5.5 Z-Scores

GOAL

Use *z*-scores to compare data, make predictions, and solve problems.

YOU WILL NEED

- calculator
- grid paper
- *z*-score tables (pages 592 to 593)

LEARN ABOUT the Math

Hailey and Serge belong to a running club in Vancouver. Part of their training involves a 200 m sprint. Below are normally distributed times for the 200 m sprint in Vancouver and on a recent trip to Lake Louise. At higher altitudes, run times improve.

Location	Altitude (m)	Club Mean Time: μ for 200 m (s)	Club Standard Deviation: σ (s)	Hailey's Run Time (s)	Serge's Run Time (s)
Vancouver	4	25.75	0.62	24.95	25.45
Lake Louise	1661	25.57	0.60	24.77	26.24

EXPLORE...

- Alexis plays in her school jazz band. Band members practise an average of 16.5 h per week, with a standard deviation of 4.2 h. Alexis practises an average of 22 h per week. As a class, discuss how you might estimate the percent of the band that, on average, practises a greater number of hours than Alexis.

? At which location was Hailey's run time better, when compared with the club results?

EXAMPLE 1 Comparing *z*-scores

Determine at which location Hailey's run time was better, when compared with the club results.

Marcel's Solution

For any given score, x, from a normal distribution,

$x = \mu + z\sigma$,

where z represents the number of standard deviations of the score from the mean.

Solving for z results in a formula for a **z-score**:

$$z = \frac{x - \mu}{\sigma}$$

Hailey's run time is less at Lake Louise, but so is the club's mean run time. I can't compare these times directly, because the means and standard deviations are different for the two locations. To make the comparison, I have to standardize Hailey's times to fit a common normal distribution.

A *z*-score indicates the position of a data value on a **standard normal distribution**.

***z*-score**

A standardized value that indicates the number of standard deviations of a data value above or below the mean.

standard normal distribution

A normal distribution that has a mean of zero and a standard deviation of one.

Vancouver:

$$z = \frac{x - \mu}{\sigma}$$

$$z = \frac{24.95 - 25.75}{0.62}$$

$$z = -1.290 \ldots$$

Lake Louise:

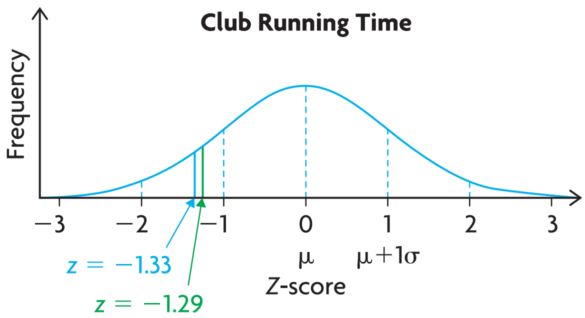
$$z = \frac{x - \mu}{\sigma}$$

$$z = \frac{24.77 - 25.57}{0.60}$$

$$z = -1.333 \ldots$$

I know that *z*-scores can be used to compare data values from different normal distributions. I calculated the *z*-score for Hailey's run times at each location.

Hailey's run time is about 1.29 standard deviations below the mean in Vancouver, and 1.33 standard deviations below the mean in Lake Louise.



I sketched the standard normal curve, which has a mean of zero and a standard deviation of 1. Then I drew a line on the graph for each *z*-score.

The *z*-score for Hailey's Lake Louise run is farther to the left than the *z*-score for her Vancouver run.

Hailey's time for 200 m was better than the club's mean in both locations. However, Hailey's *z*-score for Lake Louise was lower than her *z*-score for Vancouver, so her time was better in Lake Louise.

I can make this comparison because both times have been translated to a normal distribution that has the same mean and standard deviation.

Reflecting

A. Use *z*-scores to determine which of Serge's runs was better.

B. Explain why the lower *z*-score represents a relatively faster run.

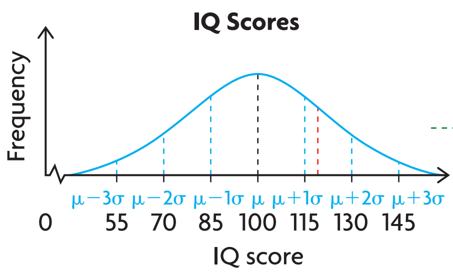
C. What can you say about a data value if you know that its *z*-score is negative? positive? zero?

APPLY the Math

EXAMPLE 2 Using *z*-scores to determine the percent of data less than a given value

IQ tests are sometimes used to measure a person's intellectual capacity at a particular time. IQ scores are normally distributed, with a mean of 100 and a standard deviation of 15. If a person scores 119 on an IQ test, how does this score compare with the scores of the general population?

Malia's Solution: Using a *z*-score table



First, I sketched a normal curve and determined the IQ scores for one, two, and three standard deviations from the mean.

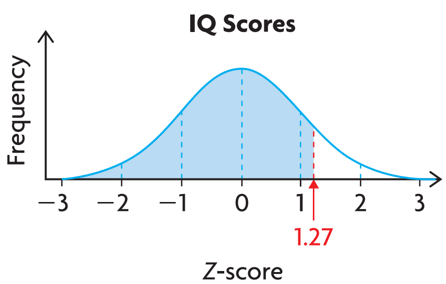
Then I drew a line that represented an IQ score of 119.

I noticed that my line was between one and two standard deviations above the mean.

$$z = \frac{x - \mu}{\sigma}$$

$$z = \frac{119 - 100}{15}$$

$$z = 1.2666...$$



I determined the *z*-score for an IQ of 119.

An IQ score of 119 is about 1.27 standard deviations above the mean. I sketched this on a standard normal curve.

I knew that I needed to determine the percent of people with IQ scores less than 119. This is equivalent to the area under the curve to the left of 1.27 on the standard normal curve.

z-score table
A table that displays the fraction of data with a z-score that is less than any given data value in a standard normal distribution.
(There is a z-score table on pages 592 to 593.)

z	0.0	0.01	0.06	0.07
0.0	0.5000	0.5040	0.5239	0.5279
0.1	0.5398	0.5438	0.5636	0.5675
1.1	0.8643	0.8665	0.8770	0.8790
1.2	0.8849	0.8869	0.8962	**0.8980**
1.3	0.9032	0.9049	0.9131	0.9147

I used a ***z*-score table**.

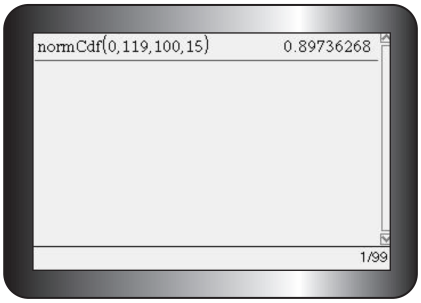
$1.27 = 1.2 + 0.07$

I used the 1.2 row and the 0.07 column.

The value in the table, 0.8980, is the fraction of the area under the curve to the left of the z-score.

The value in the z-score table is 0.8980. This means that an IQ score of 119 is greater than 89.80% of IQ scores in the general population.

Desiree's Solution: Using a graphing calculator



I used the statistics function for normal distributions on my calculator to determine the percent of the population that has an IQ score between 0 and 119.

I entered the lower bound of 0, the upper bound of 119, the mean of 100, and the standard deviation of 15.

An IQ score of 119 is greater than 89.74% of all the scores.

My solution is slightly different from Malia's because this method does not use a rounded z-score.

Your Turn

Megan determined the area under the normal curve using slightly different reasoning: "I know that the total area under a normal curve is 100%, so the area under the curve to the left of the mean is 50%. I used my graphing calculator to calculate the area between a z-score of 0 and a z-score of 1.27, by entering these as the lower and upper bounds. My calculator gave a result of 0.397... ."

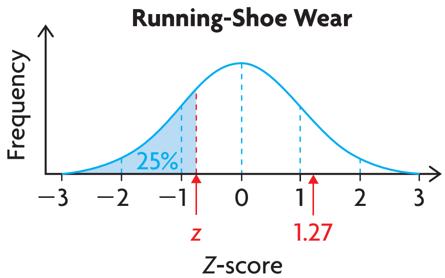
How could Megan use this result to complete her solution?

EXAMPLE 3 Using *z*-scores to determine data values

Athletes should replace their running shoes before the shoes lose their ability to absorb shock.

Running shoes lose their shock-absorption after a mean distance of 640 km, with a standard deviation of 160 km. Zack is an elite runner and wants to replace his shoes at a distance when only 25% of people would replace their shoes. At what distance should he replace his shoes?

Rachelle's Solution: Using a *z*-score table



I sketched the standard normal curve. I needed the *z*-score for 25% of the area under the curve, or 0.25.

z	0.09	0.08	0.07	0.06	0.05
–0.7	0.2148	0.2177	0.2206	0.2236	0.2266
–0.6	0.2451	0.2483	0.2514	0.2546	0.2578
–0.5	0.2776	0.2810	0.2843	0.2877	0.2912

I searched the *z*-score table for a value that is close to 0.25.

The z-score that represents an area of 0.25 is about halfway between -0.67 and -0.68, or about -0.675.

$$z = \frac{x - \mu}{\sigma}$$

$$(-0.675) = \frac{x - (640)}{(160)}$$

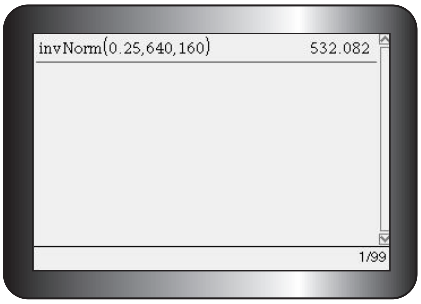
$$-108 = x - 640$$

$$532 = x$$

I substituted the values I knew into the *z*-score formula and solved for *x*.

Zack should replace his running shoes after 532 km.

Renalda's Solution: Using a graphing calculator



I used the statistics function on my calculator.

I entered the decimal value for the percent of data to the left of the *z*-score I needed. Then I entered the mean and standard deviation of the data.

Zack should replace his running shoes after 532 km.

Therefore, 25% of people would replace their shoes after 532 km.

Your Turn

Quinn is a recreational runner. He plans to replace his running shoes when 70% of people would replace their shoes. After how many kilometres should he replace his running shoes?

EXAMPLE 4 Solving a quality control problem

The ABC Company produces bungee cords. When the manufacturing process is running well, the lengths of the bungee cords produced are normally distributed, with a mean of 45.2 cm and a standard deviation of 1.3 cm. Bungee cords that are shorter than 42.0 cm or longer than 48.0 cm are rejected by the quality control workers.

a) If 20 000 bungee cords are manufactured each day, how many bungee cords would you expect the quality control workers to reject?

b) What action might the company take as a result of these findings?

Logan's Solution: Using a *z*-score table

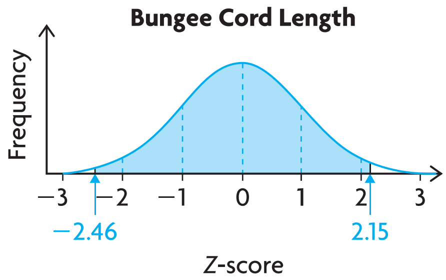
a) Minimum length = 42 cm Maximum length = 48 cm

$$z_{min} = \frac{x - \mu}{\sigma} \qquad z_{max} = \frac{x - \mu}{\sigma}$$

$$z_{min} = \frac{42.0 - 45.2}{1.3} \qquad z_{max} = \frac{48.0 - 45.2}{1.3}$$

$$z_{min} = -2.461... \qquad z_{max} = 2.153...$$

I determined the *z*-scores for the minimum and maximum acceptable lengths.



I sketched the standard normal curve. The area under the curve to the left of −2.46 represents the percent of rejected bungee cords less than 42 cm. The area under the curve to the right of 2.15 represents the percent of rejected bungee cords greater than 48 cm.

Area to left of $-2.46 = 0.0069$

Area to right of $2.15 = 1 - 0.9842$
Area to right of $2.15 = 0.0158$

I looked up each *z*-score in the *z*-score table. The *z*-score table gives the area to the left of the *z*-score, which I want for 42 cm.

Since I wanted the area to the right of the *z*-score for 48 cm, I had to subtract the corresponding area from 1.

Percent rejected = Area to the left of −2.46 + Area to the right of 2.15
Percent rejected $= 0.0069 + 0.0158$
Percent rejected $= 0.0227$ or 2.27%

I added the two areas to determine the percent of bungee cords that are rejected.

Total rejected $= (0.0227)(20\,000)$ or 454

I determined the number of bungee cords that are rejected.

b) ABC needs a more consistent process, because 454 seems like a large number of bungee cords to reject. The company should adjust its equipment so that the standard deviation is lowered.

Lowering the standard deviation will reduce the percent of rejected bungee cords.

Nathan's Solution: Using a graphing calculator

a)

I used the statistics function on my calculator to determine the percent of bungee cords that are an acceptable length. I entered the minimum and maximum acceptable lengths and then the mean and standard deviation.

Number accepted $= 20\,000 \times 0.977...$
Number accepted $= 19\,549.135...$
About 19 549 bungee cords meet the standard every day, so 451 bungee cords are rejected every day.

I determined the number of bungee cords that meet the standard. Then I subtracted to determine the number rejected.

My solution is slightly different from Logan's solution because this method does not use a rounded *z*-score value.

b) I think the company should adjust its equipment to get a lower standard deviation, so fewer bungee cords are discarded.

Your Turn

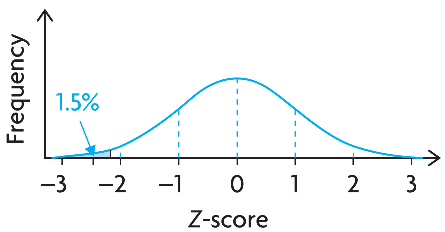
a) What percent of all the bungee cords are accepted?
b) A client has placed an order for 12 000 bungee cords, but will only accept bungee cords that are between 44.0 cm and 46.0 cm in length. Can this client's order be filled by one day's production, with the equipment operating as is? Explain.

EXAMPLE 5 Determining warranty periods

A manufacturer of personal music players has determined that the mean life of the players is 32.4 months, with a standard deviation of 6.3 months. What length of warranty should be offered if the manufacturer wants to restrict repairs to less than 1.5% of all the players sold?

Sacha's Solution

$1.5\% = 0.015$



I used my graphing calculator to determine the *z*-score that corresponds to an area under the normal curve of 0.015.

$z = -2.17$

$$z = \frac{x - \mu}{\sigma}$$

$$(-2.17) = \frac{x - (32.4)}{(6.3)}$$

$$-13.671 = x - 32.4$$

$$18.729 = x$$

I substituted the known values into the *z*-score formula and solved for *x*.

The manufacturer should offer an 18-month warranty.

Since the manufacturer wants to repair less than 1.5% of the music players, I rounded down to 18 months.

Your Turn

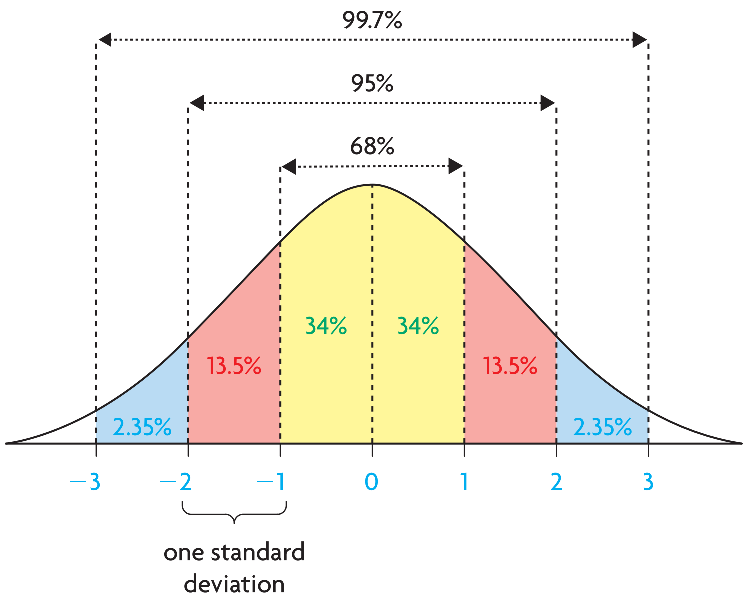
a) If 10 000 personal music players are sold, how many could the manufacturer expect to receive for repairs under warranty?

b) The manufacturer wants to offer the option of purchasing an extended warranty. If the manufacturer wants to repair, at most, 20% of the players under the extended warranty, what length of extended warranty should be offered?

In Summary

Key Ideas

- The standard normal distribution is a normal distribution with mean, μ, of 0 and a standard deviation, σ, of 1. The area under the curve of a normal distribution is 1.
- *Z*-scores can be used to compare data from different normally distributed sets by converting their distributions to the standard normal distribution.



Need to Know

- A *z*-score indicates the number of standard deviations that a data value lies from the mean. It is calculated using this formula:

$$z = \frac{x - \mu}{\sigma}$$

- A positive *z*-score indicates that the data value lies above the mean. A negative *z*-score indicates that the data value lies below the mean.
- The area under the standard normal curve, to the left of a particular *z*-score, can be found in a *z*-score table or determined using a graphing calculator.

CHECK Your Understanding

1. Determine the z-score for each value of x.
   a) $\mu = 112, \sigma = 15.5, x = 174$
   b) $\mu = 53.46, \sigma = 8.24, x = 47.28$
   c) $\mu = 82, \sigma = 12.5, x = 58$
   d) $\mu = 245, \sigma = 22.4, x = 300$
2. Using a z-score table (such as the table on pages 592 to 593), determine the percent of the data to the left of each z-score.
   a) $z = 1.24$ b) $z = -2.35$ c) $z = 2.17$ d) $z = -0.64$
3. Determine the percent of the data between each pair of z-scores.
   a) $z = -2.88$ and $z = -1.47$ b) $z = -0.85$ and $z = 1.64$
4. What z-score is required for each situation?
   a) 10% of the data is to the left of the z-score.
   b) 10% of the data is to the right of the z-score.
   c) 60% of the data is below the z-score.
   d) 60% of the data is above the z-score.

PRACTISING

In the following questions, assume that the data approximates a normal distribution.

5. Calculate the z-score for each value of x.
   a) $\mu = 24, \sigma = 2.8, x = 29.3$
   b) $\mu = 165, \sigma = 48, x = 36$
   c) $\mu = 784, \sigma = 65.3, x = 817$
   d) $\mu = 2.9, \sigma = 0.3, x = 3.4$
6. Determine the percent of the data to the left of each z-score.
   a) $z = 0.56$ b) $z = -1.76$ c) $z = -2.98$ d) $z = 2.39$
7. Determine the percent of the data to the right of each z-score.
   a) $z = -1.35$ b) $z = 2.63$ c) $z = 0.68$ d) $z = -3.14$
8. Determine the percent of the data between each pair of z-scores.
   a) $z = 0.24$ and $z = 2.53$ b) $z = -1.64$ and $z = 1.64$
9. Determine the z-score for each situation.
   a) 33% of the data is to the left of the z-score.
   b) 20% of the data is to the right of the z-score.
10. Meg wonders if she should consider a career in the sciences, because she does well in mathematics. However, she also does well in English and has thought about becoming a journalist.
    a) Determine the z-score for each of Meg's marks.
    b) Which subject is Meg better in, relative to her peers?
    c) What other factors should Meg consider?

Subject	Standard Test Results (%)		Meg's Mark (%)
	μ	σ	
English	77	6.8	93
math	74	5.4	91

11. A hardwood flooring company produces flooring that has an average thickness of 175 mm, with a standard deviation of 0.4 mm. For premium-quality floors, the flooring must have a thickness between 174 mm and 175.6 mm. What percent, to the nearest whole number, of the total production can be sold for premium-quality floors?

Speed	Treadmill (beats/min)		Water (beats/min)	
	μ	σ	μ	σ
Resting	68	8.43	71	6.15
2.55 km/h	76	9.15	130	13.50
2.77 km/h	79	11.66	146	11.96
3.02 km/h	81	11.33	160	13.50
3.31 km/h	81	10.27	167	12.58

12. Violeta took part in a study that compared the heart-rate responses of water walking versus treadmill walking for healthy college females. Violeta's heart rate was 68 on the treadmill for the 2.55 km/h walk and 145 in the water for the 3.02 km/h walk. For which event was her heart rate lower, compared with the others who took part in the study?

13. In 2006, the ages of mothers who had children aged 4 and under were approximately normally distributed, with a mean age of 32 years and a standard deviation of 5.9 years. The data is shown in the table at the right.

a) Determine the percent of mothers who were less than 40 years old.

b) Determine the percent of mothers who were less than 21 years old.

c) Determine the percent of mothers who were 18 years old or less. Why might someone want to know this?

Age of Mother (years)	2006 Census (%)
15–19	1.1
20–24	8.8
25–29	23.2
30–34	33.7
35–39	23.8
40–44	8.2
45–49	1.2
Total	100

Statistics Canada

14. In a population, 50% of the adults are taller than 180 cm and 10% are taller than 200 cm. Determine the mean height and standard deviation for this population.

15. A medical diagnostic test counts the number of blood cells in a sample. The red blood cell count (in millions per cubic microlitre) is normally distributed, with a mean of 4.8 and a standard deviation of 0.3.

a) What percent of people have a red blood cell count that is less than 4?

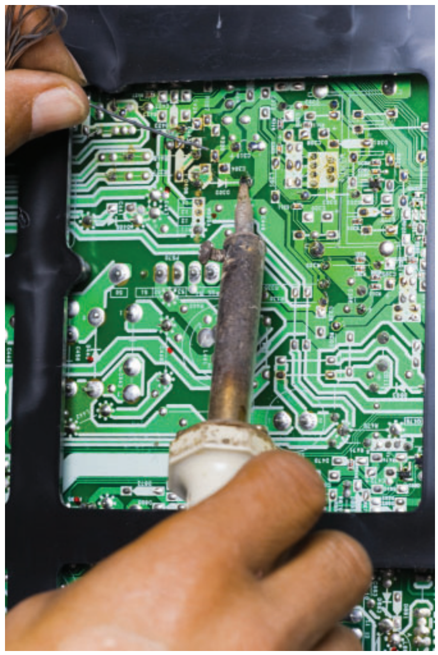
b) What percent of people have a count between 4.7 and 5.0?

c) What red blood cell count would someone have if 95% of people have a lower count?

16. An MP3 player has a one-year warranty. The mean lifespan of the player is 2.6 years, with a standard deviation of 0.48 years.

a) A store sells 4000 players. How many of these players will fail before the warranty expires?

b) Tyler is offered an extended warranty, for one extra year, when he buys a player. What is the likelihood that he will make a claim on this warranty if he takes it?



17. A manufacturer of plasma televisions has determined that the televisions require servicing after a mean of 67 months, with a standard deviation of 7.2 months. What length of warranty should be offered, if the manufacturer wants to repair less than 1% of the televisions under the warranty?

18. A tutor guarantees that 10% of her students will obtain an A on every test they write. For the last test, the mean mark is 68 and the standard deviation is 6. What mark is required to receive an A on the test?

19. In the insurance industry, standard deviation is used to quantify risk—the greater the risk, the higher the standard deviation. For example, consider the cost of a car accident for two different cars: a high-priced luxury car and a mid-priced car. The expected cost of repairs for both cars is $2500. However, the standard deviation for the high-priced car is $1000, and the standard deviation for the mid-priced car is $400. Explain why the probability that the repairs will cost more than $3000 is 31% for the high-priced car but only 11% for the mid-priced car.

20. **a)** A club accepts members only if they have an IQ score that is greater than the scores for 98% of the population. What IQ score would you need to be accepted into this club? (Recall that $\mu = 100$ and $\sigma = 15$ for the general population.)
b) Only 0.38% of the population are considered to be geniuses, as measured by IQ scores. What is the minimum IQ score that is required to be considered a genius?
c) Jarrod was told that his IQ score is in the top 30% of the population. What is his IQ score?

Closing

21. What is a *z*-score, how do you determine it, and what is it used for?

Extending

22. A company packages sugar into 5 kg bags. The filling machine can be calibrated to fill to any specified mean, with a standard deviation of 0.065 kg. Any bags with masses that are less than 4.9 kg cannot be sold and must be repackaged.
a) If the company wants to repackage no more than 3% of the bags, at what mean should they set the machine?
b) Assuming that the company sets the machine at the mean you determined in part a), what percent of the bags will have more than 5 kg of sugar? Do you think the company will be satisfied with this percent?

The LSAT must be taken by people who want to gain admission to a law school. The test focuses on logical and verbal reasoning skills.

23. Approximately 40% of those who take the LSAT, or Law School Admission Test, score from 145 to 155. About 70% score from 140 to 160.
a) Determine the mean score and standard deviation for the LSAT.
b) Harvard University also uses other methods to choose students for its law school, but the minimum LSAT score that is required is about 172. What percent of people who take the LSAT would be considered by Harvard for admission?

24. Create your own problem involving *z*-score analysis, using any of the normally distributed data from Lesson 5.4. Exchange problems with classmates, and solve the problems. Provide suggestions for improving the problems.