Consider a second sample from the same population. We record SBP on each subject in the second sample:

120 121 122 124 125 126 127

\( n = 7 \)

\( \bar{X} = \frac{865}{7} = 123.6 \)

What is different between the 2 samples?

<table>
<thead>
<tr>
<th>Continuous Variable</th>
</tr>
</thead>
</table>

### Dispersion

<table>
<thead>
<tr>
<th>X</th>
<th>(X - 123.6)</th>
<th>(X - 123.6)^2</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>-23.6</td>
<td>556.96</td>
</tr>
<tr>
<td>110</td>
<td>-13.6</td>
<td>184.96</td>
</tr>
<tr>
<td>114</td>
<td>-9.6</td>
<td>92.16</td>
</tr>
<tr>
<td>121</td>
<td>-2.6</td>
<td>6.76</td>
</tr>
<tr>
<td>130</td>
<td>6.4</td>
<td>40.96</td>
</tr>
<tr>
<td>130</td>
<td>6.4</td>
<td>40.96</td>
</tr>
<tr>
<td>160</td>
<td>36.4</td>
<td>1324.96</td>
</tr>
<tr>
<td>865</td>
<td>0</td>
<td>2247.72</td>
</tr>
</tbody>
</table>

\( s^2 = \frac{\sum (X - \bar{X})^2}{n-1} \)

\( s^2 = \frac{2247.72}{6} = 374.6 \)

### Sample Variance:

\[ s^2 = \frac{\sum (X - \bar{X})^2}{n-1} \]
Continuous Variable

- Sample Standard Deviation

\[ s = \sqrt{s^2} \]

\[ s = \sqrt{374.6} = 19.4 \]

Standard Summary: \( n=7, \bar{X} = 123.6, s=19.4 \)
### Median

Median holds 50% of values above and 50% of values below.

**Order data**

- For odd - median is middle value
- For even - median is mean of 2 middle values

100 110 114 121 130 130 160

Median

### Quartiles

- \( Q_1 \) = first quartile holds approximately 25% of the scores at or below it and
- \( Q_3 \) = third quartile holds approx. 25% of the scores at or above it
- \( Q_2 \) = ???

### Continuous Variable

<table>
<thead>
<tr>
<th>Ordered data</th>
</tr>
</thead>
<tbody>
<tr>
<td>100 110 114 121 130 130 160 180</td>
</tr>
</tbody>
</table>

- \( Q_1 \) = 125.5
- Median
- \( Q_3 \)
Box and Whisker Plot

Comparing Samples with Box and Whisker Plots

Outliers

IQR = Interquartile Range = Q3 - Q1
= range of middle half of the data
Outliers are values which either:
  exceed Q3 + 1.5 IQR, or
  fall below Q1 - 1.5 IQR
Sample: 2 3 4 5 6 100  Mean = 20, Median = 4.5, IQR = 6 - 3 = 3
100 is an outlier: It exceeds 6 + 1.5 (3) = 10.5
Summarizing Location and Variability

- When there are no outliers, the sample mean and standard deviation summarize location and variability.
- When there are outliers, the median and interquartile range (IQR) summarize location and variability, where \( \text{IQR} = Q_3 - Q_1 \).

Example

Sample: \( n=51 \) Participants in a study of cardiovascular risk factors

<table>
<thead>
<tr>
<th>Variable</th>
<th>Age (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>60 62 63 64 64 65 65 65 65 65 66 66 66 66 66 66 66 66 67 67 67 68 68 68 68 70 70 70 71 71 72 72 73 73 73 73 77 77 77 77 77 77 77 77 77 77 77 77 77 77 77 82 83 85 85 87</td>
</tr>
</tbody>
</table>

- \( Q_1 = 66, Q_3 = 76, \text{IQR} = 10 \)
  - Lower = 66 - 1.5(10) = 51
  - Upper = 76 + 1.5(10) = 91

Check for Outliers in Example
Example

Sample: n=51 Participants in a study of cardiovascular risk factors
Variable: age (years)

60 62 63 64 64 65 65 65 65 65 66 66 66 66 66 66 67 67 67 67 68 68 68 68 69 69 69 69 70 70 70 70 71 71 71 72 72 73 73 73 73 73 74 74 75 75 75 76 76 77 77 77 77 77 79 82 83 85 85 87

Sample mean: \( \bar{X} = \frac{\sum X}{n} = \frac{3637}{51} = 71.3 \)

Sample variance: \( s^2 = \frac{\sum (X - \bar{X})^2}{n-1} = 41.4 \)

Sample standard deviation: \( s = \sqrt{41.4} = 6.4 \)

Standard Summary: n=51, \( \bar{X} = 71.3, s=6.4 \)