# Week 1 glossary

**Every discipline has a vocabulary, and sometimes words have a technical meaning which is different from the everyday meaning we may be used to. Data Analysis is no exception.**

(Different subject-fields sometimes use different names for the same statistical or data-analysis idea. This is why you will often see “Alternative names: …” at the end of a definition. These are not things you need to know for the course. They’re just there in case you already know something by another name.)

Here are the words that we encounter this week. You can also [print the full glossary](https://www.stat.auckland.ac.nz/~wild/d2i/articles/data-to-inzight-glossary.pdf" \o "Download the glossary) of words and refer to them during the course. This glossary is provided for reference, not as definitions to be learned. New words are added week by week in the week they first appear.

**Categorical variable**: A variable whose values are names or codes for different groups (or categories). (Alternative names: qualitative variable, factor, class variable.) (Special cases: nominal variable, ordinal variable.)

**Entities**: The individual “things” we are recording data about. (Alternative names: individuals, units, cases, instances.)

**Frequency**: The number of times a value of a variable, or a category, occurs. (Alternative names: count, tally.)

**Missing value code**: A code that is used to tell the software program that no value has been recorded for this cell (either an empty cell, NA or NULL for iNZight).

**Missing value**: No information has been recorded for this cell.

**Numeric variable**: A variable for which all of the values are numbers (e.g. from counting or measuring). (Alternative name: quantitative variable.) (Special cases: discrete variable, continuous variable, interval variable, ratio variable.)

**Rectangular data**: Data organised and stored in such a way that each row corresponds to an individual entity and each column corresponds to a property recorded for these entities.

**Variable**: A property that we record for each entity, e.g. a measurement, or one of a set of group labels. (Alternative names: attribute, feature [in machine learning]).