



Categorical data

Data are described as categorical or qualitative when they can be categorised into distinct groups, such as ethnic group or disease severity. Although categorical data may be coded numerically, for example gender may be coded 1 for male and 2 for female, these codes have no intrinsic numerical value; it would be nonsense to calculate an average gender. Categorical data can be divided into either **nominal** or **ordinal**. Nominal data have no natural ordering and examples include eye colour, marital status and area of residence. Binary data is a special subcategory of nominal data, where there are only two possible values, for example (male/female, yes/no, treated/not treated). Ordinal data occurs when there can be said to be a natural ordering of the data values, such as better/same/worse, grades of breast cancer or social class.

Numerical data

Numerical data are data that can be quantified. Numerical data can be either count or continuous (measured). **Count** data are also known as **discrete data** and occur when the data can only take whole numbers, such as the number of children in a family or the number of visits to a GP in a year. **Continuous** data are data that can be measured and they can take any value on the scale on which they are measured; they are limited only by the scale of measurement and examples include height, weight, and blood pressure.