

Chapter 3 - 8

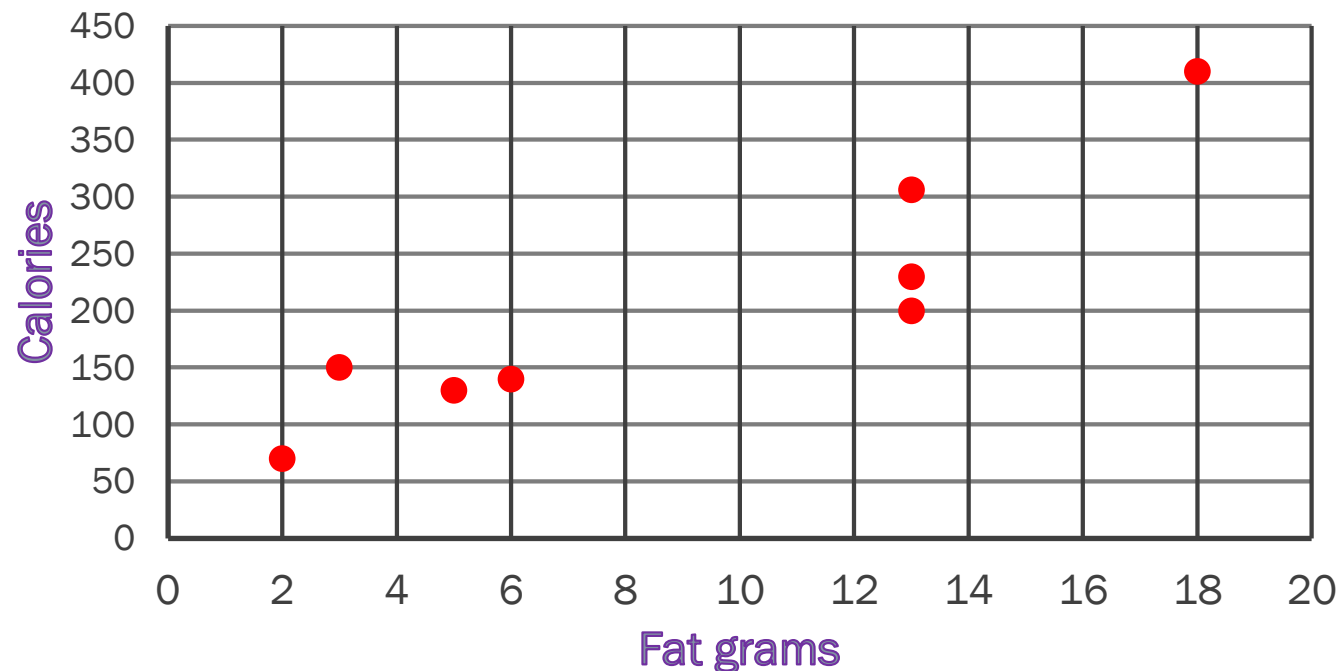
Scatter Plots and Lines of Best Fit



Vocabulary:

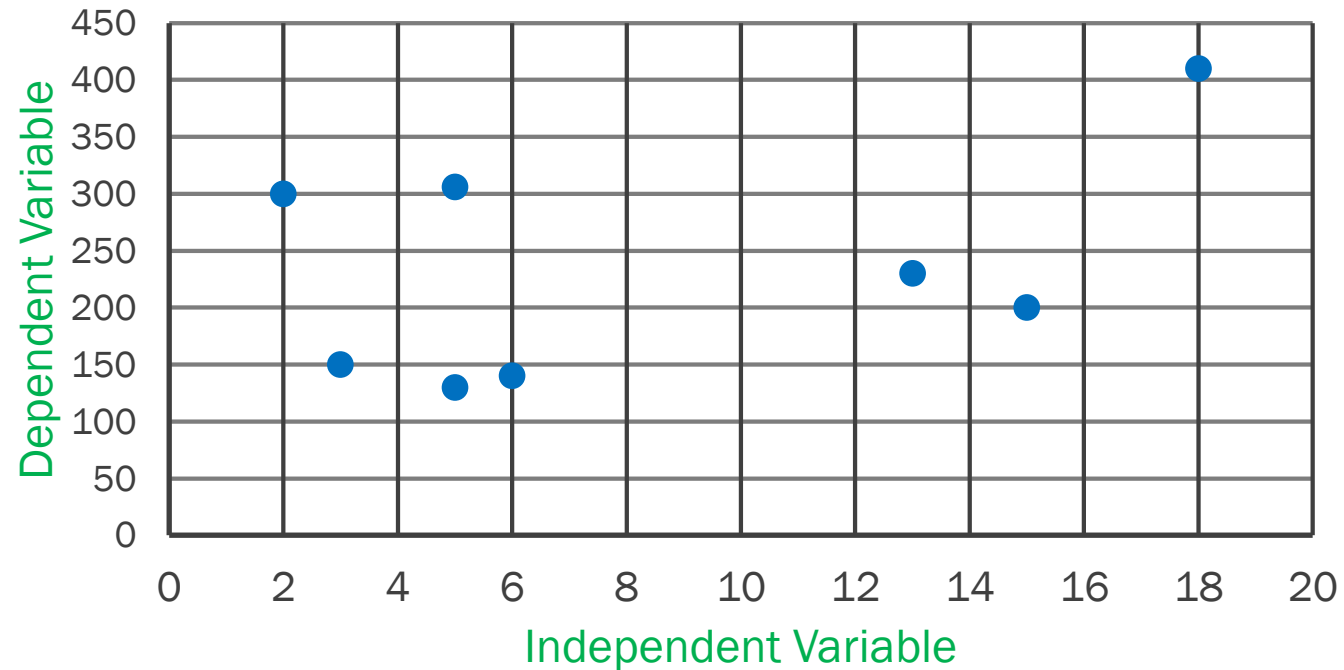
A scatter plot is a graph that shows the relationship between two sets of data. The two sets of data are graphed as ordered pairs on a coordinate plane. A scatter plot may show a pattern or relationship of the data.

This scatter plot shows the number of grams of fat and calories for a variety of potato chips. Each point has an x -value, representing fat grams, and a y -value, representing calories. This scatter plot has a positive relationship, because as x increases, y increases.



Vocabulary:

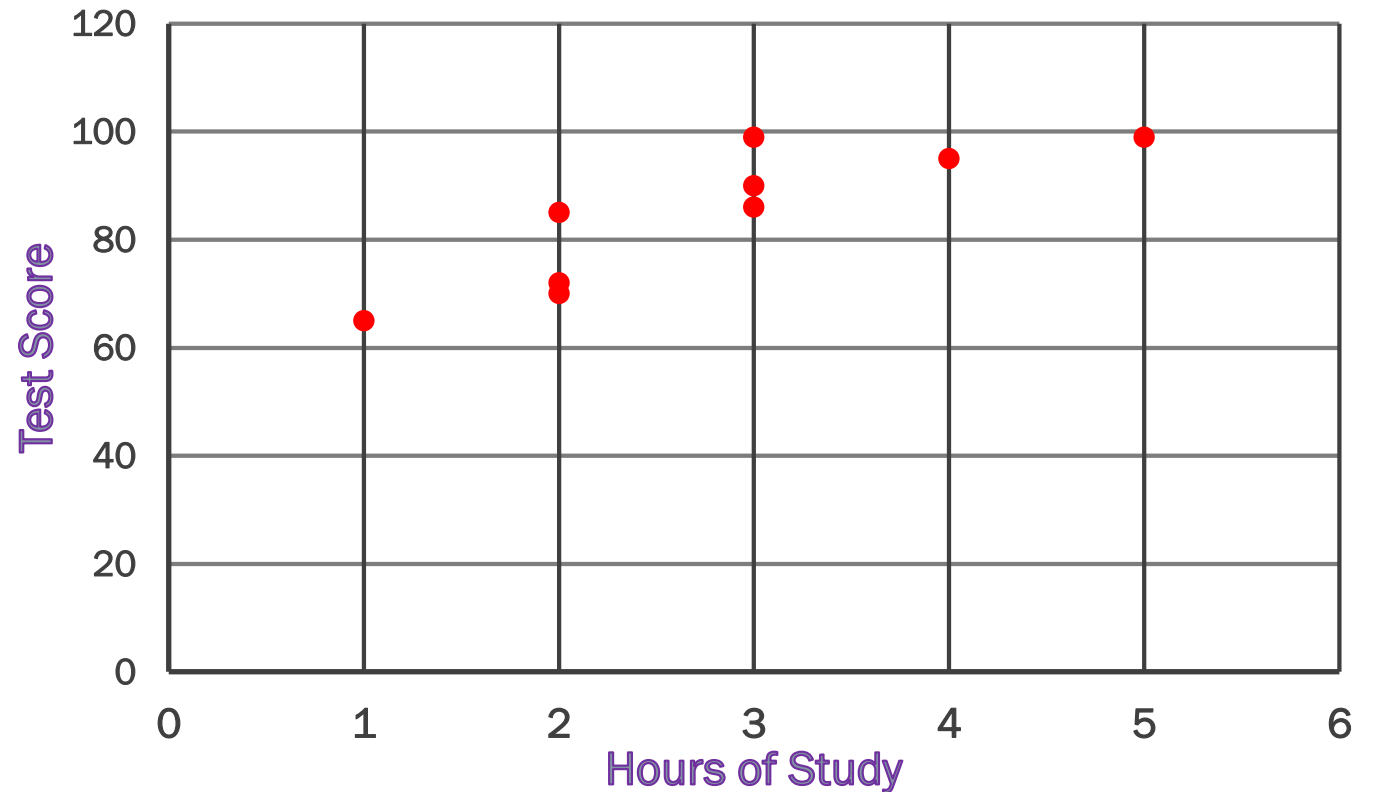
The independent variable is x and the dependent variable is y. The independent variable is the horizontal axis and the dependent variable is the vertical axis.



Vocabulary:

The scatter plot is a graph of the data in the table on the left. This again shows a positive relationship. Longer hours of study tend to mean higher test scores.

Hours of study	Test scores
1	65
2	70
2	72
2	85
3	90
3	86
3	99
4	95
5	99

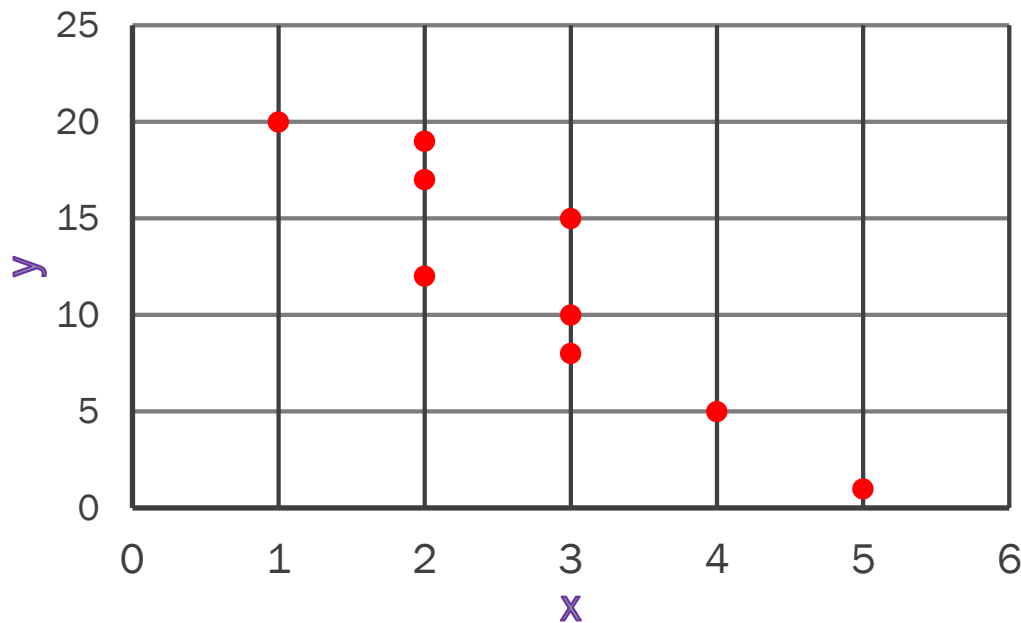


Vocabulary:

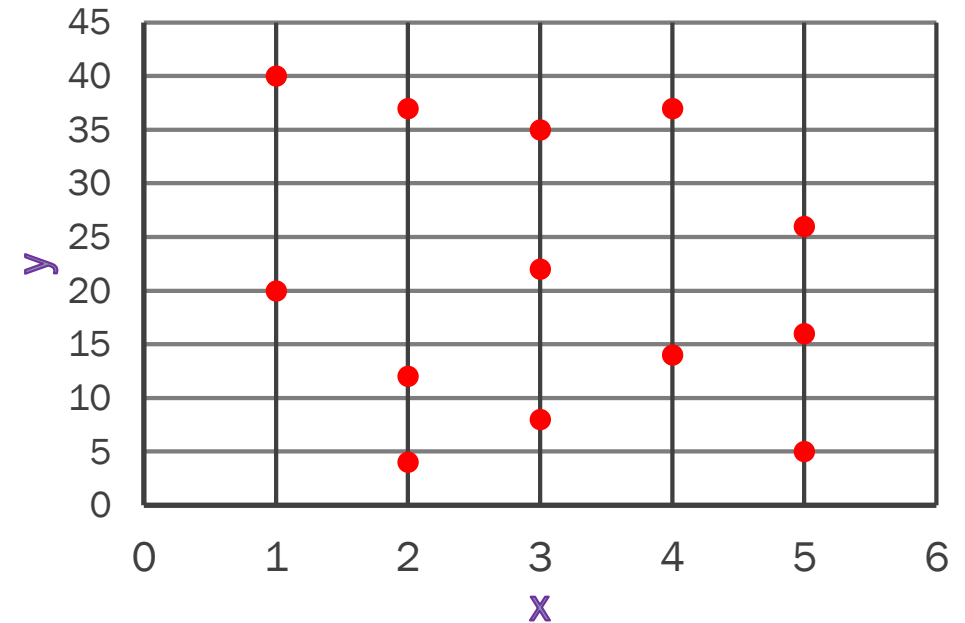
The scatter plot on the left shows a negative relationship between the x and y coordinates. As x increases, y tends to decrease.

The scatter plot on the right shows no apparent relationship.

Negative Relationship



No Relationship

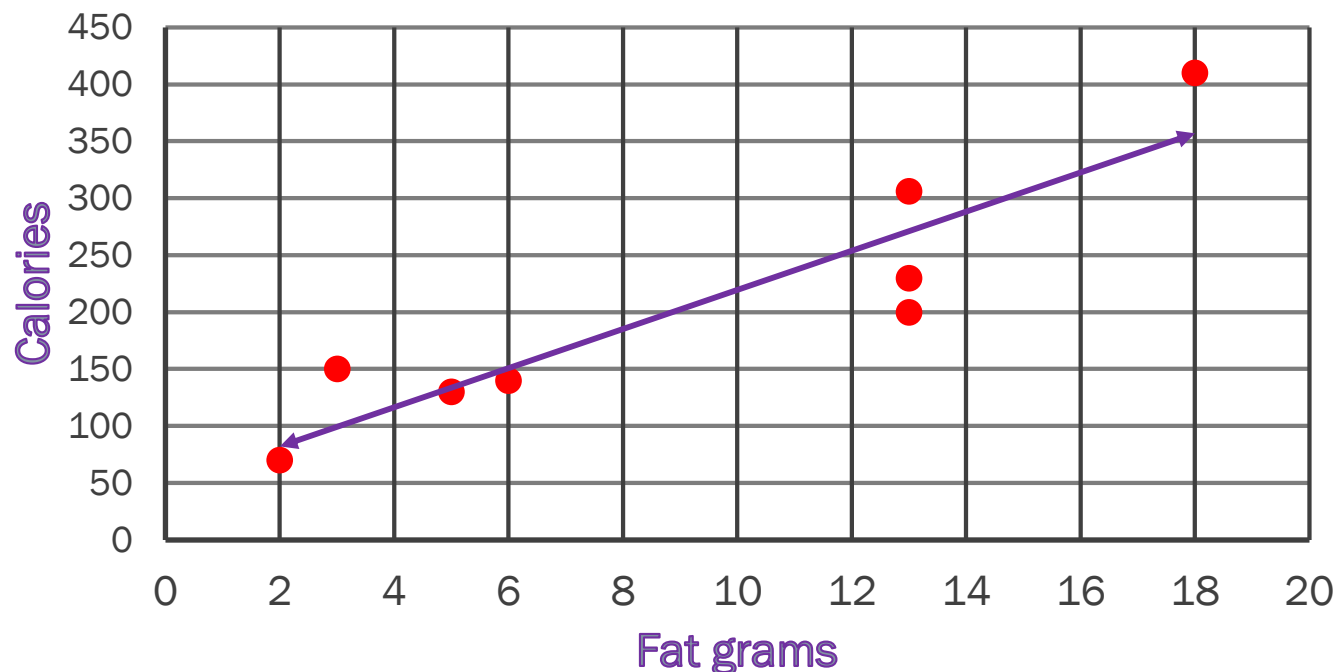


Vocabulary:

A line of best fit or trend line is a straight line that best represents the data on a scatter plot. This line may pass through some of the points, all of the points, or none of the points.

A line of best fit can help you make predictions based on the data. It shows what pattern the data follows.

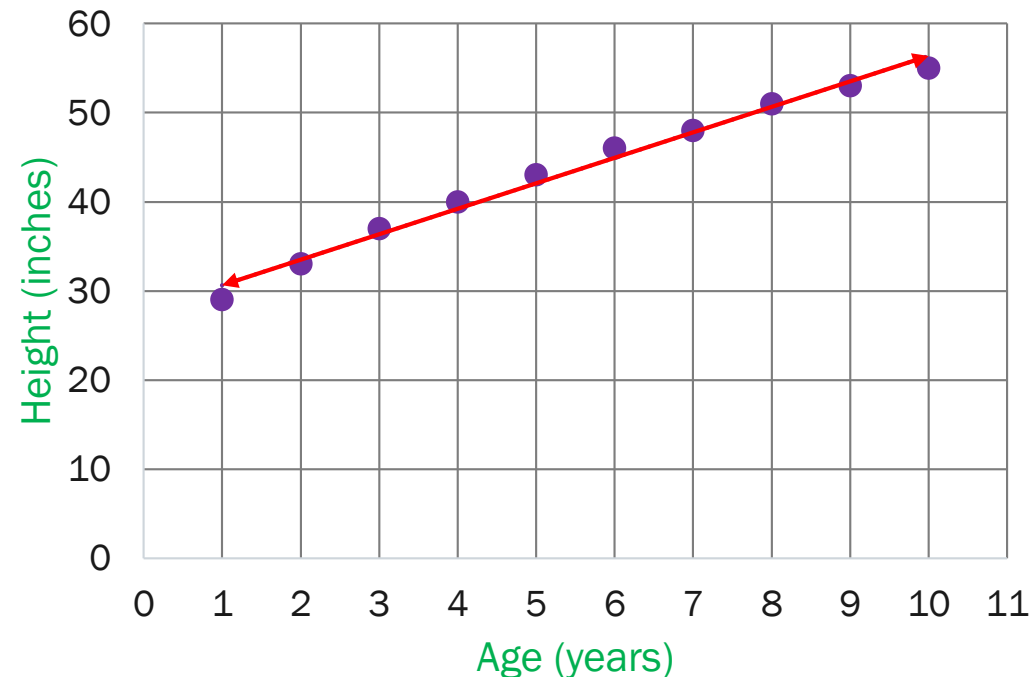
The scatter plot from earlier relating fat grams and calories is shown below. When we add a line of best fit, you can see that a couple points are very close to the line and about half of the remaining points are above the line and half are below the line.



Examples

The data set shows the height of a child from ages 1 to 10. Draw a scatter plot of the data and a line of best fit.

Age (yrs)	Height (in)
1	29
2	33
3	37
4	40
5	43
6	46
7	48
8	51
9	53
10	55



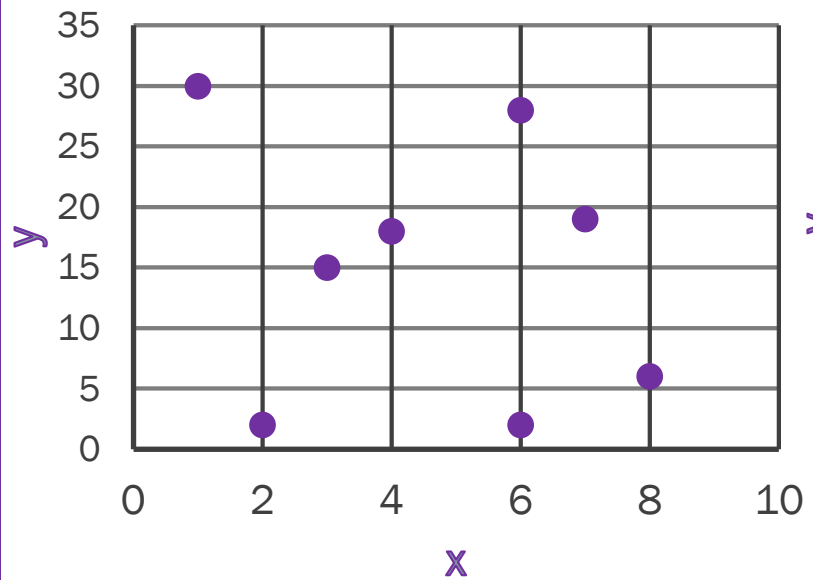
1. Label the axes
2. Choose your scale for each axis
3. Plot each point

Does this scatter plot have a positive relationship, negative relationship or no relationship?

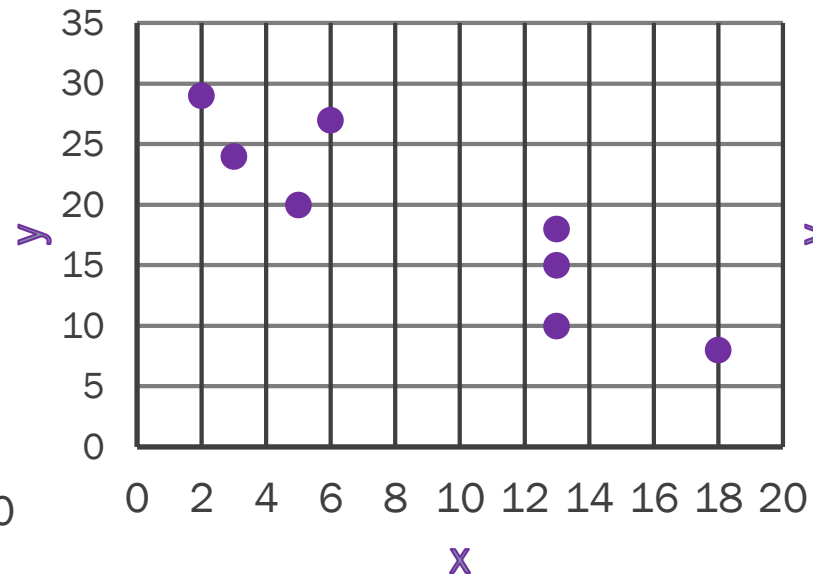
The scatter plot has a positive relationship because as x (age) increases, y (height) increases.

Examples

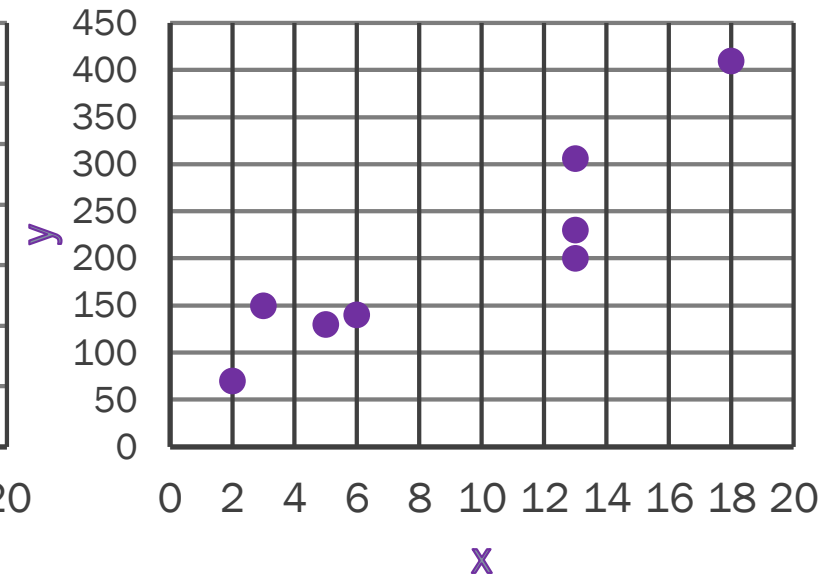
What type of relationship between x and y does each scatter plot show?



No relationship



Negative relationship



Positive relationship

Examples

Which is the independent variable (x) and which is the dependent variable (y) in each situation?

1. How many books are sold per month at a bookstore.

Independent variable: Number of months

Dependent variable: Number of books sold

1. How much money you earned from working at your job

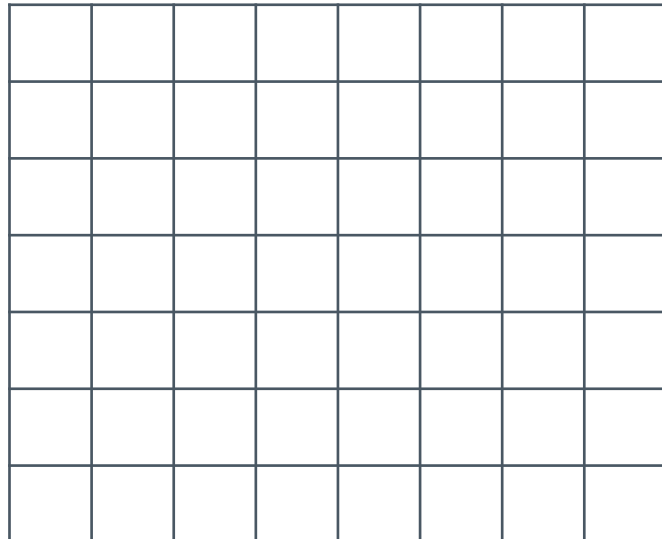
Independent variable: Time spent working

Dependent variable: Wages earned

Let's Try An Example

Using the data set shown relating year to US population.

1. What is the independent variable? What is the dependent variable?
2. Draw a scatter plot of the data.
3. Approximate the line of best fit.
4. Use the line to estimate the US population in 2020.



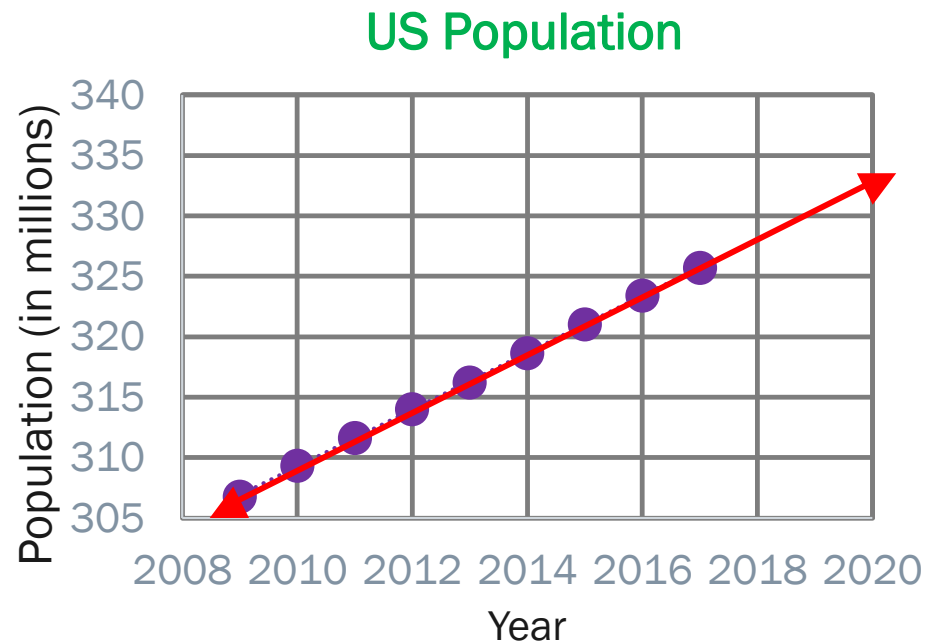
Year	Population (in millions)
2009	306.77
2010	309.34
2011	311.64
2012	313.99
2013	316.23
2014	318.62
2015	321.04
2016	323.41
2017	325.72

Let's Check our Answers

Using the data set shown relating year to US population.

1. What is the independent variable? What is the dependent variable?
2. Draw a scatter plot of the data.
3. Approximate the line of best fit.
4. Use the line to estimate the US population in 2020.

Year	Population (in millions)
2009	306.77
2010	309.34
2011	311.64
2012	313.99
2013	316.23
2014	318.62
2015	321.04
2016	323.41
2017	325.72



1. Year is the independent variable and population is the dependent variable.
2. See left
3. See left
4. If the trend continues, the US population will be about 332.5 million in 2020.