







## Standard Deviation

• The standard deviation s is the square root of the variance.

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

## Important Facts About s<sup>2</sup> and s

- Both variance and standard deviation will **ALWAYS** be non-negative values.
  - s=0 means that all observations have the same value.
- The units of \_\_\_\_\_ are the square of the units of the data.
- The units of \_\_\_\_\_\_ are the units of the data.
- Both are \_\_\_\_\_affected by outliers.

## Proposition

- Let *x*<sub>1</sub>,..., *x*<sub>n</sub> be a sample and *c* be a non-zero constant.
- If  $y_1 = x_1 + c$ , ...,  $y_n = x_n + c$ , then  $s_y = s_x$
- If  $y_1 = cx_1, ..., y_n = cx_n$ , then  $s_y = |c|s_x$



#### **Calculating Quartiles**

- 1. Arrange data in ascending order.
- 2. Locate the median.
- 3.  $Q_1$  is the median of all data to the left of the overall median.
- 4.  $Q_3$  is the median of all data to the right of the overall median.

### The Five-Number Summary

# Minimum, $Q_1, \tilde{x}, Q_3$ , Maximum

• A quick summary of both center **and** spread.









#### Fourth Spread

• The **Fourth Spread** (or Inter-quartile range) is a measure of variability that is resistant to *outliers*.

$$f_s = Q_3 - Q_1$$

- Outlier:
  - An observation more than  $1.5f_s$  from the nearest quartile
- Extreme Outlier:
  - An observation more than  $3.5f_s$  from the nearest quartile

