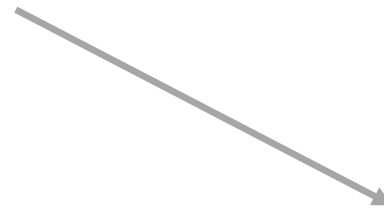# Tabular Data

The primary way that we will organize data is called the tabular data model. Data are arranged into a grid of rows and columns.

Tabular Data

The primary way that we will organize data is called the tabular data model. Data are arranged into a grid of rows and columns.

the individual entries are called **values**

breed	weight	height
Shih-Tzu	5.5	24
Labrador	33	56
Beagle	10.2	34
Newfoundland	70	69
Chihuahua	1.3	20
Affenpinscher	9.6	27

Tabular Data

The primary way that we will organize data is called the tabular data model. Data are arranged into a grid of rows and columns.

the individual entries are called **values**

Rows ⇒ Observations

each row represents one observation

breed	weight	height
Shih-Tzu	5.5	24
Labrador	33	56
Beagle	10.2	34
Newfoundland	70	69
Chihuahua	1.3	20
Affenpinscher	9.6	27

Tabular Data

The primary way that we will organize data is called the tabular data model. Data are arranged into a grid of rows and columns.

Columns ⇒ Features

each feature has name and a data type

the individual entries are called **values**

Rows ⇒ Observations

each row represents one observation

breed	weight	height
Shih-Tzu	5.5	24
Labrador	33	56
Beagle	10.2	34
Newfoundland	70	69
Chihuahua	1.3	20
Affenpinscher	9.6	27

Tabular Data

Each feature (column) in the tabular data model has a **data type** associated with it. This is not always explicit in way the dataset is saved, but will be defined when we are working with it in R.

The two most common data types we will see are:

> ↣ **numeric:** everything can be represented by a number
> ↣ **character:** arbitrary sequences of any characters

There is a single type for each feature; we cannot mix and match data types. We will see other data types as they arise in our work.

Tabular Data

Implied data types in our example:

character **\<chr\>**	numeric **\<dbl\>**	numeric **\<int\>**
↓	↓	↓
breed	**weight**	**height**
Shih-Tzu	5.5	24
Labrador	33	56
Beagle	10.2	34
Newfoundland	70	69
Chihuahua	1.3	20
Affenpinscher	9.6	27