DISCRETE AND CONTINUOUS DATA

This unit is about the effectiveness of graphs as they represent data. Be careful to look at all parts of a graph and what it intends to convey. The reader of the graph needs to see the information quickly. In addition, data will be displayed in scatter plots and conclusions will be made based on the trend of the data displayed in a scatter plot.

Discrete and Continuous Data

To place data on graphs and categorize types of data that work best with different graphs, it is necessary to distinguish between different types of data to be graphed. The first type of data to consider is **discrete data**. This is a type of data in which separate items are counted. The number of pets, the number of people at a party, and the number of children in a class are all examples of discrete data. The sets of data are viewed as distinctly separate items. A bar graph or pictograph would be wise choices to demonstrate discrete data graphically.

Another type of data is **continuous data**. This kind of data has values that may fall in between the recorded data. Through a range of values, it can be estimated where the values will fall. For example, to show how temperature changes over a period of time, the temperatures could be recorded daily at the same time and a graph of those temperatures could be made. Of course, there would be temperatures occurring in between theses reading. This is continuous data; that is, additional data that occurs between the recorded data. In this example a line graph would be a wise choice to demonstrate the data as a graph.

Discrete Data

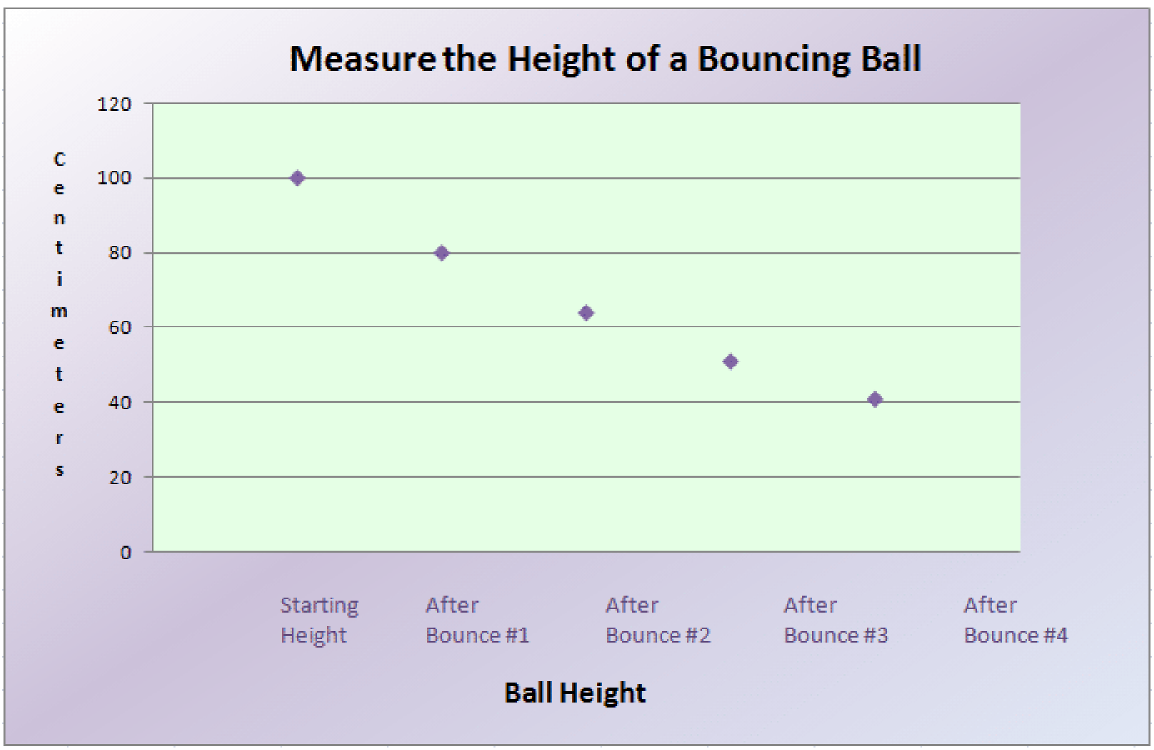
Read through the scenario* below to gain a better understanding of discrete and continuous data.

**Source: http://mathforum.org/t2t/message.taco?thread=10489&message=2*

Suppose a ball is dropped from a height of 100 cm and it bounces 4 times. The data collected is the height reached after each bounce.

Starting height:	100 cm
After bounce #1:	80 cm
After bounce #2:	64 cm
After bounce #3:	51 cm
After bounce #4:	41 cm

To graph this data, the *x*-axis would be the bounce number. The *y*-axis would be the height in centimeters. In this graph, the dots would NOT be connected.



This is an example of **DISCRETE** data.

*Note: If a line segment was drawn connecting the points (1, 80 cm) and (1, 64 cm), the line segment implies that there exists data representing bounces between the 1st and the 2nd bounce. Since a "fractional bounce number" does not exist, we should NOT connect the dots.

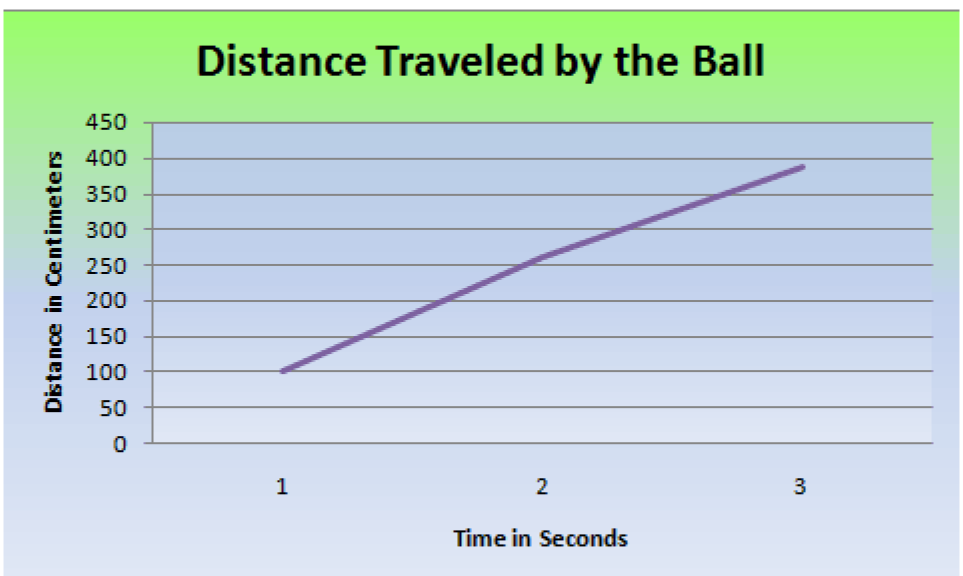
Continuous Data

Let's change the situation slightly. Now let's collect the data that represents the distance traveled by the ball over time. To make this example easier, let's assume it takes one second to travel from the initial drop to the first bounce, and one second to travel between each of the subsequent bounces.

Here's the data.

Time	Distance
1 sec.	100 cm
2 sec	260 cm (100 + 80 + 80)
3 sec	388 cm (100 + 80 + 80 + 64 + 64)
etc	etc.

The *x*-axis would represent the time in seconds and the *y*-axis would represent the distance traveled in centimeters. In this graph the dots WILL be connected. The segment connecting the points (1,100) and t(2,260) implies that there exists data between one second and two seconds. That is, there is a distance that can be measured when the ball travels for any fractional number between 1 and 2 seconds.



This is an example of **CONTINUOUS** data.

Discrete or Continuous Data?

Bicyclists: "Does there exist 1 1/2 bicyclists?" Discrete data

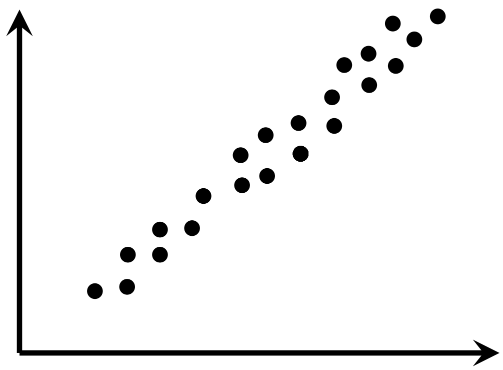
Miles: "Can we travel 2 2/3 miles?" Continuous Data

Cars: "Is there such a thing as 4.2 cars?" Discrete data

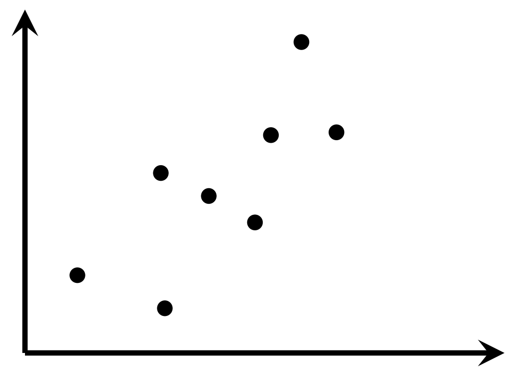
Scatter Plots

A **scatter plot** shows relationships between two sets of data in clusters of data with dots or symbols.

Positive Correlation of Data*:* If the points appear to suggest a line that slants upward and to the right, there is a positive relationship.



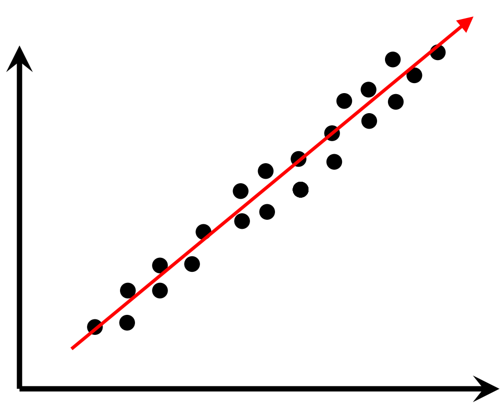
Strong Positive



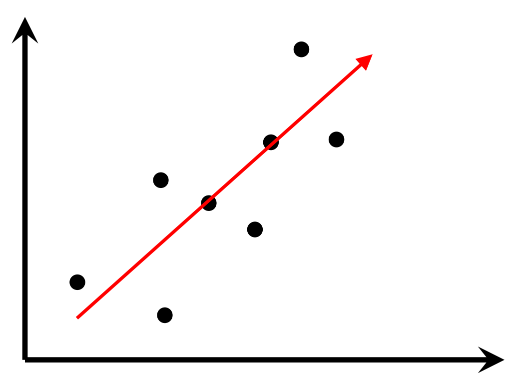
Weak Positive

Positive correlation: Both data sets increase together.

Line of Best Fit: The straight line that is located as close as possible to most of the data points is the best-fitting line. Notice the positive slope of each straight line added to the graphs below.



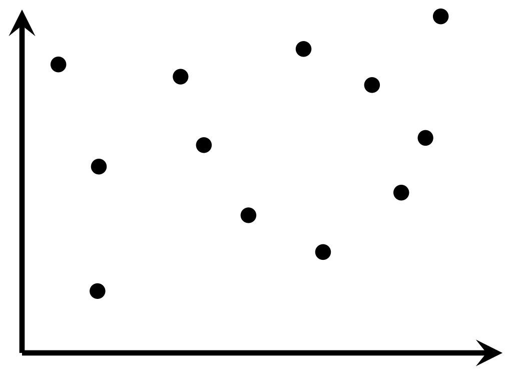
Strong Positive



Weak Positive

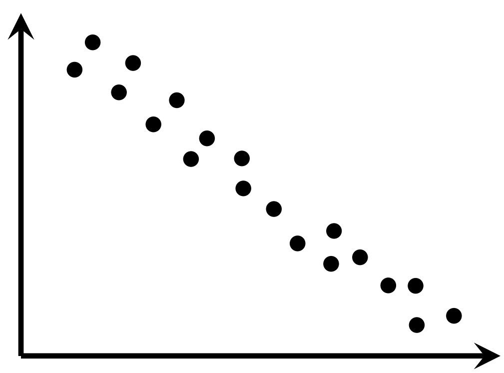
*NOTE: The best-fitting line does not need to pass through or even near all the points.

No Correlation of Data*:* If the points seem to be randomly located, then there is no relationship.

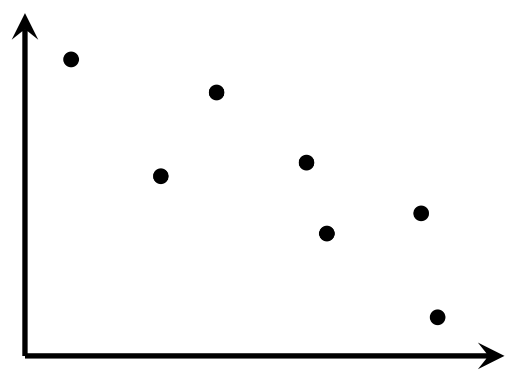


No correlation: The changes in one data set do not affect the other data set.

Negative Correlation of Data*:* If the points appear to suggest a line that slants downward and to the right, there is a negative relationship.



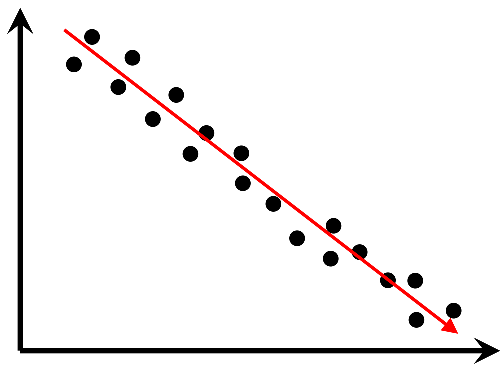
Strong Negative



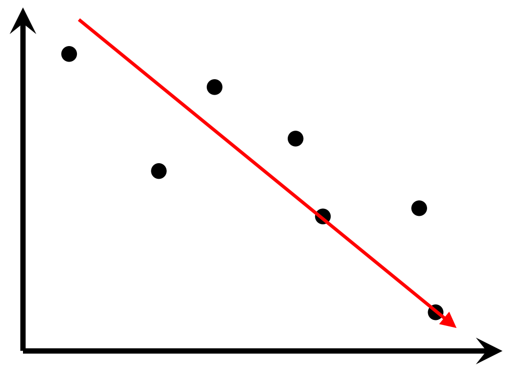
Weak Negative

Negative correlation: As one data set increases, the other decreases.

Line of Best Fit: The straight line that is located as close as possible to most of the data points is the best-fitting line. Notice the negative slope of each straight line.



Strong Negative



Weak Negative

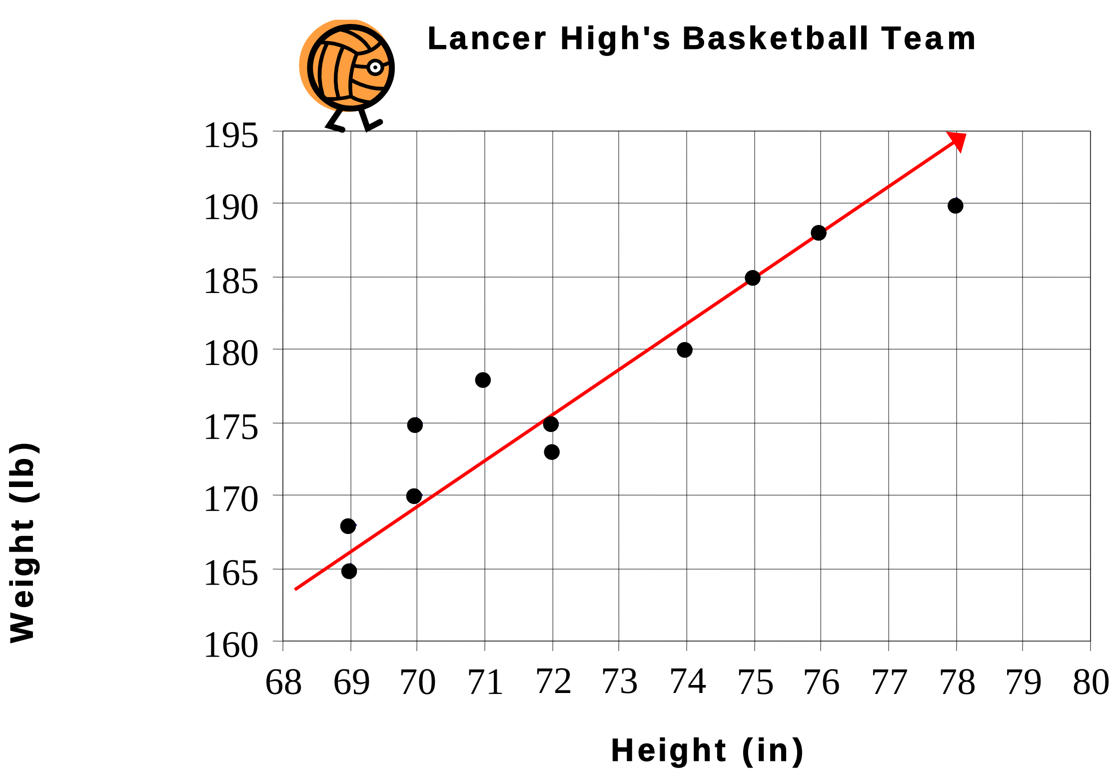
Example: Use the given data to make a scatter plot of the weight and height of each member of Lancer High's basketball team.

Step 1: Make a scatter plot of the data pairs. The points on the scatter plot are (70, 170), (69, 165), (72, 175), (74, 185), (75, 185), (70, 175), (69, 168), (72, 173), (71, 178), (78, 190), (76, 188), and (69, 165).

Step 2: Draw the line that appears to best fit the data points. There should be about the same number of points above the line as below it. The line does not have to pass through any of the data points.

Height (in)	**Weight (lb)**
70	170
69	165
72	175
74	180
75	185
70	175
69	168
72	173
71	178
78	190
76	188
69	165

What kind of correlation exists between the data sets? The line of best fit slants upward and to the right.



There is a positive correlation between the height of the team member and his weight.