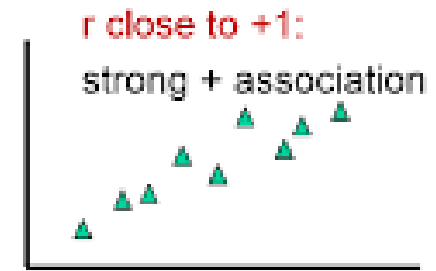Negative covariance

↑X → ↓↑y

↓X → ↓↑y

There is not a relatinoship

Correlation coefficient

The correlation coefficient = equals to their covariance divided by the product of their individual standard deviations.

It is a normalized measurement of how the two are linearly related.

$$COR(X, Y) = \frac{COV(X, Y)}{\sqrt{VAR(X)VAR(Y)}}$$

```
> cor(datWait,datDuration)
[1] 0.9008112
```

r = +1:
Perfect + correlation

r close to +1:
strong + association

r close to -1:
strong - association

r close to 0: Weak or
no association
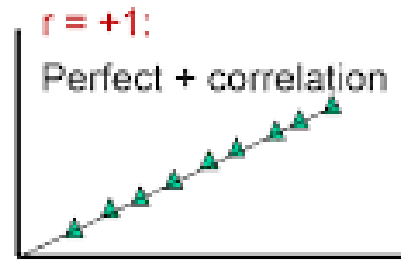
Direction and strength [-1,1]

Correlation is no causation
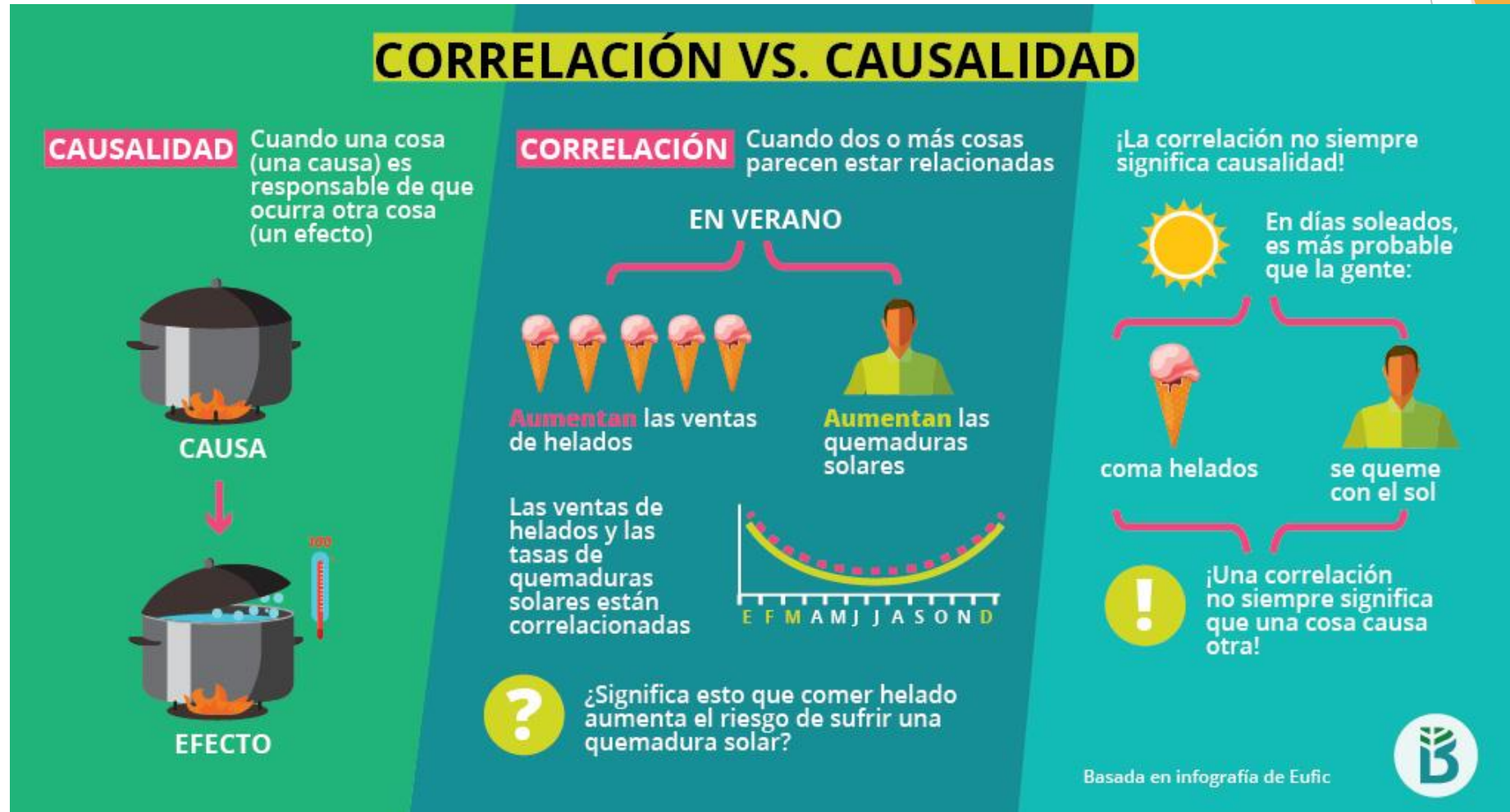
Correlation is no causation

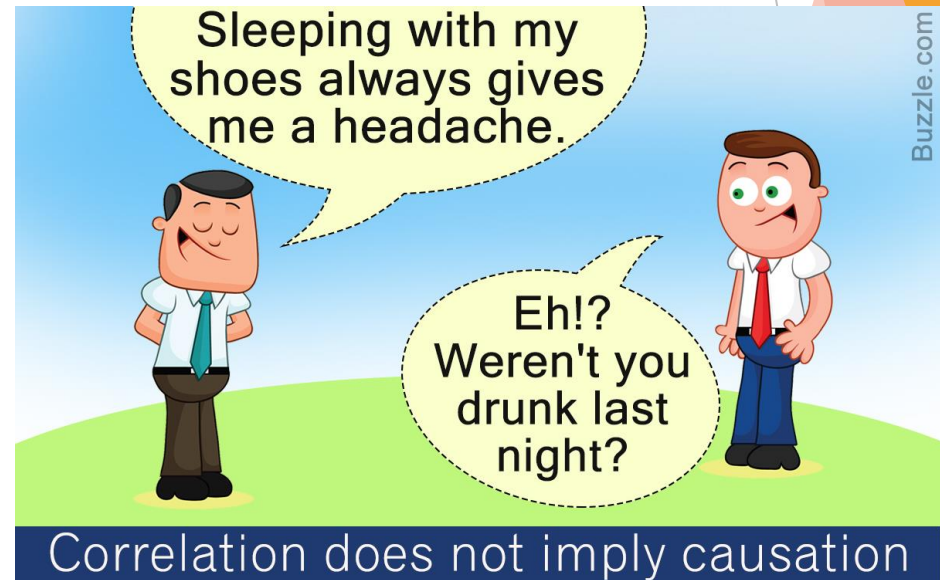The faster windmills are observed to rotate, the more wind is observed to be.

Therefore wind is caused by the rotation of windmills. (Or, simply put: windmills, as their name indicates, are machines used to produce wind.)

In this example, the correlation (simultaneity) between windmill activity and wind velocity does not imply that wind is caused by windmills. It is rather the other way around, as suggested by the fact that wind doesn't need windmills to exist, while windmills need wind to rotate. Wind can be observed in places where there are no windmills or non-rotating windmills—and there are good reasons to believe that wind existed before the invention of windmills.

When a country's debt rises above 90% of GDP, growth slows.

Therefore, high debt causes slow growth.

This argument by Carmen Reinhart and Kenneth Rogoff was refuted by Paul Krugman on the basis that they got the causality backwards: in actuality, slow growth causes debt to increase.

Coronavirus Causes GDP Contraction Around the World
2020 Q2 GDP growth in selected countries (in percent)

Country	Value
CHN	+11.5
KOR	-3.2
JAP	-7.9
USA	-9.1
GER	-9.7
FRA	-13.8
ESP	-18.5
GBR	-20.4
IND	-25.2

Quarter-on-quarter growth, not annualized
Source: OECD

statista
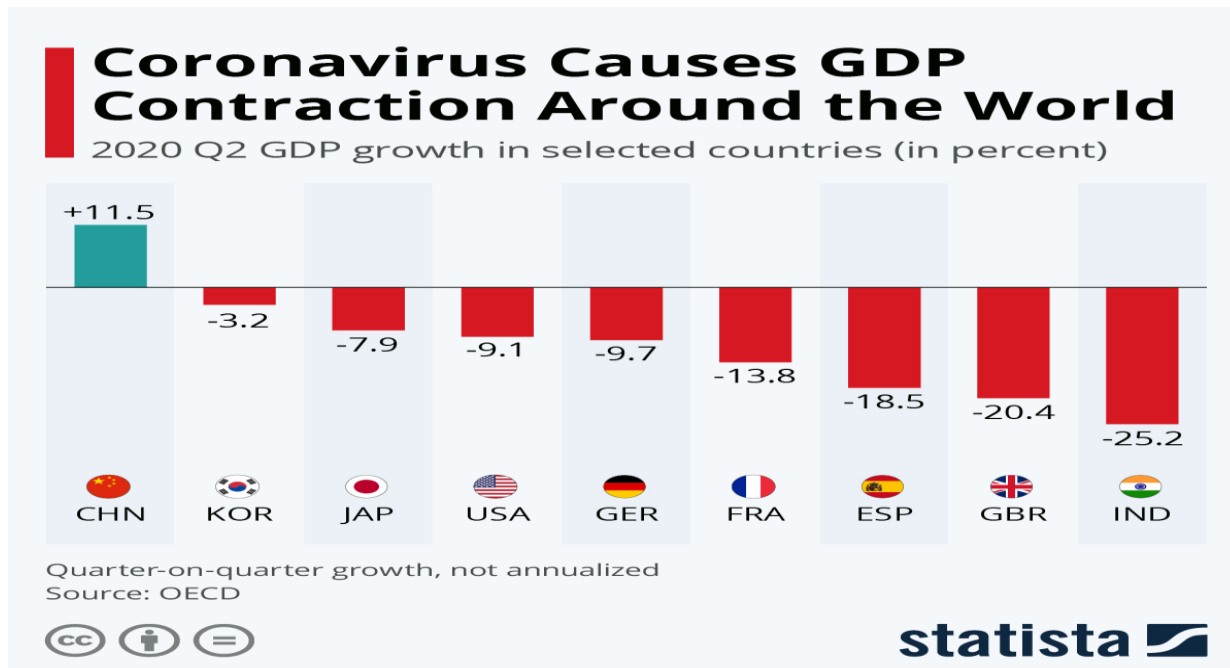
GDP Gross Domestic Product	GNP Gross National Product	GNI Gross National Income
Value of national output produced in a country. - National income - National output - National expenditure	GNP = GDP + net property income from abroad.	Similar to GNP - Sum value of output by resident producers + net receipts of primary income from abroad + any product taxes (less subsidies) not included in the valuation of output.
Includes income of foreign multinationals	Excludes income earnt by multinational when profit is sent back to other country	A country which earns positive net income on FDI will include in GNI, but not GDP

www.economicshelp.org

Example 1: Increase in product scans, rise in loyalty programs participants
For one of the NeuroTags' clients, we observed that the number of scans increased, followed by more customers claiming the loyalty points.

While investigating the cause, we found that earlier, the product code placement was not catching the buyer's attention. After tweaking the NeuroTags code placement to be more prominent, customers started to scan a lot more, and also claim loyalty points. Visibility of the tags was the cause for both the increase in scans and the increase in loyalty points claimed.