

## Savoir repérer et ne pas confondre corrélation et causalité

En Alsace, il existe une corrélation entre le nombre de naissances et le nombre de cigognes par village. En effet, plus il y a de cigognes dans un village, plus il y a de naissances, et inversement.

Est-ce à dire que cette corrélation est une causalité ? C'est-à-dire, est-ce que c'est le nombre de cigognes qui explique le nombre de naissances, autrement dit que ce sont les cigognes qui apportent les bébés ? Bien sûr que non !

Il y a une corrélation entre la variable A (nombre de cigognes) et la variable B (nombre de naissances) mais pas de causalité entre les 2.

Alors comment expliquer cette corrélation sans causalité ? Est-ce un heureux hasard ?

Il faut trouver, par le raisonnement, une troisième variable C, variable qui explique à la fois le nombre de cigognes et le nombre de naissances par village alsacien.

Cette troisième variable "cachée" est le nombre d'habitants par village. En effet, plus il y a d'habitants, plus il y a de maisons, donc de cheminées, donc de cigognes ; et, plus il y a d'habitants, plus il y a de couples et donc de naissances.

### A retenir :

On parle de corrélation positive quand 2 variables évoluent dans le même sens, et de corrélation négative quand 2 variables évoluent en sens opposé.

On parle de causalité quand 2 variables corrélées ont en plus un lien de cause à effet : A explique B ou B explique A.

Toutes les corrélations entre 2 variables ne sont pas des causalités.

Toutes les causalités entre 2 variables sont des corrélations.

