Introduction to correlation

Introduction to correlation

We often want to investigate the relationship between pairs of variables.

- Vegetation height and mean annual temperature
- Animal body size and metabolic rate
- Number of automobiles in a location and carbon emissions

The **correlation** between pairs of variables, such as those listed above, describes how the variation of each variable is related to the other variable.

Visualising the correlation between two variables

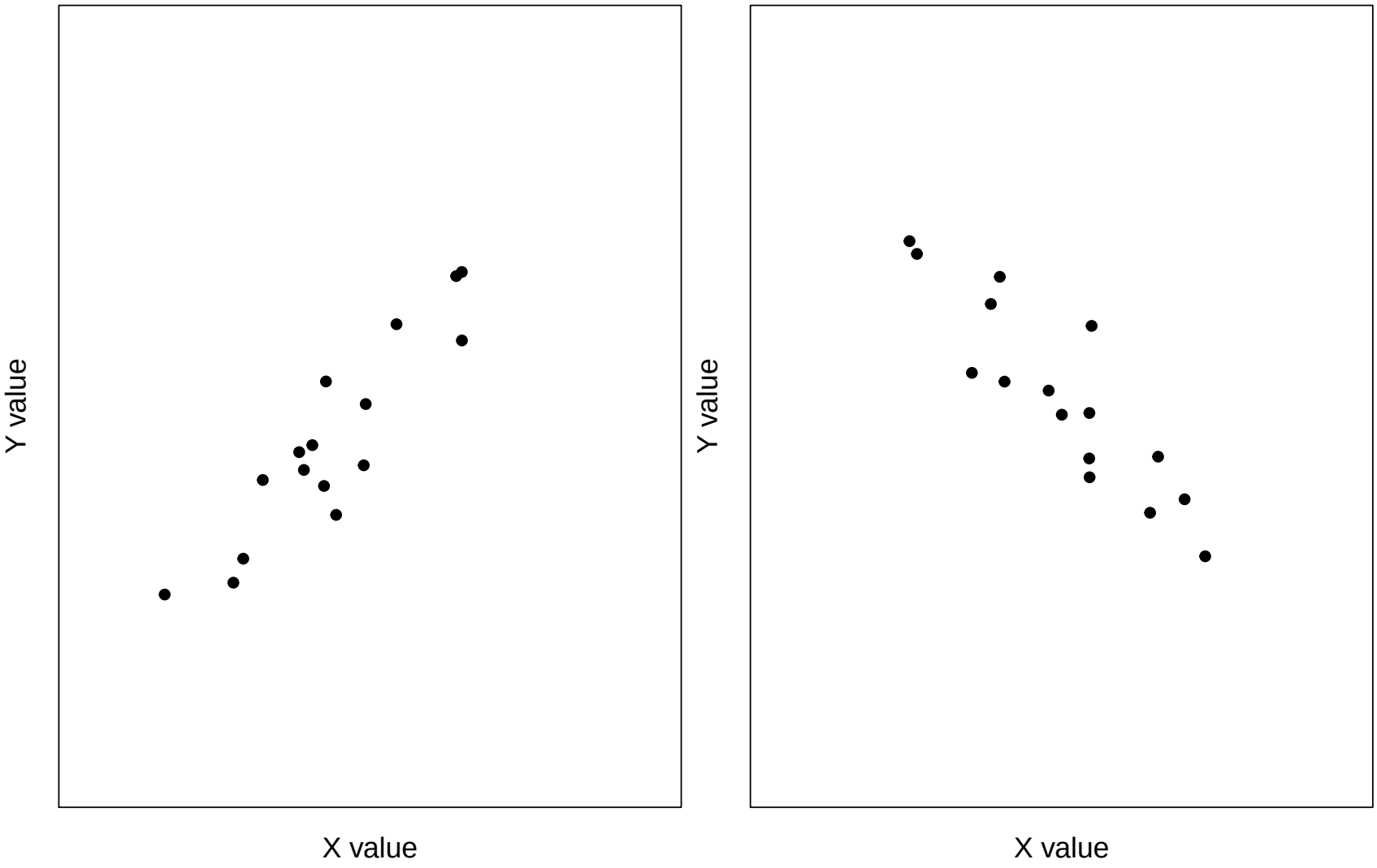

Figure 1: Two plots of hypothetical variables illustrating a positive (left) and negative correlation

Visualising two variables that are not correlated

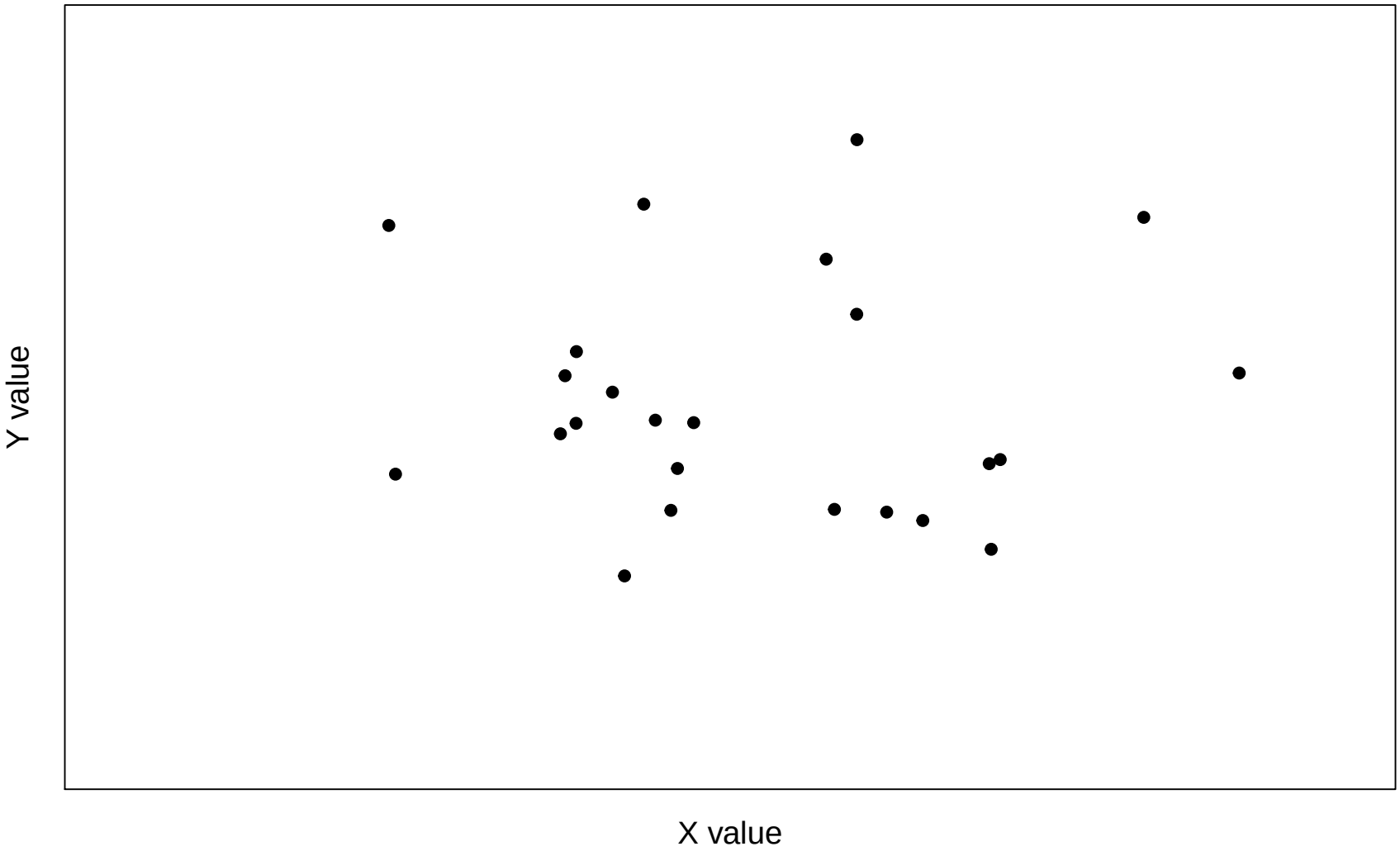

Figure 2: A plot of two hypothetical variables that are not correlated.

Getting a more intuitive sense of correlation

Formalised with the **correlation coefficient** (r)

- Provides a statistical measure of strength and direction of correlation
- Only describes association between variables (**not** cause and effect)

The value r ranges between -1 and 1

- Negative numbers indicate a negative correlation
- Positive numbers indicate a positive correlation
- Avalue of zero indicates no correlation

We can get a more intuitive understanding of the correlation coefficient with **[this application]**.