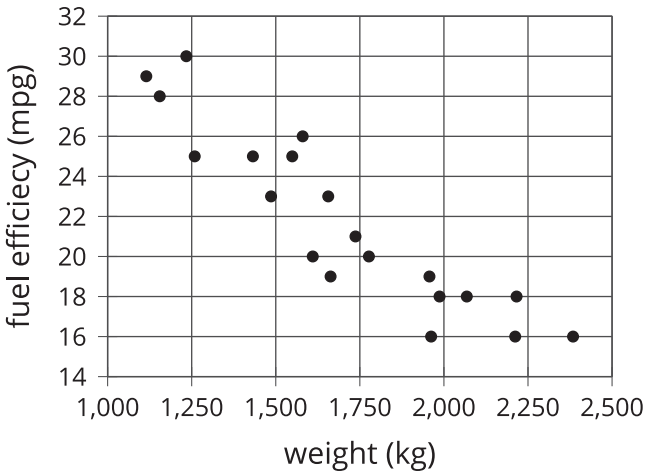
Unit 6, Lesson 4

Fitting a Line to Data

Let's look at the scatter plots as a whole.

4.1 Predict This

Here is a scatter plot that shows weights and fuel efficiencies of 20 different types of cars.



If a car weighs 1,750 kg, would you expect its fuel efficiency to be closer to 22 mpg or to 28 mpg? Explain your reasoning.

4.2 Battery Life

In an experiment, a group gathers different rechargeable devices like phones, watches, electric toothbrushes, headphones, and cordless vacuums. They use each device until the battery runs out and record the initial battery life in hours. They use the devices for a year and then do the same experiment 1 year later, and record the battery life in hours again.

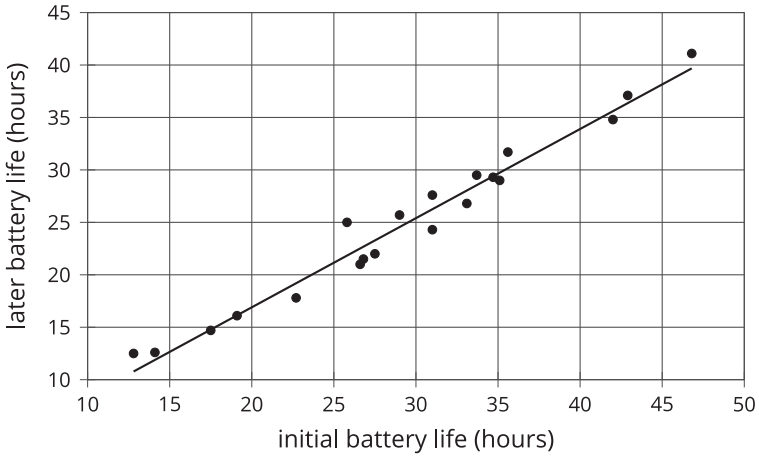
In this table, the first column shows battery life for 20 devices initially and the second column shows battery life 1 year later.

initial battery life (hours)	later battery life (hours)	predicted later battery life (hours)
12.8	12.5	10.78
14.1	12.6	11.885
17.5	14.7	14.775
19.1	16.1	16.135
22.7	17.8	19.195
25.8	25	21.83
26.6	21	22.51
26.8	21.5	22.68
27.5	22	23.275
29	25.7	24.55
31	27.6	26.25
31	24.3	26.25
33.1	26.8	28.035
33.7	29.5	28.545
34.7	29.3	29.395
35.1	29	29.735
35.6	31.7	30.16
42	34.8	35.6
42.9	37.1	36.365
46.8	41.1	39.68

The scatter plot shows the battery life measurements for 20 devices together with the graph of $y = 0.85x - 0.1$.

The function described by the equation $y = 0.85x - 0.1$ is a *model* of the relationship between the initial and later battery life for a device.

This model *predicts* the later battery life from the device's initial battery life. These predicted battery lives are shown in the third column of the table.



1. Two devices that both have an initial battery life of 31 hours have different later battery life. What are their later battery lives? How can you see this in the table? How can you see this in the graph?

2. The model predicts that when the initial battery life is 31 hours, the later battery life will be 26.25. How can you see this in the graph? How can you see this using the equation?

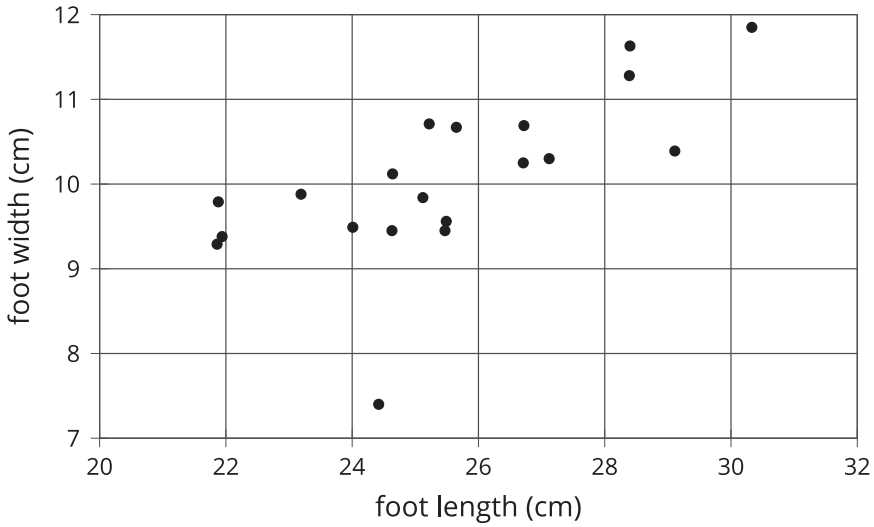
3. One of the devices has an initial battery life of 29 hours. What does the model predict for its later battery life? How does that compare to the actual later battery life?

4. Find a device for which the model makes a very good prediction of the actual later battery life. How can you see this in the table? In the graph?

5. Find a device for which the model's prediction is not very close to the actual later battery life. How can you see this in the table? In the graph?

4.3 The Agony of the Feet

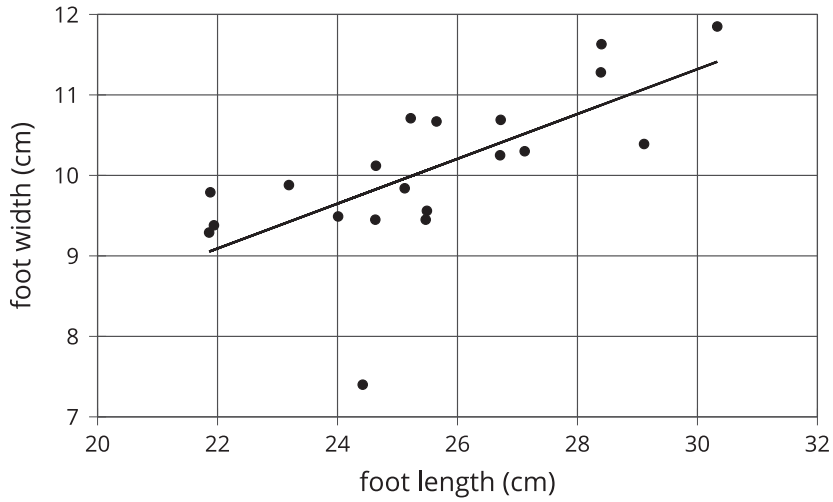
Here is a scatter plot that shows lengths and widths of 20 different left feet.



1. Estimate the widths of the longest foot and the shortest foot.

2. Estimate the lengths of the widest foot and the narrowest foot.

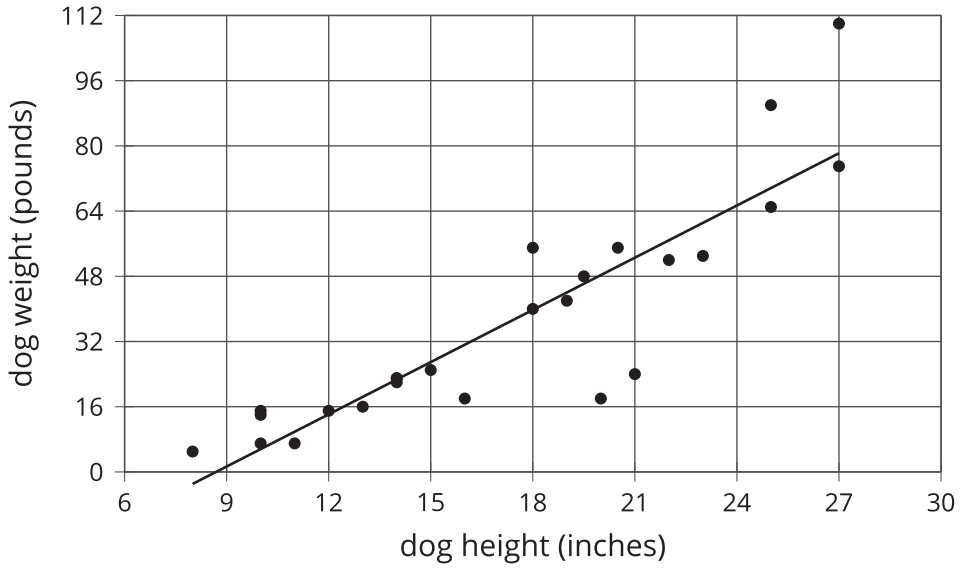
3. Here is the same scatter plot together with the graph of a model for the relationship between foot length and width.



Circle the data point that seems weird when compared to the model. What length and width does that point represent?

Lesson 4 Summary

Sometimes, we can use a linear function as a model of the relationship between two variables. For example, here is a scatter plot that shows heights and weights of 25 dogs together with the graph of a linear function which is a model for the relationship between a dog's height and its weight.



For some dogs, we can see that the model does a good job of predicting the weight given the height. These correspond to points on or near the line. The model doesn't do a very good job of predicting the weight given the height for the dogs whose points are far from the line.

For example, there is a dog that is about 20 inches tall and weighs a little more than 16 pounds. The model predicts that the weight would be about 48 pounds. We say that the model overpredicts the weight of this dog. There is also a dog that is 27 inches tall and weighs about 110 pounds. The model predicts that its weight will be a little less than 80 pounds. We say the model underpredicts the weight of this dog. For most of the dogs in this data set, though, the model does a good job of predicting the weight from the height.

Sometimes a data point is far away from the other points or doesn't fit a trend that all the other points fit. We call these **outliers**.

CC BY NC Illustrative Mathematics, based on IM 6–8 Math, CC BY Open Up Resources.