# Deviation from average

# Deviation from average

**Roughly how far are the numbers from their average?**

# Deviation from average

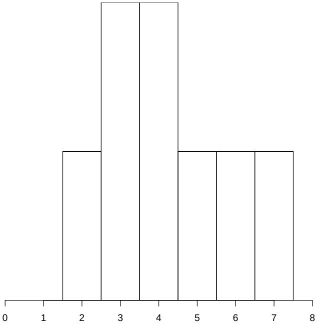
**Roughly how far are the numbers from their average?**

List: 2, 3, 3, 4, 4, 5, 6, 7 **average = 4.25**

# Deviation from average

**Roughly how far are the numbers from their average?**

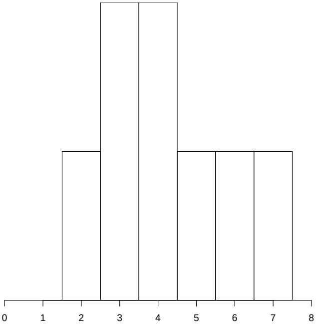
List: 2, 3, 3, 4, 4, 5, 6, 7 **average = 4.25**



# Deviation from average

**Roughly how far are the numbers from their average?**

List: 2, 3, 3, 4, 4, 5, 6, 7 **average = 4.25**




| value | deviation |
|---|---|
| 2 | $-2.25$ |
| 3 | $-1.25$ |
| 3 | $-1.25$ |
| 4 | $-0.25$ |
| 4 | $-0.25$ |
| 5 | 0.75 |
| 6 | 1.75 |
| 7 | 2.75 |

# About how big are the deviations?

**deviation from average = value − average**

# About how big are the deviations?

**deviation from average = value − average**

| | value | deviation |
|---|---|---|
| | 2 | −2.25 |
| | 3 | −1.25 |
| | 3 | −1.25 |
| | 4 | −0.25 |
| | 4 | −0.25 |
| | 5 | 0.75 |
| | 6 | 1.75 |
| | 7 | 2.75 |
| **average** | **4.25** | **0** |

# About how big are the deviations?

**deviation from average = value − average**

| | value | deviation | squared deviation |
|---|---|---|---|
| | 2 | −2.25 | 5.06 |
| | 3 | −1.25 | 1.56 |
| | 3 | −1.25 | 1.56 |
| | 4 | −0.25 | 0.06 |
| | 4 | −0.25 | 0.06 |
| | 5 | 0.75 | 0.56 |
| | 6 | 1.75 | 3.06 |
| | 7 | 2.75 | 7.56 |
| **average** | **4.25** | **0** | **2.44** |

# About how big are the deviations?

**deviation from average = value − average**

| | value | deviation | squared deviation |
|---|---|---|---|
| | 2 | −2.25 | 5.06 |
| | 3 | −1.25 | 1.56 |
| | 3 | −1.25 | 1.56 |
| | 4 | −0.25 | 0.06 |
| | 4 | −0.25 | 0.06 |
| | 5 | 0.75 | 0.56 |
| | 6 | 1.75 | 3.06 |
| | 7 | 2.75 | 7.56 |
| **average** | **4.25** | **0** | **2.44** |

**standard deviation** $= \sqrt{2.44} = 1.56$

# About how big are the deviations?

**deviation from average = value − average**

| | value | deviation | squared deviation |
|---|---|---|---|
| | 2 | −2.25 | 5.06 |
| | 3 | −1.25 | 1.56 |
| | 3 | −1.25 | 1.56 |
| | 4 | −0.25 | 0.06 |
| | 4 | −0.25 | 0.06 |
| | 5 | 0.75 | 0.56 |
| | 6 | 1.75 | 3.06 |
| | 7 | 2.75 | 7.56 |
| **average** | **4.25** | **0** | **2.44** |

**standard deviation** $= \sqrt{2.44} = 1.56$

**= root mean square of deviations from average**

# Terminology

Standard deviation: **SD**

# Terminology

Standard deviation: **SD**

List: \$2, \$3, \$3, \$4, \$4, \$5, \$6, \$7 **average** = \$4.25

# Terminology

Standard deviation: **SD**

List: \$2, \$3, \$3, \$4, \$4, \$5, \$6, \$7 **average** = \$4.25

**variance** = mean square of deviations from average = 2.44 squared dollars

# Terminology

Standard deviation: **SD**

List: \$2, \$3, \$3, \$4, \$4, \$5, \$6, \$7 **average** = \$4.25

**variance** = mean square of deviations from average = 2.44 squared dollars

$$\textbf{SD} = \sqrt{\text{variance}}$$
$$= \text{root mean square of deviations from average} = \$1.56$$

# Terminology

Standard deviation: **SD**

List: \$2, \$3, \$3, \$4, \$4, \$5, \$6, \$7      **average** = \$4.25

**variance** = mean square of deviations from average = 2.44 squared dollars

**SD** = $\sqrt{\text{variance}}$

= root mean square of deviations from average = \$1.56

The average and the SD have the same units.

# What does the SD measure?

**The standard deviation (SD) measures roughly how far off the entries are from their average.**

# What does the SD measure?

**The standard deviation (SD) measures roughly how far off the entries are from their average.**

List: 2, 3, 3, 4, 4, 5, 6, 7

# What does the SD measure?

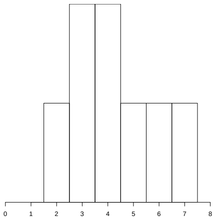
**The standard deviation (SD) measures roughly how far off the entries are from their average.**

List: 2, 3, 3, 4, 4, 5, 6, 7
average $= 4.25$
SD $= 1.56$

# What does the SD measure?

**The standard deviation (SD) measures roughly how far off the entries are from their average.**

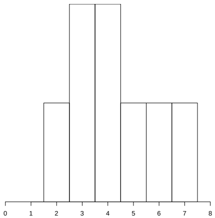
List: 2, 3, 3, 4, 4, 5, 6, 7
average $= 4.25$
SD $= 1.56$



# What does the SD measure?

**The standard deviation (SD) measures roughly how far off the entries are from their average.**

List: 2, 3, 3, 4, 4, 5, 6, 7
average = 4.25
SD = 1.56

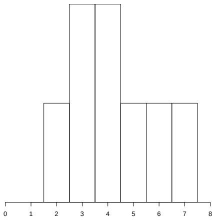



The interval **average ± SD** is roughly 2.75 to 5.75.

# What does the SD measure?

**The standard deviation (SD) measures roughly how far off the entries are from their average.**

List: 2, 3, 3, 4, 4, 5, 6, 7
average $= 4.25$
SD $= 1.56$

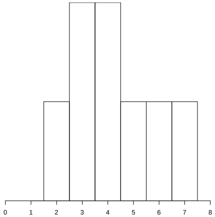



The interval **average $\pm$ SD** is roughly 2.75 to 5.75.
It picks up a good chunk of the list, but not all.

# What does the SD measure?

**The standard deviation (SD) measures roughly how far off the entries are from their average.**

List: 2, 3, 3, 4, 4, 5, 6, 7
average = 4.25
SD = 1.56




The interval **average ± SD** is roughly 2.75 to 5.75.
It picks up a good chunk of the list, but not all.

No matter what the list, the vast majority of entries will be in the range **average ± a few SDs.**

# The formula in two languages

# The formula in two languages

SD $=$

# The formula in two languages

SD = average

# The formula in two languages

SD = deviations from average

# The formula in two languages

SD = square of deviations from average

# The formula in two languages

SD = mean square of deviations from average

# The formula in two languages

SD = root mean square of deviations from average

# The formula in two languages

**The formula in English**

SD = root mean square of deviations from average

# The formula in two languages

**The formula in English**

SD = root mean square of deviations from average

**In another language,**

# The formula in two languages

**The formula in English**

SD = root mean square of deviations from average

**In another language,** the list might be called $x_1, x_2, \ldots, x_n$,

# The formula in two languages

**The formula in English**
SD = root mean square of deviations from average

**In another language,** the list might be called $x_1, x_2, \ldots, x_n$, the average $\bar{x}$,

# The formula in two languages

**The formula in English**
SD = root mean square of deviations from average

**In another language,** the list might be called $x_1, x_2, \ldots, x_n$, the average $\bar{x}$, and then

# The formula in two languages

**The formula in English**
SD = root mean square of deviations from average

**In another language,** the list might be called $x_1, x_2, \ldots, x_n$, the average $\bar{x}$, and then

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

# The formula in two languages

**The formula in English**

SD = root mean square of deviations from average

**In another language,** the list might be called $x_1, x_2, \ldots, x_n$, the average $\bar{x}$, and then

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

The calculation is the same in both languages.

# The vexed question of *n-1*

Some of you might have seen the “definition”

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

# The vexed question of *n-1*

Some of you might have seen the "definition"

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

Please don't use it in descriptive statistics; that's not what it's for.

# The vexed question of *n-1*

Some of you might have seen the "definition"

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

Please don't use it in descriptive statistics; that's not what it's for.

In some situations where you are trying to use the SD of a sample to estimate the SD of the population from which the sample was drawn, then according to some criteria, it might be better to use $n-1$ instead of $n$ in the denominator.

# The vexed question of *n-1*

Some of you might have seen the "definition"

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

Please don't use it in descriptive statistics; that's not what it's for.

In some situations where you are trying to use the SD of a sample to estimate the SD of the population from which the sample was drawn, then according to some criteria, it might be better to use $n - 1$ instead of $n$ in the denominator.

But unless you're in those situations (and you won't be, in Stat 2.1X), there's no need to fuss about using $n - 1$.

# A powerful little list

**0, 1**

# A powerful little list

**0, 1**

**average** $= 0.5$

# A powerful little list

**0, 1**

**average** $= 0.5$
**deviations:** $0 - 0.5 = -0.5$, and $1 - 0.5 = 0.5$

# A powerful little list

**0, 1**

**average** $= 0.5$
**deviations:** $0 - 0.5 = -0.5$, and $1 - 0.5 = 0.5$

**The SD is the rough size of the deviations.**

# A powerful little list

**0, 1**

**average** $= 0.5$
**deviations:** $0 - 0.5 = -0.5$, and $1 - 0.5 = 0.5$

**The SD is the rough size of the deviations.**
What's the rough size of $-0.5$ and $0.5$?

# A powerful little list

**0, 1**

**average** $= 0.5$
**deviations:** $0 - 0.5 = -0.5$, and $1 - 0.5 = 0.5$

**The SD is the rough size of the deviations.**
What's the rough size of $-0.5$ and $0.5$? **Surely it's** $0.5$?

# A powerful little list

**0, 1**

**average** $= 0.5$
**deviations:** $0 - 0.5 = -0.5$, and $1 - 0.5 = 0.5$

**The SD is the rough size of the deviations.**
What's the rough size of $-0.5$ and $0.5$? **Surely it's** $0.5$?
Check, using the **root mean square of deviations:**

# A powerful little list

**0, 1**

**average** $= 0.5$
**deviations:** $0 - 0.5 = -0.5$, and $1 - 0.5 = 0.5$

**The SD is the rough size of the deviations.**
What's the rough size of $-0.5$ and $0.5$? **Surely it's** $0.5$?
Check, using the **root mean square of deviations:**

$$\text{SD} = \sqrt{\frac{(-0.5)^2 + 0.5^2}{2}} = \sqrt{\frac{2 \times 0.5^2}{2}} = \sqrt{0.5^2} = 0.5$$

# A powerful little list

**0, 1**

**average** $= 0.5$
**deviations:** $0 - 0.5 = -0.5$, and $1 - 0.5 = 0.5$

**The SD is the rough size of the deviations.**
What's the rough size of $-0.5$ and $0.5$? **Surely it's** $0.5$?
Check, using the **root mean square of deviations:**

$$\text{SD} = \sqrt{\frac{(-0.5)^2 + 0.5^2}{2}} = \sqrt{\frac{2 \times 0.5^2}{2}} = \sqrt{0.5^2} = 0.5$$

**Find the average and SD of the list** 480, 480, 480, 500, 500, 500.

# A powerful little list

**0, 1**

**average** $= 0.5$
**deviations:** $0 - 0.5 = -0.5$, and $1 - 0.5 = 0.5$

**The SD is the rough size of the deviations.**
What's the rough size of $-0.5$ and $0.5$? **Surely it's** $0.5$?
Check, using the **root mean square of deviations:**

$$\text{SD} = \sqrt{\frac{(-0.5)^2 + 0.5^2}{2}} = \sqrt{\frac{2 \times 0.5^2}{2}} = \sqrt{0.5^2} = 0.5$$

**Find the average and SD of the list** 480, 480, 480, 500, 500, 500.

- Half the list is 480, and the other half is 500. So the **average is** 490.

# A powerful little list

**0, 1**

**average** $= 0.5$
**deviations:** $0 - 0.5 = -0.5$, and $1 - 0.5 = 0.5$

**The SD is the rough size of the deviations.**
What's the rough size of $-0.5$ and $0.5$? **Surely it's** $0.5$?
Check, using the **root mean square of deviations:**

$$\text{SD} = \sqrt{\frac{(-0.5)^2 + 0.5^2}{2}} = \sqrt{\frac{2 \times 0.5^2}{2}} = \sqrt{0.5^2} = 0.5$$

**Find the average and SD of the list** 480, 480, 480, 500, 500, 500.

- Half the list is 480, and the other half is 500. So the **average is** 490.
- So all the deviations are either $-10$ or 10. So the **SD is** 10.

# A powerful little list

**0, 1**

**average** $= 0.5$
**deviations:** $0 - 0.5 = -0.5$, and $1 - 0.5 = 0.5$

**The SD is the rough size of the deviations.**
What's the rough size of $-0.5$ and $0.5$? **Surely it's** $0.5$?
Check, using the **root mean square of deviations:**

$$\text{SD} = \sqrt{\frac{(-0.5)^2 + 0.5^2}{2}} = \sqrt{\frac{2 \times 0.5^2}{2}} = \sqrt{0.5^2} = 0.5$$

**Find the average and SD of the list** 480, 480, 480, 500, 500, 500.

- Half the list is 480, and the other half is 500. So the **average is** 490.
- So all the deviations are either $-10$ or 10. So the **SD is** 10. **Done!**