

## 5. Types of variables

A variable is any property that is measured in an observation ([Sokal and Rohlf, 1995](#)); i.e., anything that varies among things that we can measure ([Dytham, 2011](#)). We can summarise how these measurements vary with summary statistics, or visually with figures. Often, we will want to predict one variable from a second variable. In this case, the variable that we want to predict is called the **response variable**, also known as the **dependent variable** or **Y** variable (‘dependent’ because it *depends* on other variables, and ‘Y’ because this is the letter we often use to represent it). The variable that we use to predict our response variable is the **explanatory variable**, also known as the **independent variable** or **X** variable (‘independent’ because it does not depend on other variables, and ‘X’ because this is the letter most often used to represent it). There are several different types of variables:

- **Categorical** variables take on a fixed number of discrete values ([Spiegelhalter, 2019](#)). In other words, the measurement that we record will assign our data to a specific category. Examples of categorical variables include species (e.g., “Robin”, “Nightingale”, “Wren”) or life history stage (e.g., “egg”, “juvenile”, “adult”). Categorical variables can be either nominal or ordinal.
  - **Nominal** variables do not have any inherent order (e.g., classifying land as “forest”, “grassland”, or “urban”).
  - **Ordinal** variables do have an inherent order (e.g., “low”, “medium”, and “high” elevation).
- **Quantitative** variables are variables represented by numbers that reflect a magnitude. That is, unlike categorical variables, we are collecting numbers that really mean something tangible (in contrast, while we might represent low, medium, and high elevations with the numbers 1, 2, and 3, respectively, this is just for convenience; a value of ‘2’ does not always mean ‘medium’ in other contexts). Quantitative variables can be either discrete or continuous.
  - **Discrete** variables can take only certain values. For example, if we want to record the number of species in a forest, then our variable can only take discrete counts (i.e., integer values). There could conceivably be any natural number of species (1, 2, 3, etc.), but there could not be 2.51 different species in a forest; that does not make sense.

## 5. Types of variables

- **Continuous** variables can take any real value within some range of values (i.e., any number that can be represented by a decimal). For example, we could measure height to as many decimals as our measuring device will allow, with a range of values from zero to the maximum possible height of whatever it is we are measuring. Similarly, we could measure temperature to any number of decimals, at least in theory, so temperature is a continuous variable.

The reason for organising variables into all of these different types is that different types of variables need to be handled in different ways. For example, it would not make sense to visualise a nominal variable in the same way as a continuous variable. Similarly, the choice of statistical test to apply to answer a statistical question will almost always depend on the types of variables involved. If presented with a new data set, it is therefore very important to be able to interpret the different variables and apply the correct statistical techniques (this will be part of the assessment for this module).