Solutions

1. Adam is building a doorway and wants the height of the door to be three standard deviations above the mean Canadian height. How high must the door be if the mean is 210 cm with a standard deviation of 10 cm?
   - A 230 cm
   - B 250 cm
   - C 200 cm
   - D 240 cm

D

= Mean Height + 3 x SD

= 210 + 3(10)

= 210 + 30

= 240 cm

2. Which is an incorrect statement about standard deviation?

A The variance is the square root of the standard deviation.

B The standard deviation is often called the average distance of the measurements from the mean.

C The standard deviation is expressed in the same units as the data.

D The standard deviation is always a positive quantity.

The standard deviation is the square root of the variance, not the other way around.

A

3. The mean of a data set is 25.3 cm, and the standard deviation is 3.6. Determine the z-score of each of the following and interpret the results.

a) 27.2

b) 24.1

c) 21.9

d) 29.8

$$z = \frac{x - \overline{x}}{s}$$

x = value

$\overline{x}$ = mean

s = standard deviation

a) z = (27.2 - 25.3)/3.6

z = 1.9/3.6

z = 0.5278

27.2 is 0.5278 standard deviations GREATER than the mean.

b) z = (24.1 - 25.3)/3.6

z = -1.2/3.6

z = -0.3333

24.1 is 0.3333 standard deviations LESS than the mean.

b) z = (21.9 - 25.3)/3.6

z = -3.4/3.6

z = -0.9444

21.9 is 0.9444 standard deviations LESS than the mean.

d) z = (29.8 - 25.3)/3.6

z = 4.5/3.6

z = 1.25

29.8 is 1.25 standard deviations GREATER than the mean.

4. Calculate the standard deviation for each data set and interpret the results.

a) Lengths, in centimetres, of fish caught on a fishing trip.

15.4	12.3	18.2	9.9
17.4	12.6	16.3	11.8
12.3	12.6	16.7	

Population standard deviation:

$$\sigma = \sqrt{\frac{\sum x^2 - N \cdot \mu^2}{N}}$$

desmos

$f = [15.4, 12.3, 18.2, 9.9, 17.4, 12.6, 16.3, 11.8$

N f = 11 element list

mean(f)

μ = 14.1363636364

mean$(f)^2$

μ^2 = 199.83677686

total$(f^{(2)})$

$\sum x^2$ = 2273.09

$\sigma = \sqrt{[(2273.09 - 11(14.13636^2)) / 11]}$

$\sigma = \sqrt{[(2273.09 - 2198.203415) / 11]}$

$\sigma = \sqrt{[74.88658545 / 11]}$

$\sigma = \sqrt{6.807871405}$

$\sigma = 2.609189799$

$\sigma = 2.6092$ cm

4. Calculate the standard deviation for each data set and interpret the results.

a) Lengths, in centimetres, of fish caught on a fishing trip.

15.4	12.3	18.2	9.9
17.4	12.6	16.3	11.8
12.3	12.6	16.7	

Population standard deviation:

$$\sigma = \sqrt{\frac{\sum x^2 - N \cdot \mu^2}{N}}$$

μ = sum of lengths / N

μ = 14.1363636...

$\sum x^2$ = sum of lengths^2

$\sum x^2$ = 2273.09

N = 11

Excel

Length	Length^2
15.4	237.16
17.4	302.76
12.3	151.29
12.3	151.29
12.6	158.76
12.6	158.76
18.2	331.24
16.3	265.69
16.7	278.89
9.9	98.01
11.8	139.24

$\mu = 155.5 / 11$

$\mu = 14.13636...$

$\mu = 14.136$ cm

$\sigma = \sqrt{[(2273.09 - 11(14.13636^2)) / 11]}$

$\sigma = \sqrt{[(2273.09 - 2198.203415) / 11]}$

$\sigma = \sqrt{[74.88658545 / 11]}$

$\sigma = \sqrt{6.807871405}$

$\sigma = 2.609189799$

$\sigma = 2.6092$ cm

4. Calculate the standard deviation for each data set and interpret the results.

b) Number of home runs in a season by the players on a team.

3 10 0 12 5 6 10 16 34 11
6 7 21

Population standard deviation:

$$\sigma = \sqrt{\frac{\sum x^2 - N \cdot \mu^2}{N}}$$

desmos

$g = [3,10,0,12,5,6,10,16,34,11,6,7,21]$

N g = 13 element list

mean(g)

μ = 10.8461538462

mean$(g)^2$

μ^2 = 117.639053254

total$(g^{(2)})$

$\sum x^2$ = 2473

$\sigma = \sqrt{[(2473 - 13(10.84615^2)) / 13]}$

$\sigma = \sqrt{[(2473 - 1529.306608) / 13]}$

$\sigma = \sqrt{[943.6933923 / 13]}$

$\sigma = \sqrt{72.59179941}$

$\sigma = 8.520082...$

$\sigma = 8.5201$ HR

---

4. Calculate the standard deviation for each data set and interpret the results.

b) Number of home runs in a season by the players on a team.

3 10 0 12 5 6 10 16 34 11
6 7 21

Population standard deviation:

$$\sigma = \sqrt{\frac{\sum x^2 - N \cdot \mu^2}{N}}$$

μ = sum of HR / N

μ = 10.84615...

$\sum x^2$ = sum of HR^2

$\sum x^2$ = 2473

N = 13

Excel

HR	HR^2
3	9
6	36
10	100
7	49
0	0
21	441
12	144
5	25
6	36
10	100
16	256
34	1156
11	121

μ = 141 / 13

μ = 10.84615...

μ = 10.846 HR

$\sigma = \sqrt{[(2473 - 13(10.84615^2)) / 13]}$

$\sigma = \sqrt{[(2473 - 1529.306608) / 13]}$

$\sigma = \sqrt{[943.6933923 / 13]}$

$\sigma = \sqrt{72.59179941}$

$\sigma = 8.520082...$

$\sigma = 8.5201$ HR

4. Calculate the standard deviation for each data set and interpret the results.

c) Final scores by the figure skaters in a competition.

168.3	178.2	186.1	134.5
156.7	156.4	167.1	132.0
154.7	149.8	126.2	134.8
154.0	175.2	159.2	

Population standard deviation:

$$\sigma = \sqrt{\frac{\sum x^2 - N \cdot \mu^2}{N}}$$

desmos

$h = [168.3, 178.2, 186.1, 134.5, 156.7, 156.4, 1$

N h = 15 element list

mean(h)

μ = 155.546666667

mean(h)^2

μ^2 = 24194.7655111

total($h^{(2)}$)

$\sum x^2$ = 367391.14

$\sigma = \sqrt{[(367391.14 - 15(155.54666^2)) / 15]}$

$\sigma = \sqrt{[(367391.14 - 362921.4516) / 15]}$

$\sigma = \sqrt{[4469.688443 / 15]}$

$\sigma = \sqrt{297.9792295}$

$\sigma = 17.26207489...$

$\sigma = 17.2621$

---

4. Calculate the standard deviation for each data set and interpret the results.

c) Final scores by the figure skaters in a competition.

168.3	178.2	186.1	134.5
156.7	156.4	167.1	132.0
154.7	149.8	126.2	134.8
154.0	175.2	159.2	

Population standard deviation:

$$\sigma = \sqrt{\frac{\sum x^2 - N \cdot \mu^2}{N}}$$

μ = sum of Scores / N

μ = 155.54666...

$\sum x^2$ = sum of Scores^2

$\sum x^2$ = 367391.14

N = 15

Score	Score^2
168.3	28324.9
156.7	24554.9
154.7	23932.1
154	23716
178.2	31755.2
156.4	24461
149.8	22440
175.2	30695
186.1	34633.2
167.1	27922.4
126.2	15926.4
159.2	25344.6
134.5	18090.3
132	17424
134.8	18171

Excel

μ = 2333 / 15

μ = 155.54666...

μ = 155.5467

$\sigma = \sqrt{[(367391.14 - 15(155.54666^2)) / 15]}$

$\sigma = \sqrt{[(367391.14 - 362921.4516) / 15]}$

$\sigma = \sqrt{[4469.688443 / 15]}$

$\sigma = \sqrt{297.9792295}$

$\sigma = 17.26207489...$

$\sigma = 17.2621$

5. For each of the situations, decide whether you would use the sample or population standard deviation formula. Explain your decisions.

a) A researcher recruits females ages 35 to 50 years old for an exercise training study to investigate risk markers for heart disease (e.g., cholesterol).

b) One of the questions on a national survey asks for the respondent's age. Researchers want to describe the variability in all ages received from the survey.

c) A teacher administers a test to her students. The teacher wants to summarize the results the students attained as a mean and standard deviation.

a) Use the **SAMPLE** standard deviation formula. Researchers are not going to ask every female aged 35 to 50, so they are going to ask a sample of them.

b) Use the **POPULATION** standard deviation formula. It is a national survey, so they are asking everybody.

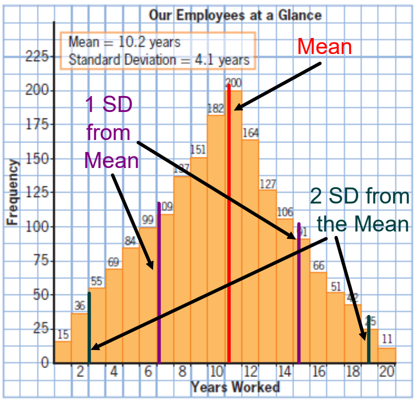
c) Use the **POPULATION** standard deviation formula. The teacher is using results from everyone who took the test (the whole class).

6. As part of a report on its employees, a company published this graph.

a) The standard deviation is given as 4.1 years. Identify which numbers of years worked are within one standard deviation of the mean.

b) What percent of the employees are within two standard deviations of the mean?

c) How does this graph help to explain z-scores?



c) The graph helps to explain z-scores by allowing us to visualize the number of standard deviations an observation (value) is from the mean.

a) Lower limit = Mean - SD

= 10.2 - 4.1

= 6.1 years.

Upper limit = Mean + SD

= 10.2 + 4.1

= 14.3 years.

b) Lower limit = Mean - 2 x SD

= 10.2 - 2(4.1)
= 10.2 - 8.2
= 2.0 years.

Upper limit = Mean + 2 xSD

= 10.2 + 2(4.1)
= 10.2 + 8.2
= 18.4 years.

Total number of employees is 1826 (total of numbers on top of each column).

There are a total of 1775 employees who have worked from between 2 and 19 years.

So there are 1775/1826 = 0.9721

97% of employees are within two standard deviations of the mean.

10. The actual volume of milk in 1-L cartons of milk was checked by measuring a selection of 120 cartons. The chart shows the results.

a) Calculate the mean and standard deviation, accurate to three decimal places.

b) Did you use the population or sample formulas? Why?

c) The company has decided that a sample that is within two standard deviations of the mean is acceptable. A random sample was taken and the volume was 0.98 L. Would this be an acceptable sample?

d) On the following day, the mean volume of milk per carton was 1.012 L, with a standard deviation of 0.009 L. Compare the two days' test results.

Volume (L)	Frequency
0.98	6
0.99	18
1.00	30
1.01	35
1.02	19
1.03	9
1.04	0
1.05	3

a) Sample standard deviation:

$$s = \sqrt{\frac{\sum x^2 - n \cdot \overline{x}^2}{n-1}}$$

n = 120

$\overline{x}$ = 1.0070833 $\overline{x}$ = Sum of (Vol x Freq) / n

$\sum x^2$ = 121.7309 $\sum x^2$ = Sum of (Freq x Vol²)

s = √[(121.7309 - 120(1.0070833²)) / (120 - 1)]

s = √[(121.7309 - 121.7060128) / 119]

s = √[0.024879972 / 119]

s = √0.0002090753978

s = 0.014459439

s = 0.014 L

$\overline{x}$ = 1.007 L

b) Using a selection of 120 cartons so this is a **SAMPLE**.

c) We can use the z-score to help us with this.

$$z = \frac{x - \overline{x}}{s} = \frac{0.98 - 1.007}{0.014} \approx -1.929$$

The random sample is within 2 standard deviations (just) of the mean, so it would be an acceptable sample.

11. The table shows the lengths of logs shipped to a lumber mill on a particular day.

a) Calculate the mean and standard deviation of the logs, accurate to three decimal places.

b) How does this data set compare to the previous day, with a mean of 8.44 m and standard deviation of 1.836 m?

c) Why would the standard deviation be important to the operators of the lumber mill?

Length (m)	Frequency
3.5–4.5	3
4.5–5.5	20
5.5–6.5	17
6.5–7.5	38
7.5–8.5	31
8.5–9.5	19
9.5–10.5	15

We have a grouped data set for this question so we are going to need the midpoint of each group to help find the mean.

a) Sample standard deviation:

$$s = \sqrt{\frac{\sum x^2 - n \cdot \overline{x}^2}{n-1}}$$

$\overline{x}$ = Sum of (Midpoint length x Freq) / n

$\overline{x}$ = 7.3356643

n = 143

$\sum x^2$ = Sum of (Freq x Midpoint length²)

$\sum x^2$ = 8045

s = √[(8045 - 143(7.3356643²)) / (143 - 1)]

s = √[(8045 - 7690.706644) / 142]

s = √[349.8881867 / 142]

s = √2.464001315

s = 1.569713768

s = 1.570 m

$\overline{x}$ = 7.336 m

b) Today's mean length had decreased compared to the previous day's (7.336m vs 8.44m). The standard deviation had also decreased (1.570m vs 1.836m). Today's lengths are less spread out.

c) The standard deviation is important to operators in that it can help to identify calibration problems with equipment.