Measurements of Disperion

- **Standard Deviation (s)**

Standard Deviation is calculated by the square root of the variance

$$s = \sqrt{s^2} = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}$$

Another disadvantage of using the sample variation and the standard deviation is that variability of the variable can't be compared in different units. Which variable has bigger variability - height or weight of an adult? To answer that, Coefficient of Variation has to be used.

- **Coefficient of Variation V_x**

· Coefficient of Variation V_x it represents relative measure of variability of the variable x and it is often expressed as a percentage

- It is the ratio of the sample standard deviation to the sample mean:

Example [3]

A table glass manufacturer has developed less expensive technology for improving the fireresistant glass. 10 glass table sheets were selected for testing. Half of them were treated by the new technology while the other half was used for comparison.

Both lots were tested by fire until they cracked. These are the results:

Critical temperature (glass cracked) [^{o}C]	
Old technology x_i	**New technology y_i**
475	485
436	390
495	520
483	460
426	488

Compare both technologies by means of basic characteristics of the exploratory analysis (mean, variation, etc.).

Solution:

- First you compare both technologies by the mean:

Mean for the old technology:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{475+436+\dots+426}{5} = 463.0 \quad [^{\circ}C]$$

Mean for the new technology:

$$\bar{y} = \frac{\sum_{i=1}^{n} y_i}{n} = \frac{485+390+\dots+488}{5} = 469.6 \quad [^{\circ}C]$$

Based on the calculated means the new technology could be recommended because the temperature it can withstand is 6^{o} C higher.

- Now you determine the measures of variability

The old technology:

Sample Variance:

$$s_x^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1} = \frac{(475-463.0)^2 + (436-463.0)^2 + \dots + (426-463.0)^2}{5-1} = 9163 \quad [^{\circ}C^2]$$

Standard Deviation:

$$s_x = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}} = \sqrt{s_x^2} = \sqrt{916.3} = 30.3 \quad [^\circ C]$$

New technology:

Sample variance:

$$s_y^2 = \frac{\sum_{i=1}^{5}(y_i - \bar{y})^2}{n-1} = \frac{(485-468.6)^2 + (390-468.6)^2 + \ldots + (488-468.6)^2}{5-1} = 2384.4 \quad [^\circ C^2]$$

Standard deviation:

$$s_y = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \bar{y})^2}{n-1}} = \sqrt{s_y^2} = \sqrt{2384.4} = 48.8 \quad [^\circ C]$$

Sample variance (standard deviation) for the new technology is significantly larger. What is the possible reason? Look at the graphical representation of the collected data. Critical temperatures are much more spread out which means this technology is not fully under control and its use can't guarantee higher production quality. In this case the critical temperature can either be much higher or much lower. For that reason it is recommended that it should be subjected to additional research. These conclusions based only on exploratory analysis. Statistics provides us with more exact methods for analyzing similar problems (hypothesis testing).