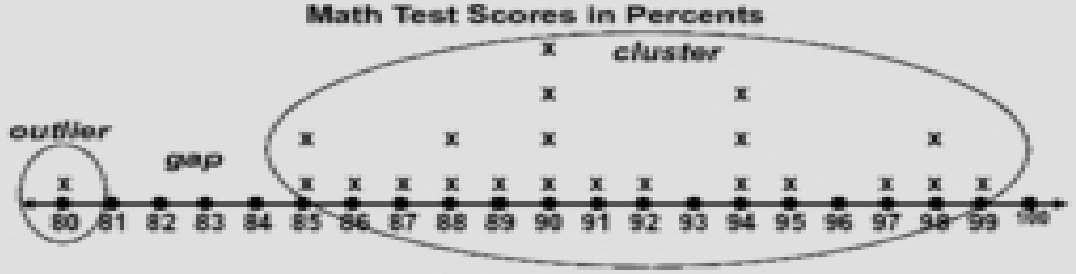
Name: _____ Date: _____

Data Distribution: The Shape of Data

Distribution is the arrangement of the values in a data set. To better understand a data set, the values may be placed (or distributed) on a graph. Once graphed, it is easy to see the shape of the data set - how values are grouped together or spread apart.

Data values that are grouped closely together are called **clusters**. Sometimes values are spread apart, leaving areas on the graph where there are no data values. Those areas with no values are called **gaps**. Data values that are much smaller or larger than the clustered values are called **outliers**.

In the dot plot below, note the cluster, gap and outlier.



Observing data on a graph gives a quick overall view of the distribution of data. The above dot plot is a graph of math test scores in percents. Here is an unordered list of the same test scores:
88, 91, 90, 85, 94, 98, 99, 86, 90, 85, 92, 88, 87, 89, 94, 98, 90, 80, 94, 90, 95, 97

By looking at the graph, you can find the answers to the following questions much more easily than by looking at the list.

What is the highest test score? What is the most common test score?
Which score is significantly lower than the others?
What numbers make up the gap? How many tests scored 94%?
What is the spread of clustered scores?