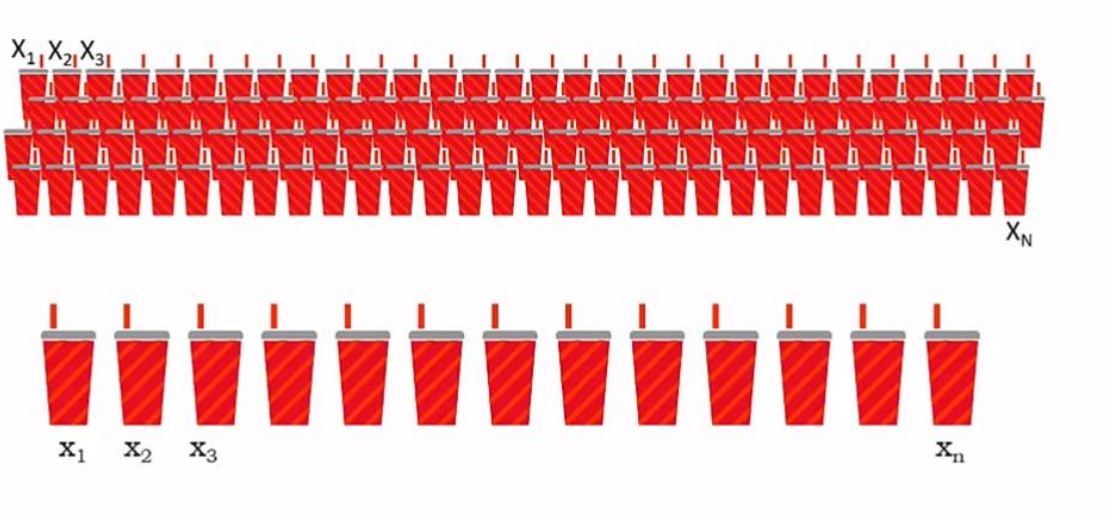
Types of variables

1. **Continuous variables** – the numbers which are continuous can be measured on a scale (temperature),
2. **Discrete** – the number which we count, rather than measure (leaflets) and
3. **Categorical = strings** – we had two selling destinations or categories of sales (beach and park).

Regarding most statistical functions, data scientist prefer to work with numbers, rather than these categorical words or strings. So, we can add some numerical labels to those variables. Those are called **dummy variables**. For example, sales on the beach we would mark with 1 and sales in the park with 2.

Population vs. sample



The big X represents a population, the little one represents a sample of data.

N is the total number of items in the population.

n is the total number of data in the sample.

It is important to use the **representative sample**, which represents the characteristics of the population in the right way.

Greek symbols -> population

Non-Greek symbols -> sample