3.2 Exercise **MATH 241** **THOMPSON**

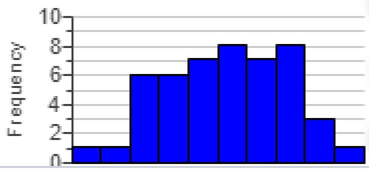
1. For an exam given to a class, the students' scores ranged from 34 to 99, with a mean of 74.

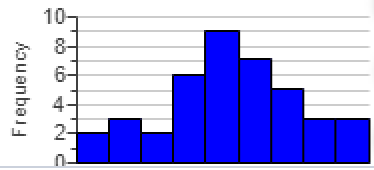
Which of the following is the most realistic value for the standard deviation: -9, 0, 74, 13, 1 ?

Choose the best answer from those given below.

✓ A. The most realistic value is 13, because the negative value is impossible, 0 would indicate no variability, 1 is too small, and 74 is too large for a typical deviation.

2. Which histogram depicts a higher standard deviation?





○ A. Histogram a depicts the higher standard deviation, because the distribution has more dispersion.

○ B. Histogram b depicts the higher standard deviation, since it is more bell shaped.

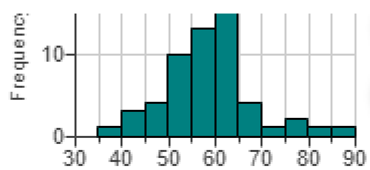
○ C. Histogram a depicts the higher standard deviation, because the bars are higher than the average bar in b.

✓ D. Histogram b depicts the higher standard deviation, because the distribution has more dispersion.

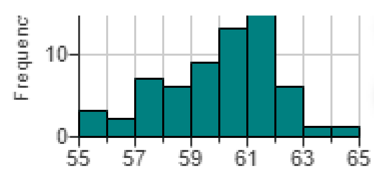
3.

	Mean	Median	Standard Deviation
I	60	60	1.5
II	66	66	11
III	60	60	10
IV	60	60	22

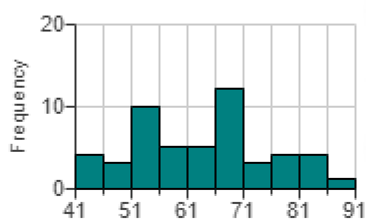
(a) III
(b) II
(c) I
(d) IV



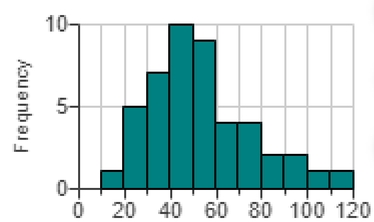
(a)



(c)



(b)



(d)

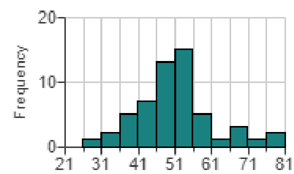
larger the standard deviation, the more dispersion the distribution has, provided that the variable of interest from the two populations has the same unit of measure.

another example

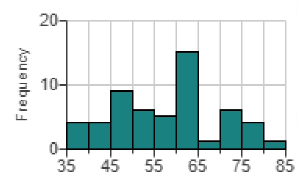
Match the histograms on the right to the summary statistics given.

	Mean	Median	Standard Deviation
I	51	51	1.8
II	60	60	10
III	51	51	9
IV	51	51	19

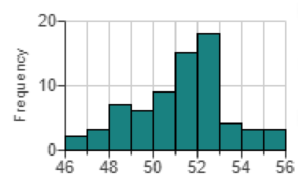
(a) III
(b) I
(c) II
(d) IV

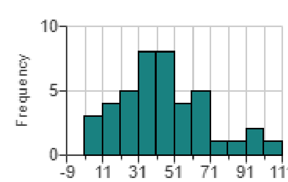


(a)



(c)





$$population\ standard\ deviation:\ \sigma = \sqrt{\frac{(x_1-\mu)^2+(x_2-\mu)^2+\ldots+(x_n-\mu)^2}{n}}$$

$$sample\ standard\ deviation:\ s = \sqrt{\frac{(x_1-\bar{x})^2+(x_2-\bar{x})^2+\ldots+(x_n-\bar{x})^2}{n-1}}$$

4. Find the population variance and standard deviation. Do this by hand.

6, 15, 27, 33, 39

population so use σ

$$\mu = \frac{6+15+27+33+39}{5} = 24$$

Choose the correct answer below. Fill in the answer box to complete your choice.
(Type an integer or a decimal. Do not round.)

A. σ^2 = 144

B. s^2 =

$$\frac{(6-24)^2+(15-24)^2+(27-24)^2+(33-24)^2+(39-24)^2}{5}=144$$

Choose the correct answer below. Fill in the answer box to complete your choice.
(Type an integer or a decimal. Do not round.)

A. σ = 12

$$\sqrt{144} = 12$$

5. The sum of the deviations about the mean always equals zero

σ means standard deviation of the population Unadj. std. dev.
σ^2 means variance of a population Unadj. variance
s means standard deviation of a sample so use Std. dev.
s^2 means variance of a sample so use variance

6. The accompanying data represent the pulse rates (beats per minute) of nine students enrolled in a statistics course. Treat the nine students as a population. .

Click the icon to view the data on the students' pulse rates.

STATCRUNCH

STAT- SUMMARY STATS-COLUMNS
highlight var1
use Ctrl and highlight
σ is Unadj. std. dev

(a) Compute the population standard deviation.

σ = 8.3 beats per minute
(Round to one decimal place as needed.)

7. Ethan and Drew went on a 10-day fishing trip. The number of smallmouth bass caught and released by the two boys each day is shown in the accompanying data table.
Work this problem on your calculator.

Click the icon to view the data table.

(a) Find the population mean and the range for the number of smallmouth bass caught per day by each fisherman. Do these values indicate any differences between the two fishermen's catches per day?

The range for the number of smallmouth bass caught per day by Ethan is 19.
(Type an integer or a decimal. Do not round.)

range = DATA – SORT highest - lowest

The population mean for the number of smallmouth bass caught per day by Ethan is 10.
(Simplify your answer.)

STATS – SUMMARY STATS - mean

The range for the number of smallmouth bass caught per day by Drew is 19.
(Type an integer or a decimal. Do not round.)

The population mean for the number of smallmouth bass caught per day by Drew is 10.
(Simplify your answer.)

Do these values indicate any differences between the two fishermen's catches per day?

- A. Yes. Ethan has a much higher range than Drew.
- B. No. Both Ethan and Drew have a similar population mean and range. ✔
- C. Yes. Ethan has a much lower population mean than Drew.

(b) Find the population standard deviation for the number of smallmouth bass caught per day by each fisherman. Do these values present a different story about the two fishermen's catches per day? Which fisherman has the more consistent record?

The population standard deviation for the number of smallmouth bass caught per day by Ethan is 4.9.
(Round to one decimal place as needed.)

The population standard deviation for the number of smallmouth bass caught per day by Drew is 7.8.

Yes, since the fishermen have similar population means and ranges while having different population standard deviations.

Which fisherman has the more consistent record?

- A. Ethan has the more consistent record, since his standard deviation is lower. ✔

- D. The range only takes into account the smallest and largest values, which does not provide any information about the dispersion of the distribution between those values. ✔

σ means standard deviation of the population Unadj. std. dev.
σ^2 means variance of a population Unadj. variance
s means standard deviation of a sample so use Std. dev.
s^2 means variance of a sample so use variance

8. Find the sample variance and standard deviation.

8, 45, 16, 47, 38, 24, 33, 32, 26, 32

- A. σ^2 =
- B. s^2 = 147.43 ✔

STATCRUNCH

STAT- SUMMARY STATS-COLUMNS
highlight var1
use Ctrl and highlight both
s^2 is Variance
s is Std. dev

Choose the correct answer below. Fill in the answer box to complete your choice.
(Round to one decimal place as needed.)

- A. s = 12.1 ✔
- B. σ =

Compute the range and sample standard deviation for strength of the concrete (in psi).

9. 3970, 4050, 3200, 3000, 2940, 3830, 4050, 4060

The range is 1120 psi.

Use DATA – SORT to get in order

Highest - lowest 4060-2940 = 1120

s = 500.1 psi (Round to one decimal place as needed.)

10. The accompanying data represent the pulse rates (beats per minute) of nine students enrolled in a statistics course. Treat the nine students as a population. Complete parts (a) through (c).

Click the icon to view the data on the students' pulse rates.

STATCRUNCH

STAT- SUMMARY STATS-COLUMNS
highlight var1
use Ctrl and highlight both
Unadj st dv

(a) Compute the population standard deviation.

σ = 7.2 beats per minute

(Round to one decimal place as needed.)

population

(b) Determine the sample standard deviation of the following three simple random samples of size 3.

Sample 1: {Jeff, Megan, Janette}
Sample 2: {Crystal, Megan, Clarice}
Sample 3: {Kathy, Tammy, Kevin}

Make separate columns for all the sample sets and find **Std. dev** for each.

The sample standard deviation, s, of sample 1, {Jeff, Megan, Janette}, is 5.8 beats per minute.
(Round to one decimal place as needed.)

samples

The sample standard deviation, s, of sample 2, {Crystal, Megan, Clarice}, is 3.6 beats pe
(Round to one decimal place as needed.)

The sample standard deviation, s, of sample 3, {Kathy, Tammy, Kevin}, is 9.1 beats per
(Round to one decimal place as needed.)

STATCRUNCH

STAT- SUMMARY STATS-COLUMNS
highlight var1
use Ctrl and highlight both
s is std. dev

(c) Which samples underestimate the population standard deviation? Which overestimate the population standard deviation?

Compare each sample with the population

Sample 1 underestimates the population standard deviation.

Sample 2 underestimates the population standard deviation.

Sample 3 overestimates the population standard deviation.

11. True or False: When comparing two populations, the larger the standard deviation, the more dispersion the distribution has, provided that the variable of interest from the two populations has the same unit of measure.

C. True, because the standard deviation describes how far, on average, each observation is from the typical value. A larger standard deviation means that observations are more distant from the typical value, and therefore, more dispersed.

12. STATS – SUMMARY – COLUMNS – mean and median

Suppose that a customer is purchasing a car. He conducts an experiment in which he puts 10 gallons of gas in the car and drives it until it runs out of gas. He conducts this experiment 15 times on each car and records the number of miles driven.

Full data set

Car 1				
212	237	243	230	229
285	256	170	292	251
154	309	261	317	275

Car 2				
205	205	216	216	258
253	240	265	258	274
272	248	258	260	256

Describe each data set, that is determine the shape, center, and spread.

Sample mean for Car 1

$\bar{x}$ = 248.1 mi / 10 gal

(Type an integer or decimal rounded to one decimal place as needed.)

Sample mean for Car 2

$\bar{x}$ = 245.6 mi / 10 gal

(Type an integer or decimal rounded to one decimal place as needed.)

Median for Car 1

M = 251 mi / 10 gal

(Type an integer or decimal rounded to one decimal place as needed.)

Median for Car 2

M = 256 mi / 10 gal

(Type an integer or decimal rounded to one decimal place as needed.)

Range is highest – lowest
DATA – SORT the column

Range for Car 1

R = 163 mi / 10 gal

(Type an integer or decimal rounded to one decimal place as needed.)

Range for Car 2

R = 69 mi / 10 gal

(Type an integer or decimal rounded to one decimal place as needed.)

s means standard deviation of a sample so use Std. dev.

Sample standard deviation for Car 1

s = 46.1 mi / 10 gal

(Type an integer or decimal rounded to one decimal place as needed.)

Sample standard deviation for Car 2

s = 23.6 mi / 10 gal

(Type an integer or decimal rounded to one

Which car would the customer buy and why?

C. Car 2, because it has a lower sample standard deviation, hence more predictable gas mileage.

13. Blocking refers to the idea that the variability in a variable can be reduced by segmenting the data by some other variable. The data in the accompanying table represent the recumbent length (in centimeters) of a sample of 10 males and 10 females who are 40 months of age. Complete parts (a) through (d).

Click the icon to view the data table.

(a) Determine the standard deviation of recumbent length for all 20 observations.

6.06 cm (Round to two decimal places as needed.)

Use this for all parts

STATCRUNCH

STAT- SUMMARY STATS-COLUMNS
highlight var1
use Ctrl and highlight both
s is std. dev

MAKE A NEW COLUMN AND COPY BOTH MALE AND FEMALE DATE IN A SEPARATE COLUMN FOR PART A

*one long list

(b) Determine the standard deviation of recumbent length for the males.

5.52 cm (Round to two decimal places as needed.)

(c) Determine the standard deviation of recumbent length for the females.

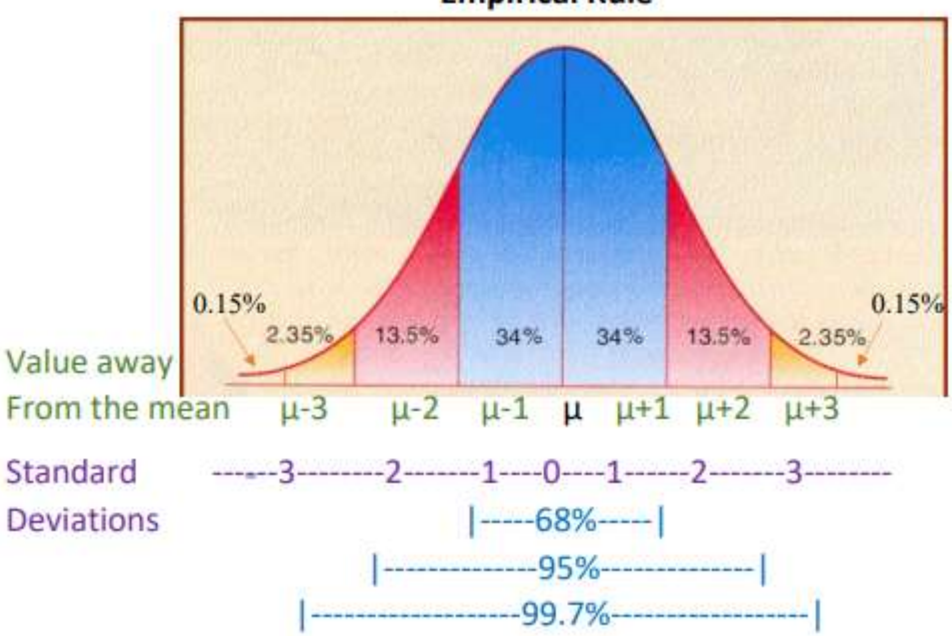
5.61 cm (Round to two decimal places as needed.)

(d) What effect does blocking by gender have on the standard deviation of recumbent length for each gender?

- A. The standard deviation for each group is the same as it is for the groups combined.
- B. The standard deviation is higher for each group than it is for the groups combined.
- ✔ C. The standard deviation is lower for each group than it is for the groups combined.

14. The standard deviation is used in conjunction with the mean to numerically describe distributions that are bell shaped. The mean measures the center of the distribution, while the standard deviation measures the spread of the distribution.

Empirical Rule



- Approximately 68% of the data will lie within 1 standard deviation of the mean.
- Approximately 95% of the data will lie within 2 standard deviations of the mean.
- Approximately 99.7% of the data will lie within 3 standard deviations of the mean.

The standard deviation describes how far, on average, each observation is from the typical value. A larger standard deviation means that observations are more distant from the typical value and more dispersed. YOU HAVE TO REMEMBER THE PERCENTAGES

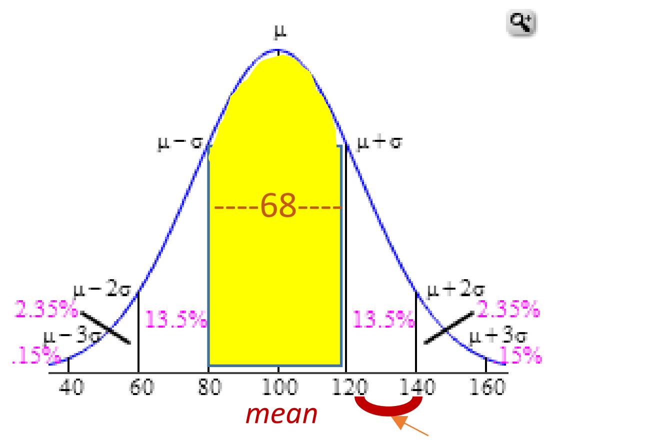
15.

Scores of an IQ test have a bell-shaped distribution with a mean of 100 and a standard deviation of 20. Use the empirical rule
(a) What percentage of people has an IQ score between 80 and 120?
(b) What percentage of people has an IQ score less than 60 or greater than 140?
(c) What percentage of people has an IQ score greater than 160?

Now use the graph to answer the questions.

(a) What percentage of people has an IQ score between 80 and 120?

68 % (Type an integer or a decimal.)

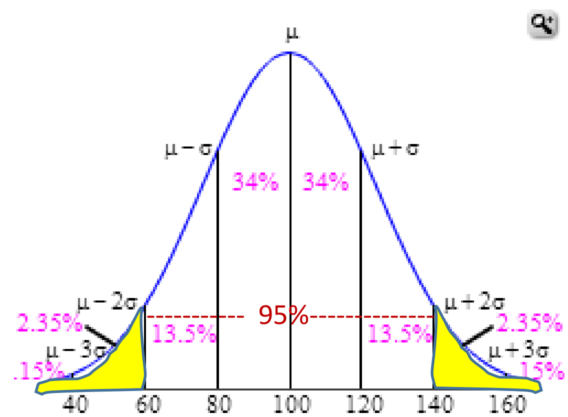


YOU FILL IN THE BOTTOM #S FROM PROBLEM

(b) What percentage of people has an IQ score less than 60 or greater than 140?

5 % (Type an integer or a decimal.)

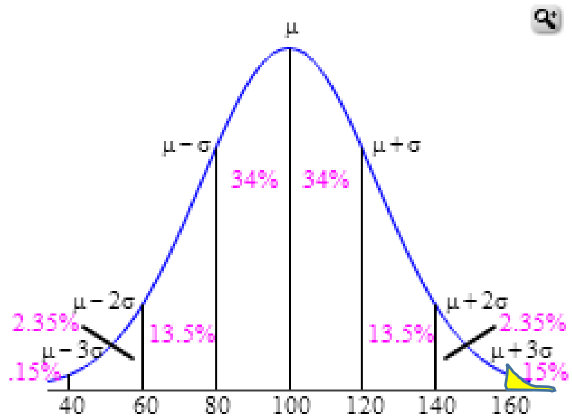
.15+2.35+2.35+.15 = 5 or 100 – 95



YOU HAVE TO REMEMBER THE PERCENTAGES IN THE BELL SHAPE

(c) What percentage of people has an IQ score greater than 160?

.15 % (Type an integer or a decimal.)



16. A certain standardized test's math scores have a bell-shaped distribution with a mean of 520 a standard deviation of 114. Complete parts (a) through (c).

(a) What percentage of standardized test scores is between 406 and 634?

*[Mean – one of the scores] → 520 – 406 = 114 which is **one** standard deviation*

The percentage of standardized test scores between 406 and 634 is 68 %.

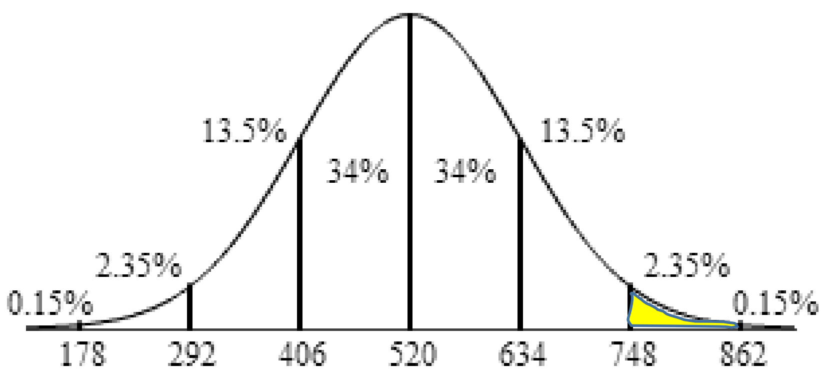
(b) What percentage of standardized test scores is less than 406 or greater than 634?

Use the results from part (a) to find the percentage of standardized test scores less than 406 or greater than 634.

Since the percentage of standardized test scores between 406 and 634 is 68%, the percentage of standardized test scores less than 406 or greater than 634 is $100\% - 68\% =$ 32 %.

(Type an integer or a decimal.)

(c) What percentage of standardized test scores is greater than 748?



2.35% + 0.15% = 2.5%

- Approximately 68% of the data will lie within 1 standard deviation of the mean.
- Approximately 95% of the data will lie within 2 standard deviations of the mean.
- Approximately 99.7% of the data will lie within 3 standard deviations of the mean.

17.

The following data represent the weights (in grams) of a random sample of 50 candies.

0.84	0.85	0.88	0.85	0.82	0.88	0.98	0.88	0.95	0.86
0.84	0.88	0.76	0.86	0.85	0.83	0.72	0.83	0.76	0.84
0.92	0.77	0.76	0.93	0.81	0.93	0.91	0.84	0.81	0.74
0.97	0.81	0.79	0.75	0.85	0.79	0.85	0.84	0.88	0.78
0.74	0.89	0.72	0.71	0.82	0.85	0.89	0.93	0.94	0.84

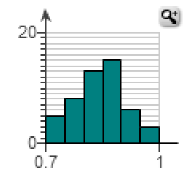
(a) Determine the sample standard deviation weight.

s = .07 gram

(Round to two decimal places as needed.)

s(sample) is Std. dev

(b) On the basis of the histogram on the right, comment on the appropriateness of using the empirical rule to make any general statements about the weights of the candies.



- ○ The histogram is not bell-shaped so the empirical rule cannot be used.
- ✔ The histogram is bell-shaped so the empirical rule can be used.

(c) Use the Empirical Rule to determine the percentage of candies with weights between 0.7 and 0.98 gram.

Hint: $\bar{x} = 0.84$.

mean = 0.84 and standard deviation is 0.07 from part a

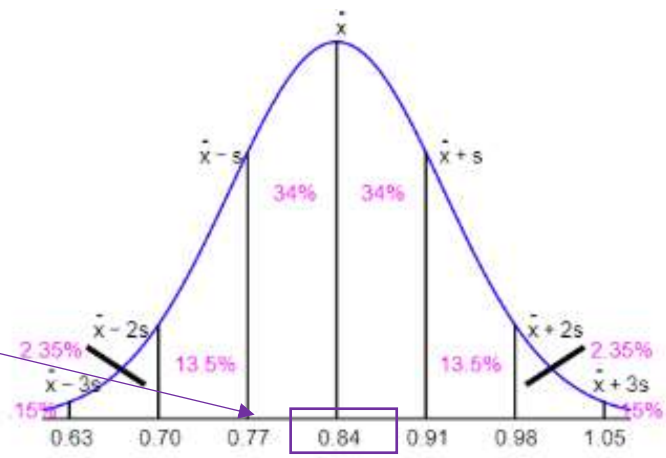
95 %

2 standard deviations away from the r

.84+.07+.07=.98

Mean is 0.84

S = 0.07



(d) Determine the actual percentage of candies that weigh between 0.7 and 0.98 gram, inclusive.

100 % **DATA – SORT to put data in order**

Count the #s in the chart between 0.7 and 0.98 and divide by total

(e) Use the Empirical Rule to determine the percentage of candies with weights more than 0.91 gram.

13.5% + 2.35% + .15%

16 %

(f) Determine the actual percentage of candies that weigh more than 0.91 gram.

16 % **DATA – SORT to put data in order**

Count the #s in the chart over 0.91and divide by total

18. True or False: Chebyshev's inequality applies to all distributions regardless of shape, but the empirical rule holds only for distributions that are bell shaped.

D. True, Chebyshev's inequality is less precise than the empirical rule, but will work for any distribution, while the empirical rule only works for bell-shaped distributions.

Chebyshev"s inequality – for a data set, regardless of the shape of distribution, at least $\left(1-\frac{1}{k^2}\right)$ **100%** of the observations will lie within k standard deviations of the mean.

19. At one point the average price of regular unleaded gasoline was $3.41 per gallon. Assume that the standard deviation price per gallon is $0.07 per gallon and use Chebyshev's inequality to answer the following.

(a) What percentage of gasoline stations had prices within 3 standard deviations of the mean?

(b) What percentage of gasoline stations had prices within 1.5 standard deviations of the mean? What are the gasoline prices that are within 1.5 standard deviations of the mean?

(c) What is the minimum percentage of gasoline stations that had prices between $3.13 and $3.69?

(a) At least 88.89 % of gasoline stations had prices within 3 standard deviations of the mean.
(Round to two decimal places as needed.) $\left(1-\frac{1}{3^2}\right) = 88.89\%$

(b) At least 55.56 % of gasoline stations had prices within 1.5 standard deviations of the mean.
(Round to two decimal places as needed.) $\left(1-\frac{1}{1.5^2}\right) = 55.56\%$

The gasoline prices that are within 1.5 standard deviations of the mean are $ 3.31 to $ 3.52
(Use ascending order.)

$\mu - 1.5\sigma = 3.41 - 1.5(.07) = 3.305$

$\mu + 1.5\sigma = 3.41 + 1.5(.07) = 3.515$

round to two decimals for $

(c) 93.75 % is the minimum percentage of gasoline stations that had prices between $3.13 and $3.69.
(Round to two decimal places as needed.)

Find the standard deviations the prices are away from the mean

[Mean – one of the scores] → *3.69 − 3.41 = 0.28 then divide by s=0.07*

$\frac{.28}{.07} = 4$ *then* $\left(1-\frac{1}{4^2}\right) = 93.75\%$

*USE $\left(1-\frac{1}{4^2}\right)$ WHEN IT SAYS MINIMUM PERCENTAGE

ANOTHER WAY ---- mean + s.d. until you get 3.69

3.41 + 0.07 = 3.48 + 0.07 = 3.55 + 0.07 = 3.62 + 0.07 = 3.69

4 standard deviations: $\left(1 - \frac{1}{4^2}\right) = 93.75\%$

20. What makes the range less desirable than the standard deviation as a measure of dispersion?

Choose the correct answer below.

- A. The range does not use all the observations. ✔
- B. The range is resistant to extreme values.
- C. The range is biased.
- D. The range describes how far, on average, each observation is from the mean.

The range is the largest value from the smallest value; therefore, it only uses two values of data.