The topic of variation is the single most important topics in statistics. An average is an attempt to summarize a set of data. However, an average taken by itself may not always be very meaningful. We need a statistical cross-reference that measures the spread of the data or variation of data.

Example 1

Suppose there are two factories producing the batteries. From each factory, 4 batteries are drawn to test for the lifetime (in hours).

Factory A: 10.1 9.9 10.1 9.9

Factory B: 16 5 7 12

The mean lifetimes of the two factories are both 10. However, by looking at the data, it is obvious that the batteries produced by factory A are much more reliable than Factory B.

*Calculate the Range for Factory A and Factory B*

Although the range tells the difference between the largest and smallest values in a distribution, it doesn’t tell us how much other values vary from one another or from the mean.

**II. Variance ( ) and Standard Deviation ( )**

The variance and Standard Deviation measures the variation of the data or spread of the data around the mean.
III. How to find the Variance and Standard Deviation using a graphing calculator.

Step 1: Enter data: press STAT → EDIT → ENTER
Step 2: Calculate the mean: press STAT → CALC → 1-VAR STATS → ENTER → $s_x$

Example 2:

Calculate the Variance and Standard Deviation for Factory A using a calculator and Factory B without a calculator

Factory A: 10.1 9.9 10.1 9.9
Factory B: 16 5 7 12

V. Range RULE OF THUMB

(1) The vast majority of sample values lie within 2 standard deviation from the mean.

And if we know the mean and the standard deviation of the sample, we can estimate the minimum and maximum usual values by adding or subtracting 2 times the standard deviation to or from the mean.

Minimum “usual” value:  
Maximum “usual” value: 
Example 5
If the mean height of all professional basketball players is 81 inches and the standard deviation is 3 inches, between what two heights would the majority of the players be?

Example 6
Would a height of 88 inches be unusual?

Example 7
Consider the data for the volumes of soda shown at the right. Within how many standard deviations of the mean do all of the data values fall?

<table>
<thead>
<tr>
<th>Volumes of soda (oz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>75       65       80</td>
</tr>
<tr>
<td>70       70       75</td>
</tr>
<tr>
<td>70       65       75</td>
</tr>
<tr>
<td>80       70       70</td>
</tr>
<tr>
<td>50       85</td>
</tr>
</tbody>
</table>

Are there any Outliers?