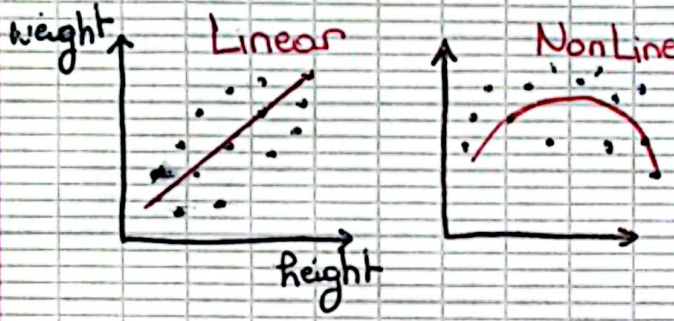
IV - Statistical Traps: How not to Fall in them?

→ Correlation IS NOT Causation

. Association: Measurements of one tells you something about the measurements of the other

. Correlation: How much 2 variables are associated

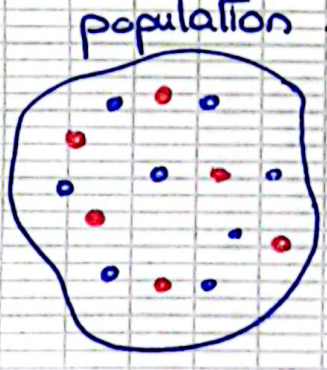


correlation coefficient
- between -1, 1
- w↗ h↗ + corr ≃ 1
- w↗ h↘ - corr ≃ -1
- = 0, no corr

. correlation ↛ causality
causality → correlation
(correlation is a necessary condition for causality but not a sufficient one)

. cyclic causation (both ways)
Depression causes → alcohol
alcohol causes → depression

→ Selection Bias and Data Attrition



. Selection Bias: Individuals we Sample differsystematically from the population
↳ To avoid it: make sure to take a random sample with respect to the question weire asking

- Attrition Bias (type of selection bias)
  ↳ Initialy random sample >> non-random subset of the sample is not taken into account

→ Importance of context
- To be Truthful, graphics must bear on the question: compared to what? → consider full historical context before drawing conclusion

→ Incorrect normalization of data
- Always check if data normalization is appropriate for fair comparisons
- Be cautious of changing conditions (e.g., inflation testing rates, ...) that affect interpretation
- Unnormalized data can lead to misleading conclusi

→ The Simpson's paradox
- Trend appears in several different groups of data but disappears or reverses when these groups are combined.
- Always visualize subgroup trends before drawing conclusions from combined data