Measures Of Shape

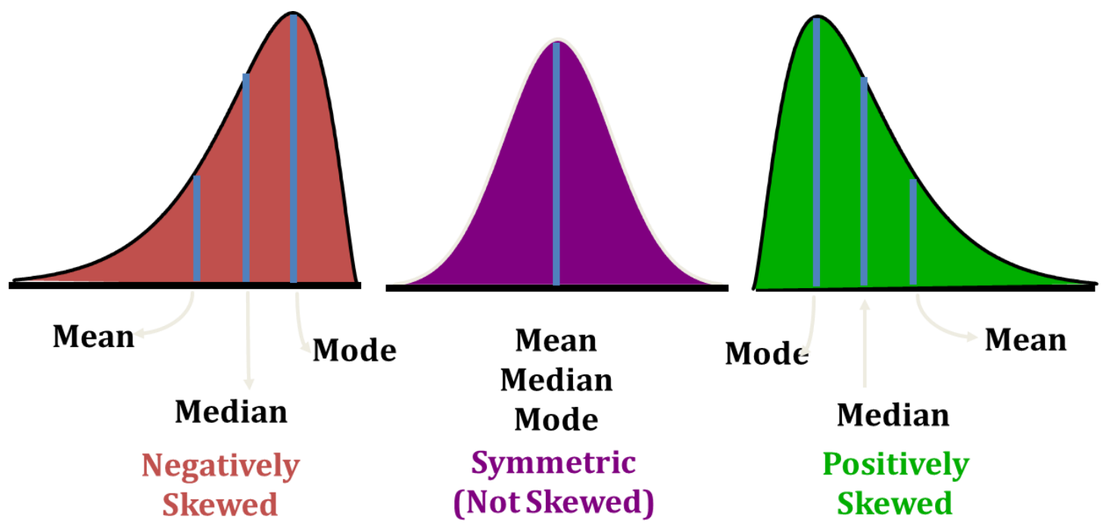
Measures Of Shape

- Skewness is a measure of symmetry or lack of symmetry

Skewness

- It can be described by its degree of asymmetry

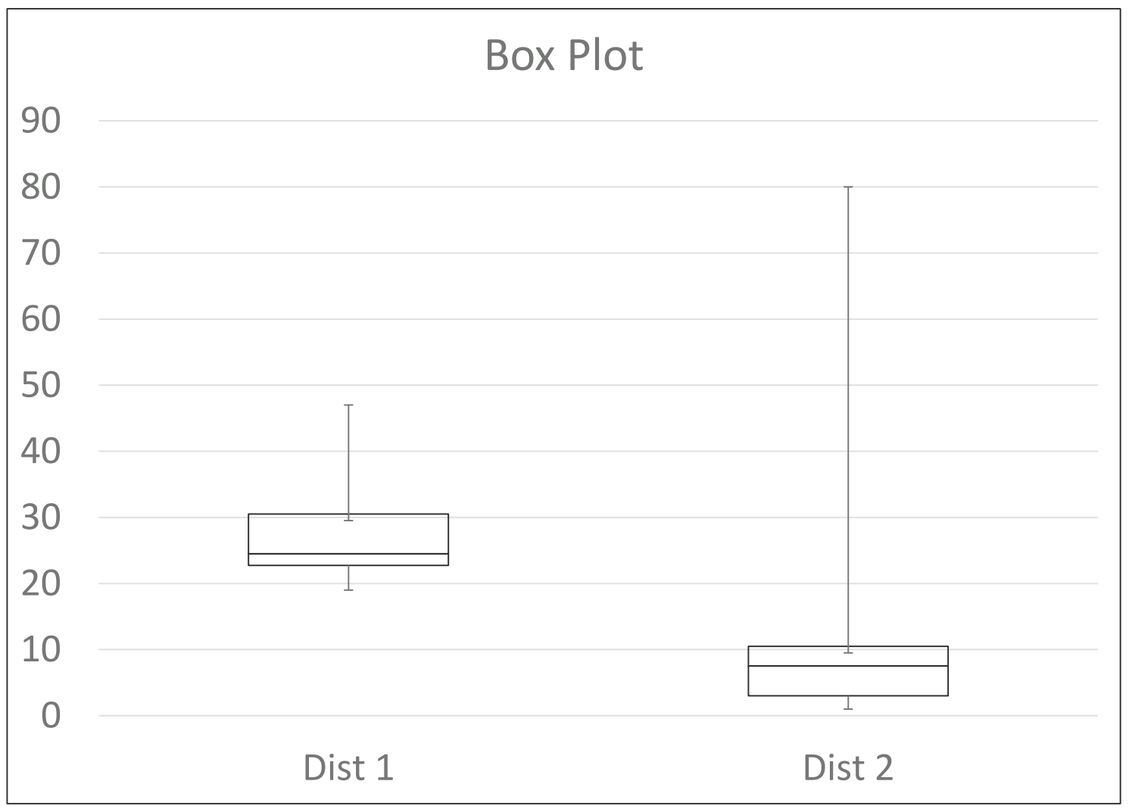
- $Mean > Median \rightarrow positive\ or\ right - skewness$
- $Mean = Median \rightarrow symmetry\ or\ zero - skewness$
- $Mean < Median \rightarrow negative\ or\ left - skewness$



- Positive skewness arises when the mean is increased by some unusually high values
- Negative skewness arises when the mean is decreased by some unusually low values

Box Plots

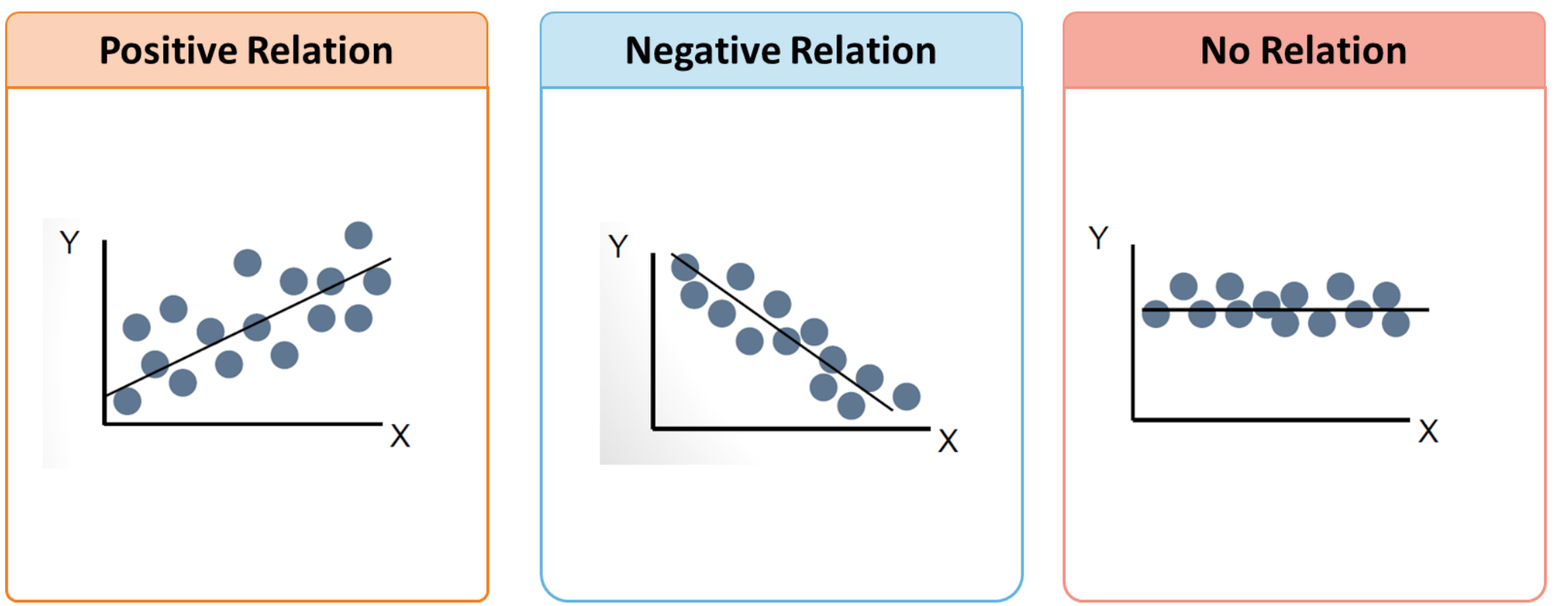
- Lower border at Q_1 or lower hinge
- Middle line is at Q_2
- Upper border at Q_3 or upper hinge
- Whiskers
- Outliers



Correlation - How to QUANTIFY relationship between variables?

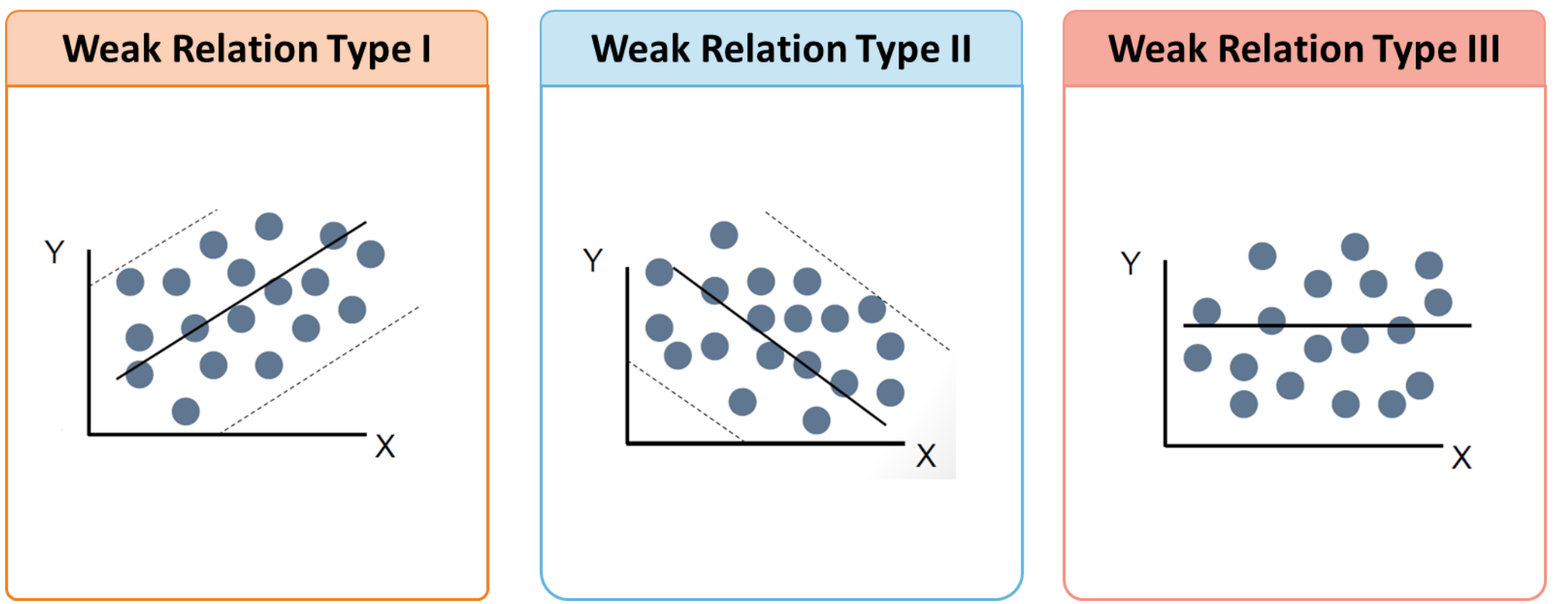
Correlation

- X and Y can exist in three different types of relations



Correlation

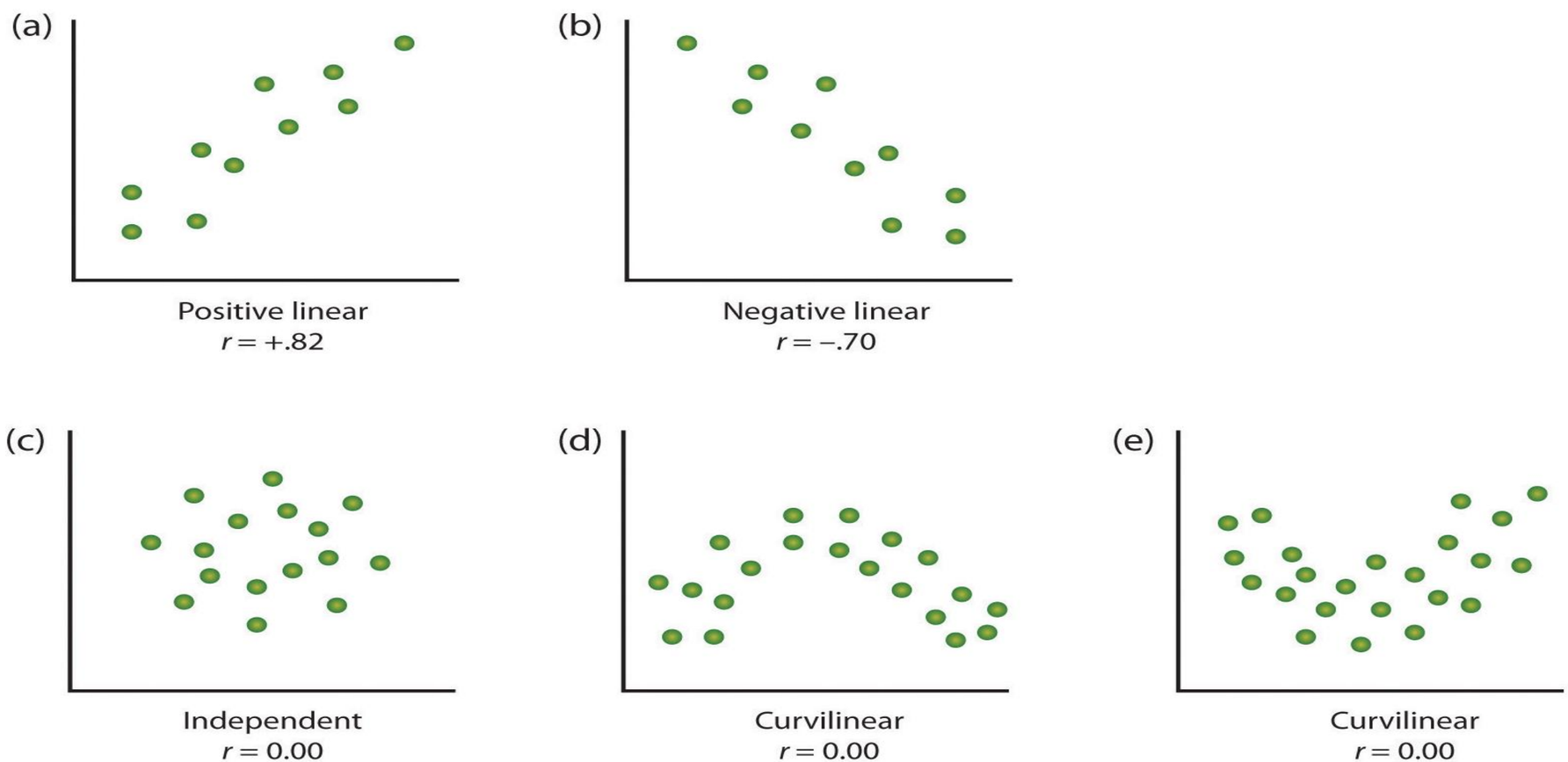
- They can also exist in a weak relation



Correlation

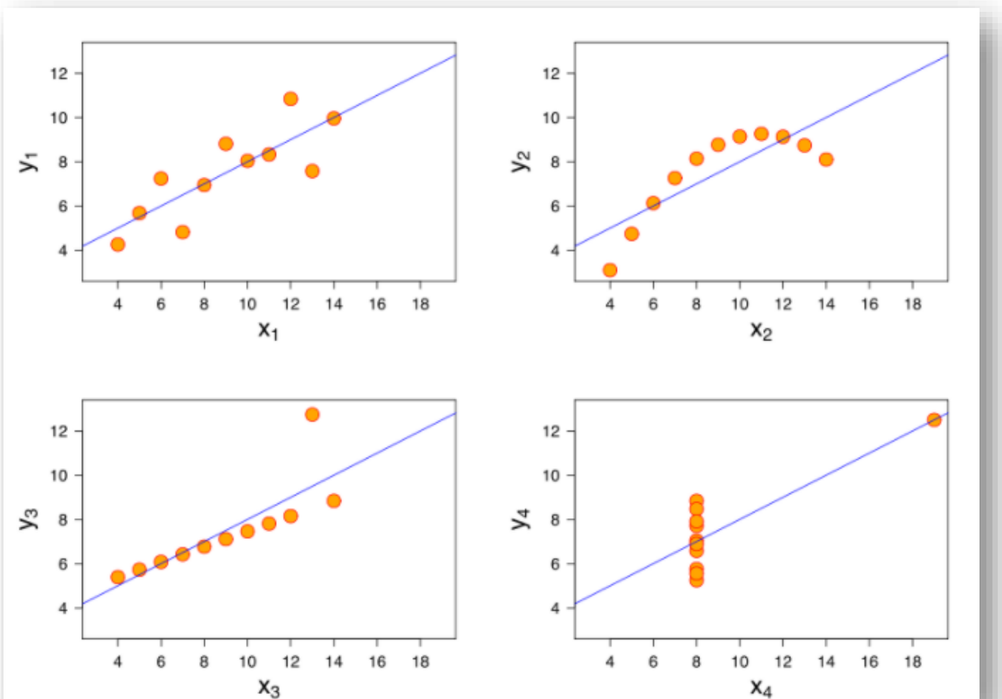
- Correlation is a statistical technique that predicts whether and how strongly pairs of variables are related.
  - The main result of a correlation is called the correlation coefficient (or "r"). It ranges from -1.0 to +1.0. The closer r is to +1 or -1, the more closely the two variables are related.
  - If r is close to 0, it means there is no relationship between the variables
  - If r is positive, it means that as one variable gets larger the other gets larger
  - If r is negative, it means that as one gets larger the other gets smaller (often called an "inverse" correlation)

Correlation



Correlation

- Four sets of data with the same correlation of 0.816



By Anscombe.svg: SchutzDerivative works of this file:(label using subscripts): Avenue - Anscombe.svg, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=9838454