

Lesson 3.2 Line of Best Fit



Relationships Summary

A **scatter plot** is a graph that shows the relationship between **two** variables.

The points in a scatter plot often show a pattern, or trend.

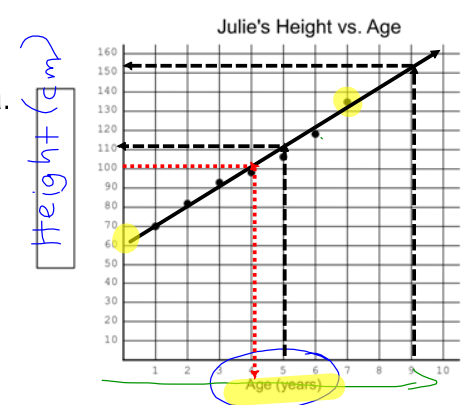
From the pattern or trend you can describe the correlation.

Example 1:

Julie gathered information about her age and height from the markings on the wall in her house.

Age (years)	1	2	3	4	5	6	7	8
Height (cm)	70	82	93	98	106	118	127	135

- a) Label the vertical axis.
- b) Describe the trend in the data.
Upward trend
(positive correlation)
- c) Describe the relationship.
As age increases
height increases



Variables

The *independent variable* is located on the x (horizontal) axis.

This variable does not depend on the other variable.

The *dependent variable* is located on the y (vertical) axis.

This variable depends on the other variable.

Note:
The independent variable comes *first* in the table of values.

Independent variable: Age

Dependent variable: Height

Line of Best Fit

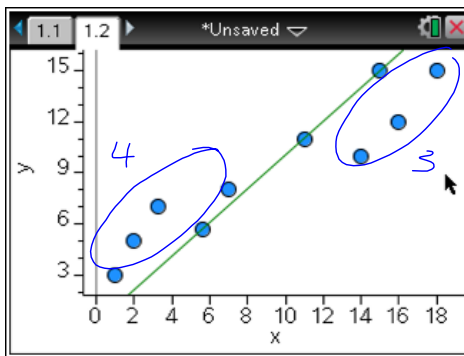
<https://www.explorelarning.com/index.cfm?method=cResource.dspView&ResourceID=308>

To be able to make predictions, we need to model the data with a line or a curve of best fit.

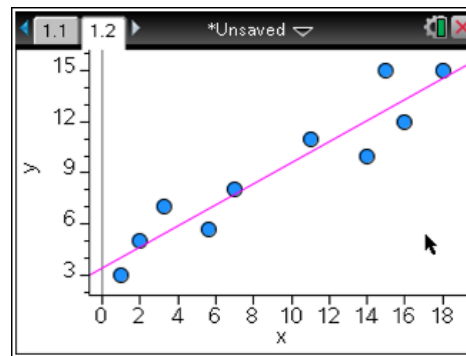
Rules for drawing a line of best fit:

1. The line must follow the Points / trend.
2. The line should pass through as many points as possible.
3. There should be the same number of points above and below the line.
4. The line should pass through points all along the line, not just at the ends. ** line going through the middle of points*

Bad LoBF



Good LoBF



Making Predictions

Use your line of best fit to estimate the following: *→ you HAVE TO USE A LINE OF BEST FIT*

Question	Answer	Method of Prediction
How tall was Julie when she was <u>5 years</u> old?	110	Interpolate
How tall will Julie be when she is 9 years old?	152/150/	extrapolate
How old was Julie at 100 cm tall?	4 yrs	interpolate
How tall was Julie when she was born?	60 cm	extra polate

Interpolate

When you interpolate, you are making a prediction inside the data.

These predictions are usually accurate (close).

Hint:

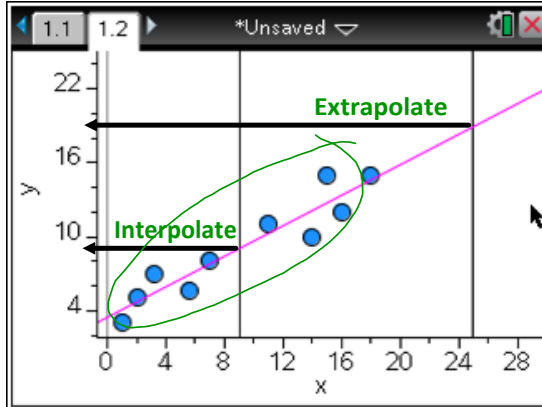
You are interpolating when the value you are finding is somewhere between the first point and the last point.

Extrapolate

When you extrapolate, you are making a prediction outside the data.

It often requires you to extend the line.

These predictions are less reliable.



You are extrapolating when the value you are finding is before the first point or after the last point. This means you may need to extend the line.

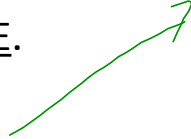
Continuous Data

- Can have an infinite number of possible values within a given range.

* Often MEASURED.

* Can have part of a whole (ex. can have half).

* Connect points with a SOLID LINE.



Discrete Data

- Can only have certain values within a given range.

* Often COUNTED.

* Can't have part of a whole (ex. Can't have half).

* Connect points with a DASHED LINE.



QUESTION

Should the scatter plot showing Julie's age and height be connected with a solid line or a dotted line?

Both variables
are continuous.

$$\frac{75-25}{16} = 3$$

How can you figure out a good scale?

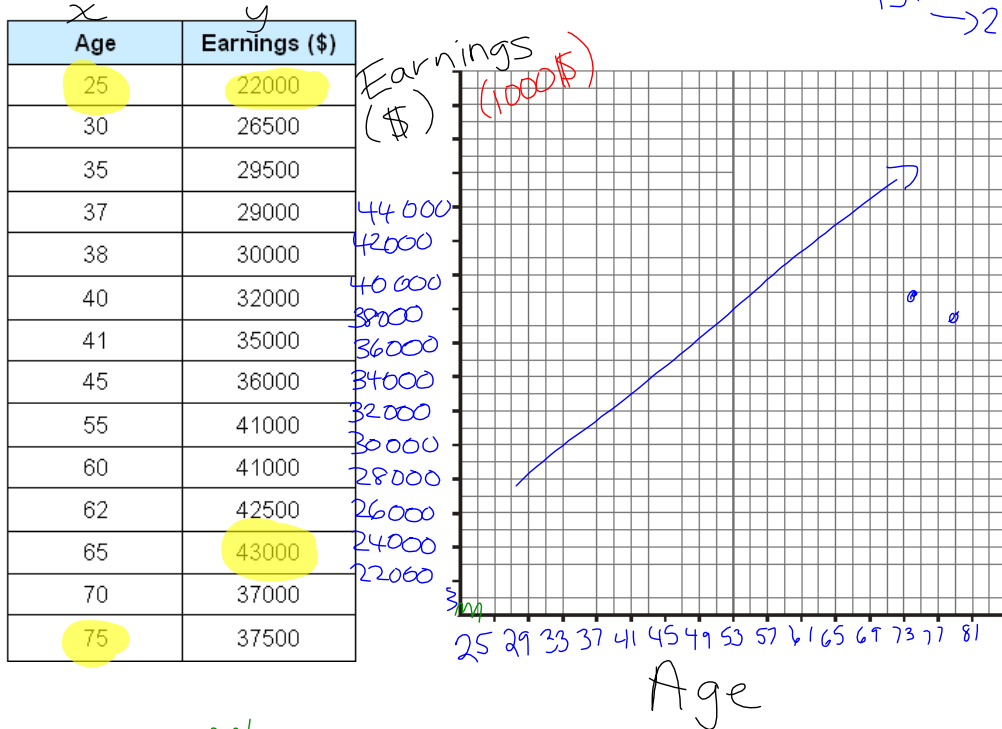
- * Determine the highest & lowest data points.
- > Subtract (This is the RANGE).
- * Count the number of grid marks to label.
- > Divide the RANGE by this number.

$$\begin{array}{r} 43000 - 22000 \\ = 21000 \\ \hline 16 \\ 1312 \rightarrow 1500 \\ \rightarrow 2000 \end{array}$$

Example 2: Creating Scatter Plots and Lines of Best Fit

Test the hypothesis: The older you are, the more money you earn.

Plot the data on the scatter plot below, choosing appropriate scales and labels.



Note: The symbol \mathbb{N} is used to signal a "break" in the axis when the scale does not start at zero to avoid a large empty space in one corner of the graph.

- 1) Draw a line of best fit. Describe the trend in the data. *upward*
- 2) Does the data support the hypothesis? Give reasons to support your answer. (Refer to the scatter plot.) Describe the relationship.
In general } yes
To ascertain }
- 3) Explain why the data for ages over 65 do not correspond with the hypothesis.
Retired
- 4) Explain what the point (41, 35000) represents.
- 5) Use the line of best fit to predict earnings when you are 50 years old. Did you interpolate or extrapolate?
- 6) Use the line of best fit to predict your age when you earn \$28000. Did you interpolate or extrapolate?

Homework
p. 153 # 1, 4

One Step Better



- What is being compared?
- Did I include the important features of a scatter plot?
- What can I conclude?