Lesson 45
Data Distributions

NAME:

Start by navigating to the Online Lesson for instructions.

Objectives

- Describe a data distribution.
- Calculate standard deviation (for a population).
- Analyze normal distributions with the Empirical Rule.

Why?

In statistics, data is collected, organized, and analyzed to explore trends. Part of descriptive statistics is visually representing data sets after summarizing the data. This process is used in all types of industries and informs nearly every major decision companies make.

Warm Up

Define the terminology using your own words.

1) mean

2) median

3) range

4) mode

To continue, return to the Online Lesson.

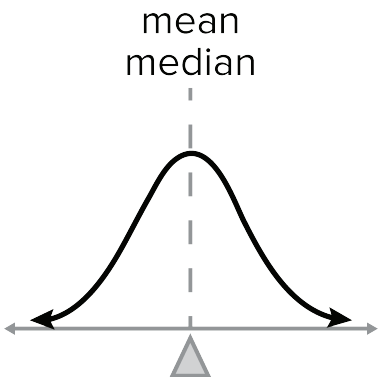
Explore

Describing Data Distributions

Fill in the notes as you watch the video in the Online Lesson.

- Measures of center help determine single values that can represent _____ _____.
- The measures of center are:
  - Mean (μ): population _____
  - Median: middle number when the data set is in _____
  - Mode: _____ number in a data set
- The closer the measures of center are to one another, the more _____ _____ they are, making the mean the representative measure of center.
- _____ should be used to calculate the mean, median, and mode of a data set, to allow for more time to _____.
- In statistics, the mean (as well as the variance and standard deviation) is rounded to _____ than the raw data.

Symmetric Distribution

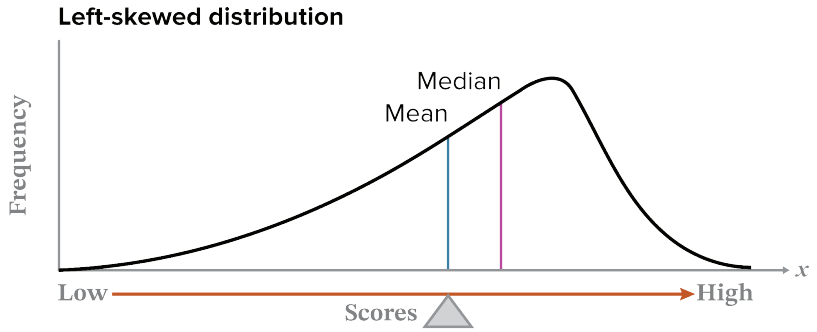


- The distribution is symmetric when _____.
- This is also called a _____.
- _____ data sets will only be approximately symmetric because no data set is perfectly normal.

Left-Skewed Distribution

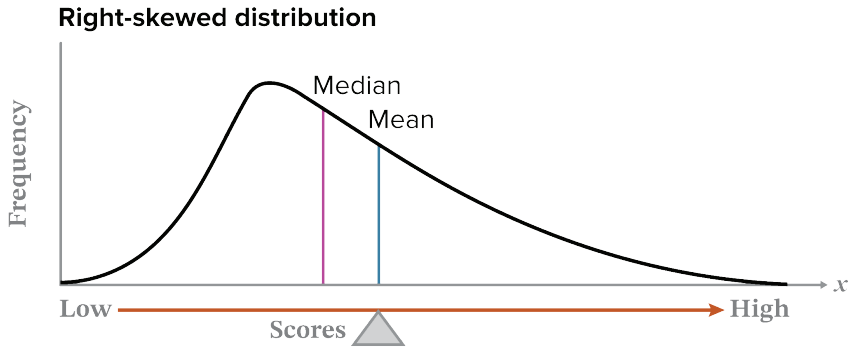
- Generally, the distribution is left-skewed when the _____.
- This distribution is also referred to as _____ skewed.

- The numbers in the data set are _____, toward negative values.



Right-Skewed Distribution

- Generally, the distribution is right-skewed when the _____.
- This distribution is also referred to as _____ skewed.
- The numbers in the data set are _____, toward positive values.



- There are some instances where the mean and median are equal, but the data is still skewed because one or more values of the data set are _____.
- No matter the distribution, it is important to look carefully at both:
  - _____
  - _____
- When distribution is skewed, the median is a better representation of center than the mean because it is _____.

Example 1

Complete the example as you watch the video in the Online Lesson.

Name and sketch the distribution.

A) $\{3, 4, 4, 5, 6, 7, 8, 15\}$

mean = 6.5

median = 5.5

mode: 4

B) $\{3, 4, 5, 6, 7, 9, 10, 10, 10, 12, 12, 12, 14\}$

mean = 8.8

median = 10

mode: 10, 12 (bimodal)

Use technology to calculate mean and standard deviation of population. For more information on how to enter data into a scientific calculator, see the More to Explore.

Example 2

Complete the example as you watch the video in the Online Lesson.

Calculate the measures of center. Then name the distribution.

$\{73, 76, 85, 85, 89, 90, 92, 95, 98\}$

Checkpoint: Describing Data Distributions

Calculate the measures of center. Name the distribution.

$\{10, 12, 15, 15, 18, 25, 40, 60\}$

To continue, return to the Online Lesson.

Calculating Standard Deviation

Fill in the notes as you watch the video in the Online Lesson.

- The measures of spread show the _____ contained in the data.
- The measures of spread are:
  - _____: The difference between the maximum and minimum value in the data set
  - _____: The measure of how data deviates from the mean

- Standard deviation (σ):
  - is a measure of spread based on _____.
  - of a population represents the average deviation, or _____ _____ the population mean.
  - is _____ because it is the distance from the mean.
  - is rounded to one decimal place further than the _____.

In this level, use the formula for standard deviation *of a population*. When using technology, be sure to use the "standard deviation of a population" option.

- The standard deviation of a population can be calculated using the formula:

$$\sigma = \sqrt{\frac{\sum \left(x_i - \mu\right)^2}{N}}$$

To find the standard deviation of a population:

1) Calculate the _____, μ.

2) Calculate the _____ between each data point, x_i, and the _____: $\left(x_i - \mu\right)^2$

3) Find the _____ of all values in step 2.

4) Calculate the _____, σ^2, by dividing the sum by the number of numbers in the data set N.

5) Take the _____ of the variance.

Example 3

Complete the example as you watch the video in the Online Lesson.

Piano teacher, Ms. Mbuy, tracked the number of students who registered for lessons over ten semesters. She wants to know how the class sizes vary from term to term. The class sizes are: $\{5, 6, 6, 7, 7, 8, 8, 9, 9, 10\}$.

A) Calculate the standard deviation without technology.

B) Over the same semesters, Mr. Jacono, the strings teacher, determined that his classes have a $\mu = 10.9$, $\sigma = 2.1$. Which teacher's classes have greater variance? Explain.

x	$(x-\mu)^2$
5	$(5-7.5)^2 =$
6	$(6-7.5)^2 =$
6	$(6-7.5)^2 =$
7	
7	
8	
8	
9	
9	
10	

See More to Explore for information on how to use technology to check your work.

Checkpoint: Calculating Standard Deviation

Mrs. Jensen's statistics unit includes six quizzes, each worth 55 points. Avery's scores are $\{47, 49, 50, 50, 51, 53\}$. Determine the standard deviation for Avery's quizzes.

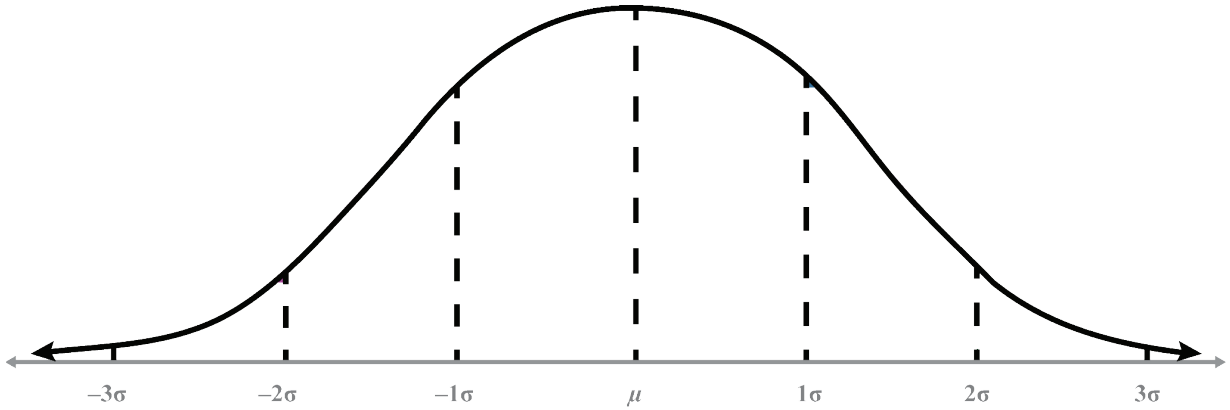
To continue, return to the Online Lesson.

The Empirical Rule (for Normal Distributions)

Fill in the notes as you watch the video in the Online Lesson.

- The Empirical Rule is also called the _____.
- For _____ distributed data only, the Empirical Rule says:
  - The area to _____ represents 50% of the data.
  - When given the mean (μ) and standard deviation (σ) of a population, you can approximate the area any single data will fall into by considering how far above/below it is from the mean.
  - The range of data can be written using either _____ notation or a _____ inequality.

Empirical Rule



Example 4

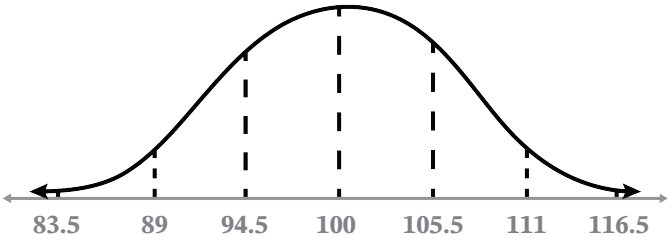
Complete the example as you watch the video in the Online Lesson.

A snack bag company fills bags using a machine. The production line manager's goal is for each filled bag to weigh 100 grams. Because the filling machine is very accurate, most bags are close to the target weight, with only a few bags being slightly under or slightly over the goal weight.

A) Name the mean and standard deviation.

B) How likely is it that a bag will weigh between 89 grams and 111 grams?

Snack Bag Weights



C) If 250 bags were selected randomly from the production lines, how many would weigh less than 111 grams?

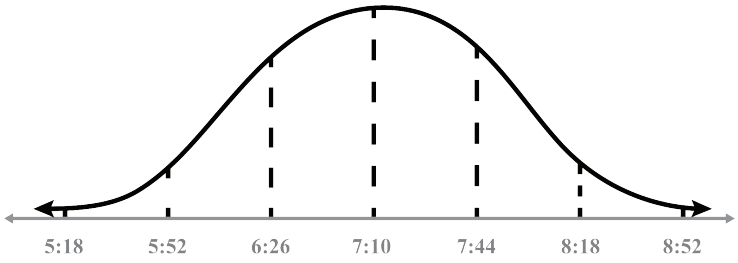
D) The manager notices that another filling machine has a standard deviation of 2.6, but has the same mean. Explain which machine should be used.

Example 5

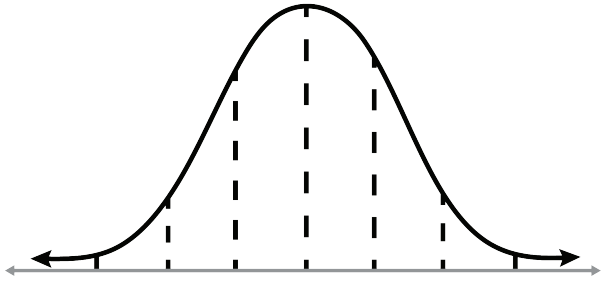
Complete the example as you watch the video in the Online Lesson.

A high school track coach compiled data over her coaching career, comparing pre-season and post-season mile times.

Pre-Season Mile Times



Post-Season Mile Times



A) Name the mean and standard deviation for the pre-season.

B) Complete the post-season normal distribution when $\mu = 6{:}25$, $\sigma = 0{:}20$.

C) Name the interval for 95% of the pre- and post-season mile times.

D) Explain how the data shows that mile times have improved.

Checkpoint: The Empirical Rule

The technology supervisor at Fair View High School determined that teachers receive $\mu = 11$, $\sigma = 2$ emails each day.

A) Sketch and label the normal distribution for emails received.

B) How likely is it that less than or equal to 13 emails are received daily?

C) What interval represents 95% of the population?

To continue, return to the Online Lesson.

Practice 1

Complete problems on a separate sheet of paper.

For problems 1–3, calculate the measures of center. Then sketch the distribution.

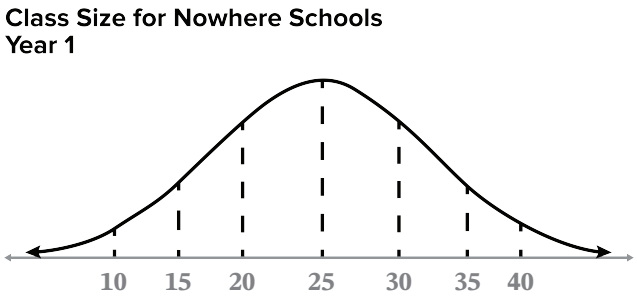
1) $\{95, 98, 92, 85, 99, 94, 91, 100, 70, 96\}$

2) $\{1, 2, 4, 20, 4, 25, 5\}$

3) $\{20, 80, 50, 60, 60, 65, 85\}$

For problems 4–10, use the given normal distribution.

4) Name the mean and standard deviation.



5) Determine the percentage of students with a class size on the interval $[15, 35]$.

6) What is the likelihood that a class contains fewer than 35 students?

7) Nowhere Schools want to include the range of students per class so parents know what to expect. What makes the most sense to include on their site (68%, 95%, or 99.7%)? Explain.

8) In Year 2, Nowhere Schools hired more teachers and added more classrooms. The average number of students per class decreased to 23 with a standard deviation of 3. Sketch the normal distribution for Year 2.

9) Name the interval where 68% of the class sizes occurred for both Year 1 and Year 2.

10) Explain how hiring more teachers and adding classrooms affected class size.

Calculate the standard deviation. Round to the nearest tenth.

11) $\{18, 40, 31, 28, 36, 22, 33, 30, 25, 15\}$

12) $\{8, 25, 12, 15, 20\}$

To continue, return to the Online Lesson.

Mastery Check

Show What You Know

The scientists measured the wingspans in centimeters of two groups of Monarch butterflies in different parts of the world.

Pennsylvania, USA
mean: 9.5, median: 9.5, mode: 9.5, 9.6

Michoacán, Mexico
mean: 9.1, median: 9.5, mode: 9.8

A) Explain the type of distribution for each group.

B) The average wingspan of a Pennsylvania Monarch butterfly is 9.5 cm with a standard deviation of 0.23. Sketch the normal curve. Name the intervals for the Empirical Rule.

C) Scientists measured the wingspan of eight Monarch butterflies found on a plant in Florida, USA. Calculate the mean and standard deviation.
$\{8.8, 9.2, 9.4, 9.5, 9.5, 9.7, 9.8, 10\}$

D) What is the interval for 99.7% of Florida Monarchs? Explain if they vary more than the Pennsylvania butterflies.

Say What You Know

In your own words, talk about what you have learned using the objectives for this lesson and your work on this page.

To continue, return to the Online Lesson.

Practice 2

Complete problems on a separate sheet of paper.

For problems 1–2, calculate the measures of center. Then sketch the distribution.

1) {40, 46, 48, 54, 57, 57, 60, 62, 73, 74}

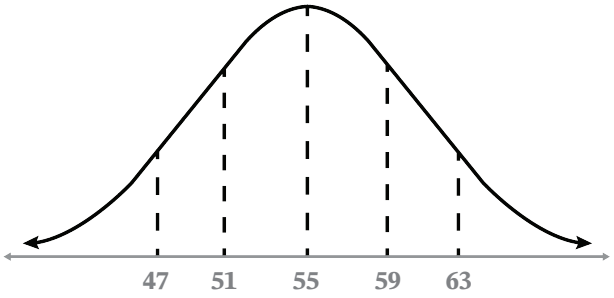
2) {6.5, 6.8, 9.0, 9.2, 9.4, 9.5, 9.6, 9.7, 9.8, 9.9, 10.0, 10.0}

3) Calculate the standard deviation.
{18, 20, 20, 21, 23}

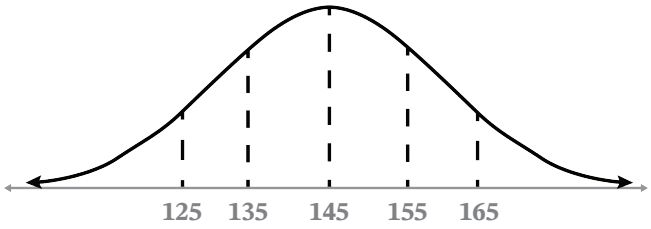
For problems 4–11, use the given normal distribution.

An athletic trainer measures athletes' heart rates at rest and immediately after a workout. Measurements are in beats per minute (bpm).

Resting heart rate (bpm)



Post-workout heart rate (bpm)



4) Name the mean and standard deviation.

5) What percentage of athletes' heart rates were higher than the mean?

6) If 125 athletes were monitored, approximately how many had a post-workout heart rate between 155 and 165 bpm?

7) If 125 athletes were monitored, how many have a resting heart rate less than 51 bpm?

8) If 150 athletes were monitored, how many fell between one standard deviation from the mean?

9) Write an inequality that represents 34% of the post-workout data above the mean (i.e., __ < x < __).

10) The coach created a new workout for the athletes. It increased their average heart rate by 5 bpm, but decreased the variance by 3 bpm. Sketch a graph for the new workout (Workout 2).

11) The team coach tells the athletic trainer that during a game, the heart rates of the athletes vary greatly. Which workout should the trainer recommend to the coach?

12) Calculate the standard deviation.
{12, 16, 18, 22, 24, 28}

To continue, return to the Online Lesson.

Targeted Review

Complete problems on a separate sheet of paper.

Multiply.

1) $(a-4b)^2$

2) $(a-4b)^3$

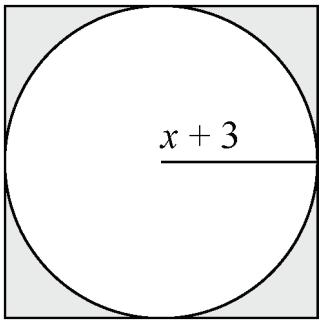
3) $(a-4b)^4$

4) In problems 1–3, what patterns do you notice in the exponent of the first term, a, and the second term, b?

5) Write $\log_a 4 + 2\log_a x - \log_a y$ as a single logarithm.

6) Convert to logarithmic form: $2^5 = 32$

Use the figure to complete problems 7–8.



7) Determine the geometric probability of landing on the circle with the given radius drawn inside (inscribed in) the square.

8) Write the geometric probability of landing on the shaded region in terms of π.

Multiple Choice

____ 9) Select the equations that best match the description.

The natural log, $g(x)$, is reflected over $y = x$, shifted right 3 units, and shifted up 2 units from $f(x)$

A) $f(x) = \ln(x-3) + 2 \qquad g(x) = e^x$

B) $f(x) = e^x \qquad g(x) = \ln(x-3) + 2$

C) $f(x) = e^x \qquad g(x) = \ln(x+3) + 2$

D) $f(x) = e^{(x-3)} + 2 \qquad g(x) = \ln x$

____ 10) What is the solution to the equation $7^x = 15$?

A) $x = 2.14$

B) $\log_5 x = 17$

C) $x = \dfrac{\log 7}{\log 15}$

D) $x = \dfrac{\log 15}{\log 7}$

_____ **11)** Evaluate: $e^{\ln 5 - \ln 8}$

A) $\ln -3$

B) $\frac{5}{8}$

C) e^3

D) $\frac{8}{5}$

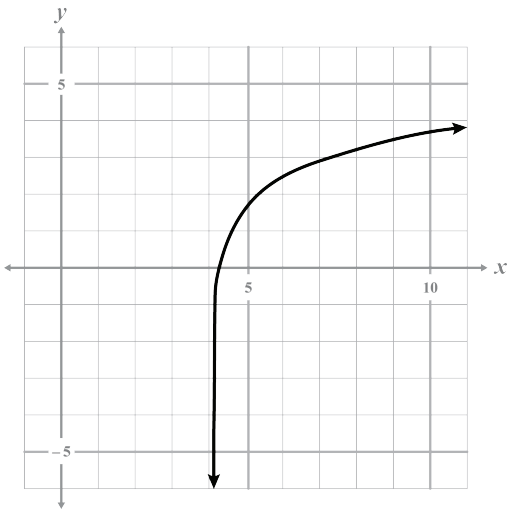
_____ **12)** Name the domain for $y = \ln(x - 4) + 2$.

A) $\{x \mid x \in \mathbb{R}, x > 4\}$

B) $\{x \mid x \in \mathbb{R}, x < 2\}$

C) $\{x \mid x \in \mathbb{R}, x < 4\}$

D) $\{x \mid x \in \mathbb{R}\}$



Problem	1	2	3	4	5	6	7	8	9	10	11	12
Origin	L3	L3	L3	–	L40	L38	L8	L8	L43	L41	L42	43

L = Lesson in this level, A1 = Algebra 1: Principles of Secondary Mathematics

To continue, return to the Online Lesson.