

# Scatter Plots and Two-Way Tables: math lesson, Grade 8 – download the PDF

Do taller people really have longer arms? Does more studying lead to better scores? To answer questions like these, you need to look at **two** pieces of information about each person or object at the same time. In this chapter you will draw scatter plots, spot patterns, build a line that models the data, and organize categories in two-way tables.

## 1. Bivariate data and scatter plots

When you record two numerical measurements for each individual (for example, hours studied and test score), you have **bivariate data**. Each individual gives an ordered pair  $(x, y)$ .

### **D** Definition: scatter plot

A **scatter plot** is a graph in the coordinate plane where each ordered pair of bivariate data is shown as one point. The **independent variable** (the one you think may influence the other) goes on the horizontal axis, and the **dependent variable** goes on the vertical axis.

**M****Method: construct a scatter plot**

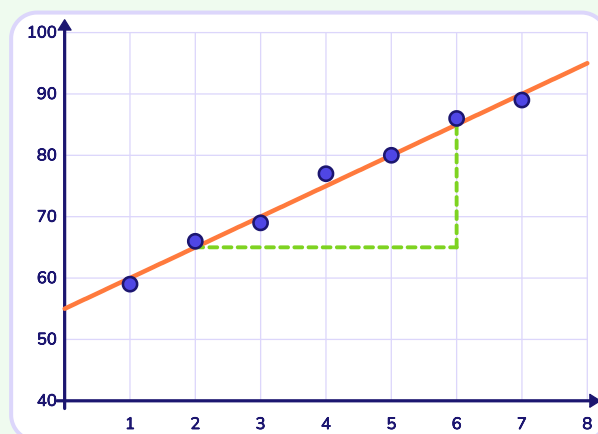
1. Choose the variable for each axis.
2. Pick a scale that fits the smallest and largest values, with equal spacing between tick marks.
3. Plot one point for each pair. Do not connect the points.
4. Label both axes with names and units, and give the graph a title.

**E****Example 1**

Seven students recorded hours studied for a quiz and their score out of 100.

| Hours studied $x$ | 1  | 2  | 3  | 4  | 5  | 6  | 7  |
|-------------------|----|----|----|----|----|----|----|
| Score $y$         | 59 | 66 | 69 | 77 | 80 | 86 | 89 |

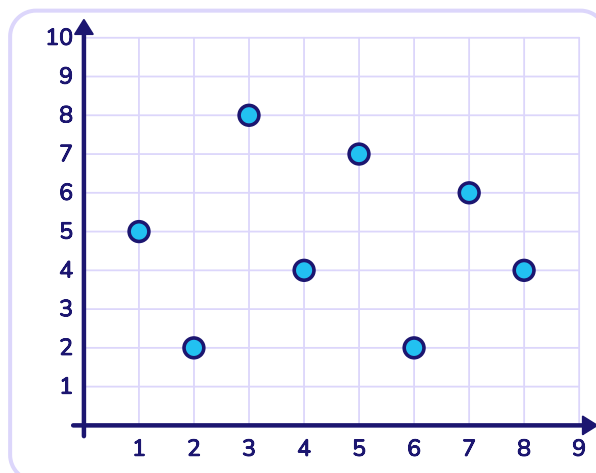
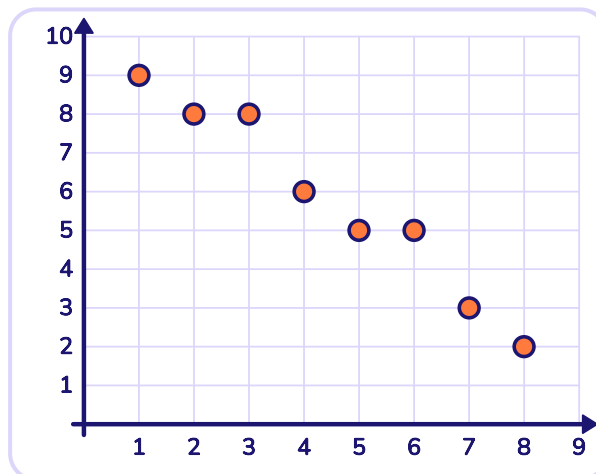
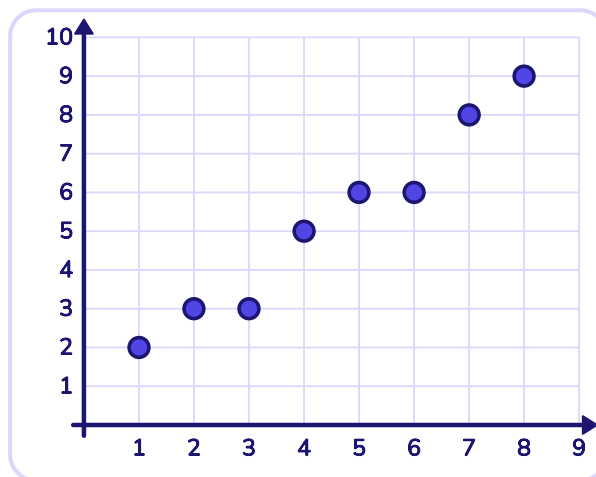
The first point is  $(1, 59)$ : move 1 unit right and 59 units up. The seven points are shown below.



## 2. Patterns: positive, negative, or no association

The shape of the cloud of points tells you how the two variables are related. This relationship is called an **association**.

- **Positive association:** as  $x$  increases,  $y$  tends to increase.
- **Negative association:** as  $x$  increases,  $y$  tends to decrease.
- **No association:** the points show no clear trend.



If the points stay close to a straight line, the association is **linear**. If they follow a curve instead, it is **nonlinear**. The closer the points are to the line, the stronger the association.

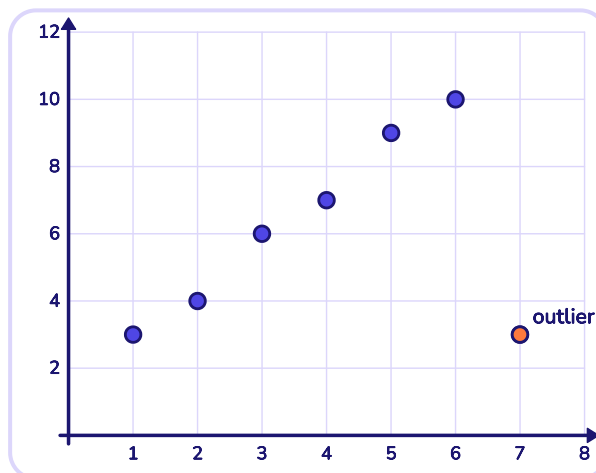


### Careful

An association does not prove that one variable *causes* the other. Ice cream sales and sunburns both rise in summer, but ice cream does not burn anyone. Hot, sunny weather affects both.

## 3. Outliers and clusters

An **outlier** is a point that is far away from the overall pattern of the other points. A **cluster** is a group of points that sit close together, often separated from other groups by a **gap**.



### Example 2

In the graph above, six points rise steadily, but the point **(7, 3)** falls far below the trend. It is an outlier. Always ask why: a measuring mistake, a special event, or a truly unusual individual?

An outlier can pull a trend line toward itself and make predictions less accurate. Clusters can reveal that your data comes from two different groups.

## 4. The line of best fit

When a scatter plot shows a linear association, you can draw a straight line that models the trend. This is called a **line of best fit** or **trend line**.

### M Method: draw a line of best fit by eye

1. Use a clear straightedge.
2. Place it so it follows the direction of the points.
3. Try to have about the same number of points above and below the line.
4. The line does not have to pass through any data point, and it does not have to start at the origin.

In the graph of Example 1, the orange line passes through  $(2, 65)$  and  $(6, 85)$ . These two points are on the *line*, so you can use them to find its equation.

### E Example 3: find the equation

$$\text{Slope: } m = \frac{85 - 65}{6 - 2} = \frac{20}{4} = 5.$$

Use  $y = 5x + b$  with the point  $(2, 65)$ :  $65 = 5(2) + b$ , so  $b = 55$ . The model is  $y = 5x + 55$ .

## 5. Using a linear model to predict

Once you have an equation, substitute a value of  $x$  to **predict**  $y$ .

**E****Example 4**

Predict the score of a student who studies 4.5 hours:  $y = 5(4.5) + 55 = 77.5$ . The model predicts about 77.5 points.

Predicting inside the range of your data is called **interpolation** and is usually reliable. Predicting far outside the range is **extrapolation**, which can be risky. For 10 hours, the model gives  $5(10) + 55 = 105$ , which is impossible on a test scored out of 100.

**6. Interpreting slope and intercept in context**

In a model  $y = mx + b$ , the numbers have a meaning in the real situation.

**P****Rule**

- The **slope**  $m$  is the predicted change in  $y$  for each increase of 1 unit in  $x$ .
- The **y-intercept**  $b$  is the predicted value of  $y$  when  $x = 0$ .

For  $y = 5x + 55$ : each extra hour of study is associated with about 5 more points. A student who studied 0 hours is predicted to score about 55. Be careful: sometimes the intercept does not make sense in context, for example when  $x = 0$  is far from the data.

**!****Zyro says**

On my planet we write units next to every number. Say "5 points per hour," not just "5." The units make your interpretation clear.

## 7. Two-way frequency tables

Scatter plots show two *numerical* variables. To study two *categorical* variables (such as yes/no answers), use a **two-way frequency table**. Rows show one variable, columns show the other, and each cell counts the individuals in both categories.

### E Example 5

One hundred twenty eighth graders were asked how long they sleep and whether they use a phone in bed.

|                        | Phone in bed | No phone | Total |
|------------------------|--------------|----------|-------|
| Sleeps under 8 hours   | 45           | 15       | 60    |
| Sleeps 8 hours or more | 20           | 40       | 60    |
| Total                  | 65           | 55       | 120   |

Each **total** (marginal frequency) is the sum of its row or column. The number 45 is a **joint frequency**: the students who sleep under 8 hours *and* use a phone in bed.

## 8. Relative frequencies and association

A **relative frequency** is a count divided by a total, usually written as a percent. To check for association, compare relative frequencies *within each row* (or each column).

**E****Example 6**

Among the 60 students who sleep under 8 hours,  $\frac{45}{60} = 75\%$  use a phone in bed.

Among the 60 who sleep 8 hours or more,  $\frac{20}{60} \approx 33.3\%$  do. These percents are very different, so there is an **association** between sleeping less and using a phone in bed.

**P****Rule**

If the row relative frequencies are about equal, there is no association. If they differ noticeably, there is an association.

**⚠****Careful**

Divide by the total of the *group you are asking about*. "What percent of phone users sleep under 8 hours?" divides by the column total 65, giving  $\frac{45}{65} \approx 69.2\%$ , not  $\frac{45}{120}$ .

## Key takeaways

- A scatter plot shows pairs  $(x, y)$  as points; it is not connected.
- Association can be positive, negative, or none; linear or nonlinear.
- An outlier is far from the pattern; a cluster is a tight group of points.
- A line of best fit follows the trend; find its equation from two points on the line.
- Slope is the change in  $y$  per 1 unit of  $x$ ; the intercept is the value of  $y$  when  $x = 0$ .
- Predict inside the data range; be careful when extrapolating.
- In a two-way table, compare relative frequencies to look for association.
- Association does not prove causation.



## Test yourself: quick challenge for Grade 8

[Change level](#)

Level : **Grade 8**

6 random questions, instant feedback.



**Start the challenge**



 **Speed drill for Grade 8: how many in 60 seconds?**

## Keep exploring with Zyro

### Math practice

Scatter Plots and Two-Way Tables: math practice, Grade 8

### Math tests

Scatter Plots and Two-Way Tables: math test, Grade 8

### Math quizzes

Scatter Plots and Two-Way Tables: math quiz, Grade 8

### Math practice

Rational and Irrational Numbers: math practice, Grade 8

 **Math practice**

Exponents and Scientific Notation: math practice, Grade 8

 **Math tests**

Rational and Irrational Numbers: math test, Grade 8

**Zyro vous emmène sur nos sites partenaires**

Embarquez pour d'autres planètes du savoir et leurs applis.

Décollage