



# Dataset

## 1 Why

We want language to talk about a list of elements from a set.

## 2 Definition

A *dataset* is a finite sequence of elements from a set. We say that the data set is *in* the set. We refer to the length of the sequence as the *size* of the data set and we refer to an element of the sequence as a *record* of the data set.

### 2.1 Notation

Let  $A$  be a non-empty set. We usually denote sequences by  $(a_1, \dots, a_n)$ , but for data sets we will use a superscript,  $(a^1, \dots, a^n)$ . Often  $A$  is itself a set of sequences, and we will want to refer to the  $i$ th component of  $a^1$  by  $a_i^1$ .