**Descriptive Analysis**

4 min

Let's imagine we're looking at a field of flowers like the one in the learning environment. There are flowers with different shapes, colors, and heights. If you had to describe these flowers, how would you quantify all these characteristics? What does a typical blue flower look like? Are blue flowers all similar in height? It's hard to say just looking at a mix of all the flowers, just like it's hard to describe a big dataset just by looking at a spreadsheet of values.

**Descriptive analysis** lets us describe, summarize, and visualize data so that patterns can emerge. Sometimes we'll only do a descriptive analysis, but most of the time a descriptive analysis is the first step in our analysis process.

Descriptive analyses include measures of central tendency (e.g., mean, median, mode) and spread (e.g., range, quartiles, variance, standard deviation, distribution), which are referred to as **descriptives** or **summary statistics**.

Typically, data visualization is also included in descriptive analysis.

Descriptive analysis helps us understand our data and decide what steps to take next, but we cannot extend what we learn to other datasets.

**Instructions**

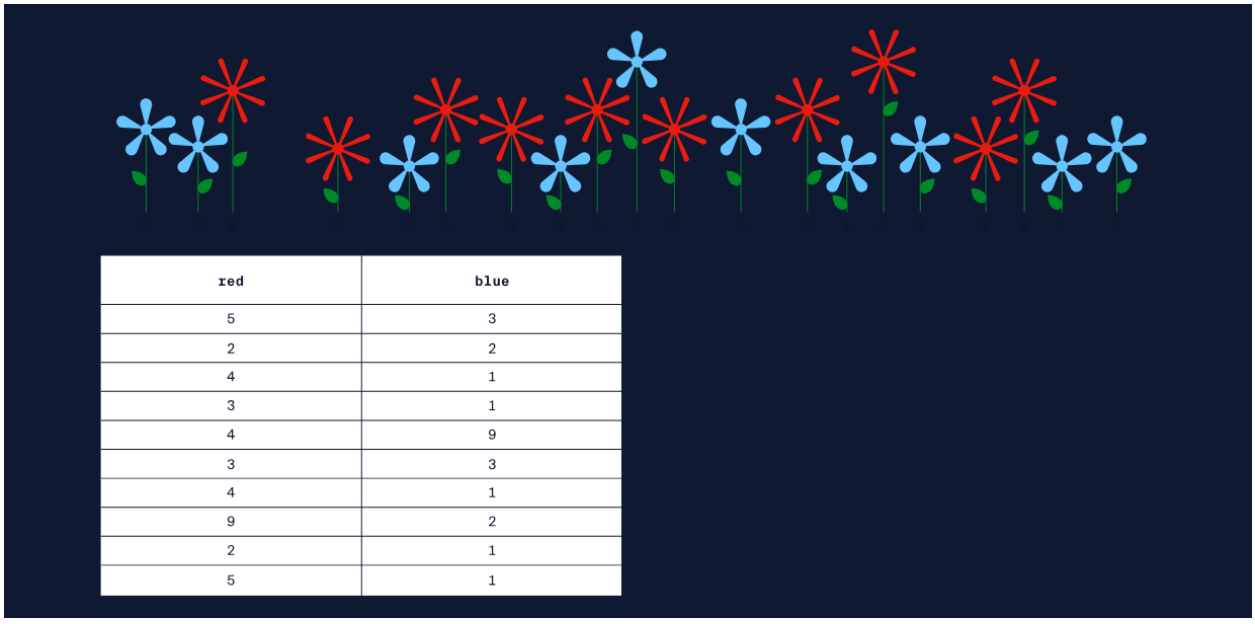
Take a look at the different stages of descriptive analysis of the flowers in the learning environment by selecting 'Next' or playing the slides using the center button. Consider the following:

1. Does calculating the mean and median height for blue and red flowers make it easier to see trends in the data? What about the boxplot?
2. Which part of the descriptive analysis shown is most useful?
3. Based on our analysis of red and blue flowers, can we conclude that red flowers are taller than blue flowers *everywhere*?

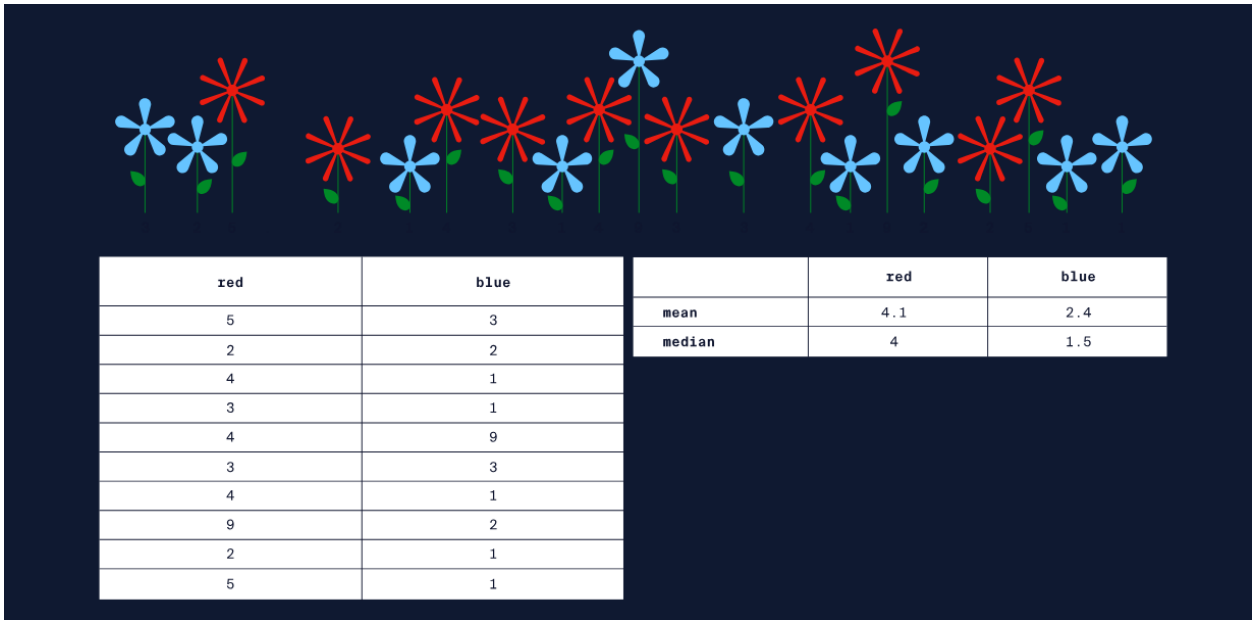
Remember:

- Descriptive analyses let us draw basic conclusions about the data at hand.
- We cannot extend findings from descriptive analyses to other datasets.

In the next exercise, we will learn about another type of data analysis: exploratory analysis!

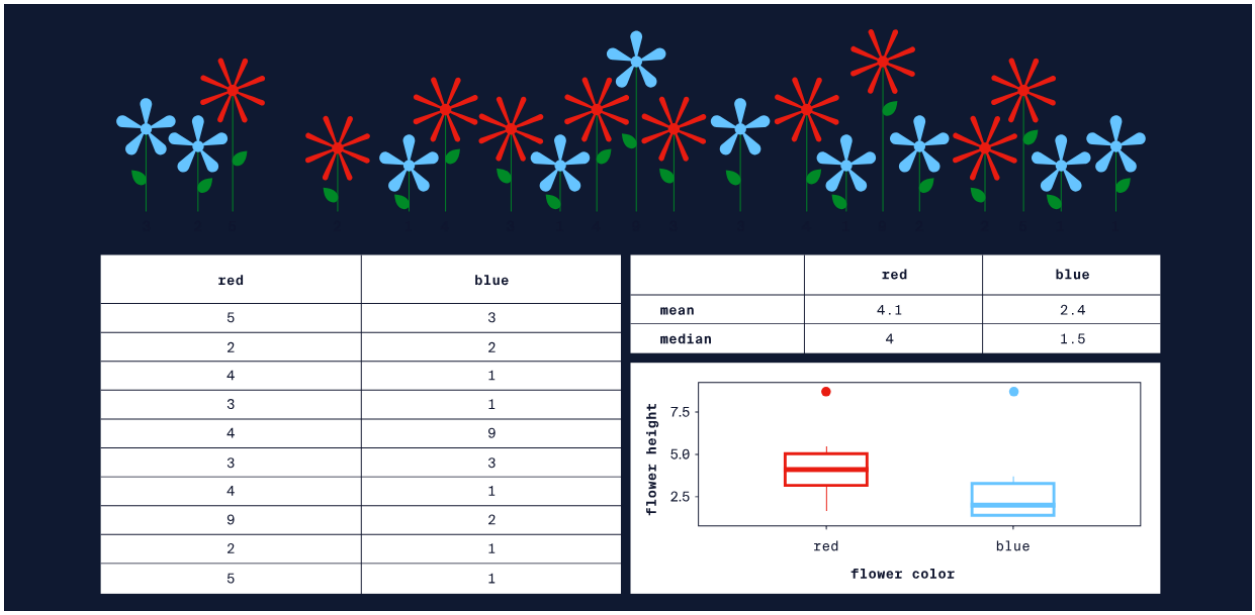


| red | blue |
|---|---|
| 5 | 3 |
| 2 | 2 |
| 4 | 1 |
| 3 | 1 |
| 4 | 9 |
| 3 | 3 |
| 4 | 1 |
| 9 | 2 |
| 2 | 1 |
| 5 | 1 |



| red | blue |
|---|---|
| 5 | 3 |
| 2 | 2 |
| 4 | 1 |
| 3 | 1 |
| 4 | 9 |
| 3 | 3 |
| 4 | 1 |
| 9 | 2 |
| 2 | 1 |
| 5 | 1 |

| | red | blue |
|---|---|---|
| mean | 4.1 | 2.4 |
| median | 4 | 1.5 |



| red | blue |
|---|---|
| 5 | 3 |
| 2 | 2 |
| 4 | 1 |
| 3 | 1 |
| 4 | 9 |
| 3 | 3 |
| 4 | 1 |
| 9 | 2 |
| 2 | 1 |
| 5 | 1 |

| | red | blue |
|---|---|---|
| mean | 4.1 | 2.4 |
| median | 4 | 1.5 |