HONORS PRECALCULUS

CHAPTER 1

SEQUENCES & BIVARIATE DATA

Chapter 1: Sequences and Bivariate Data
Portfolio Items

1.1 Explain the meaning of a **mathematical sequence**. Include a discussion of subscript notation. Include examples. (Sec 1.1)

1.2 Give a complete discussion of **sigma notation**. Include examples. (Sec 1.1)

1.3 Explain the meaning of an **arithmetic sequence**. Include examples and appropriate formulae. (Sec 1.1)

1.4 Explain the meaning of a **geometric sequence**. Include a discussion of the **common ratio**. Include examples and appropriate formulae. (Sec 1.2)

1.5 Explain the meaning of an **infinite sequence**. Include examples along with appropriate formulae. (Sec 1.2).

1.6 Discuss **convergent** and **divergent** sequences. Give with examples and appropriate formulae. (Sec 1.2)

1.7 Explain the meaning of an **infinite series**. Include examples and appropriate formulae. (Sec 1.2)

1.8 Explain the meaning of a **recursive sequence**. Include a discussion of **explicit and recursive forms**. Include examples. (Sec 1.3)

1.9 Using an example of your choice, explain the steps involved in finding a **least squares regression line**. Discuss how the line would be used for making predictions in a real-life setting. Give the real-life interpretations of the slope and y-intercept from your example. (Sec 1.4)

1.10 Explain the meanings of **extrapolation** and **interpolation**. Use your example in Item 1.9 to give examples. (Sec 1.4)

1.11 Explain the meaning of **residual**. Include an example using your scenario in Item 1.9. (Sec 1.4)

1.12 Explain the meaning of **outlier**. (Sec 1.5)

1.13 Discuss what a **residual plot** is and how it is used. Include an example using Item 1.9. (Sec 1.5)

1.14 Discuss the meaning of **correlation coefficient** and how it is used. Include an example using your scenario in Item 1.9. (Sec 1.5)

UNIT 1 - SEQUENCES + BIVARIATE DATA

GETTING READY

1. Triangle ABC is an isosceles right triangle with $\angle C = 90°$ and $BC = \sqrt{5}$. Find the measure of
   - a. AC
   - b. AB
   - c. $\angle A$
2. In triangle DEF, $\angle F = 30°$, $\angle E = 60°$, and $DE = 2$. Find
   - a. EF
   - b. DF
3. Solve the system of equations:

   $3x + 7y = 6$
   $2x + 9y = 4$
4. Given the equation $2x + 3y = 6$
   - a. Tell the slope.
   - b. Graph the equation.
   - c. Tell the slope of the line parallel to $2x + 3y = 6$.
   - d. Graph the line that passes through (1, 3) and is parallel to $2x + 3y = 6$.
5. Write the equation of line that passes through (−2, 1) and (5, 4).
6. Simplify the following:
   - a. $\sqrt{3}(2 + \sqrt{3})$
   - b. $(3 + \sqrt{5})^2$
   - c. $\left(\frac{2}{\sqrt{3}}\right)^3$
7. High temperatures the first seven days of February in Miami are displayed in the table below.

Day	1	2	3	4	5	6	7
Temp. (°F)	66	63	71	73	75	75	76

   - a. Make a scatter plot of this data.
   - b. Estimate a line of best fit for the data.
8. Tell the next term in each of the following and explain the pattern the generates the sequence.
   - a. 5, 7, 9, 11,
   - b. $\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \ldots$
   - c. 8, 7, 5, 2, −2, . . .

Bopper's DVD Store has the latest in DVDs for rental or purchase. To attract more customers, Bopper's introduces the following promotion:

> **Bopper's DVD Store**
> Earn A Free DVD Rental!
> Earn 3 DVD Points With Your 1st DVD Rental
> Earn 2 DVD Points For Every Additional Rental
> Redeem 24 DVD Points For 1 FREE DVD Rental!

1. Complete the following table to indicate the total number of DVD points after each indicated DVD rental.

Bopper's DVD Store						
DVD Rentals n	1	2	3	4	5	6
Total DVD Points B_n						

2. The table shows that DVD points earned depends on the number of DVD rentals. Let B_n denote the total number of Bopper's DVD points earned after n rentals. What is the total number of DVD points for one, two, and three rentals?

$B_1 =$ $\qquad$ $B_2 =$ $\qquad$ $B_3 =$

The notation B_n is called **subscript notation**. This notation can be used to describe a **sequence**. In a sequence, B_n denotes the value of the n^{th} term in the sequence, as well as the number of DVD points earned after n rentals. The values $B_1, B_2, B_3, B_4, B_5, \ldots$ form a sequence of values. This sequence of values can be denoted as $\{B_n\}$.

> **READING MATH**
> B_1, B_2, and B_n are read as "B sub 1" and "B sub n", where "sub" represents *subscript*.
> In B_n, n is the term number or *index*.

3. Use $\{B_n\}$ to answer the following.

a. List the first eight terms of the sequence $\{B_n\}$.

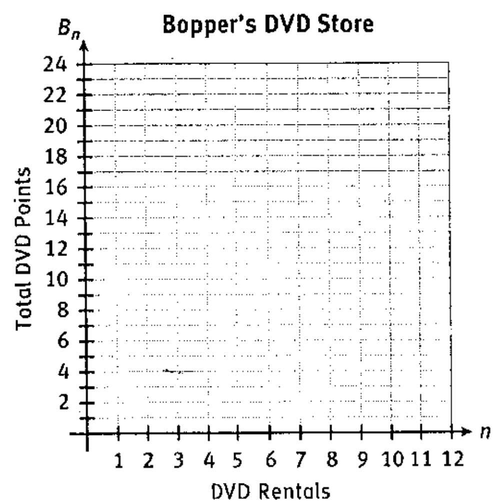
b. Explain the meaning of B_7.

c. Write an algebraic expression for the n^{th} term, B_n in terms of n, the number of DVD rentals at Bopper's DVD Store.

d. In the context of Bopper's DVD rentals, explain the meaning of n and the algebraic expression written in Part c.

4. On the coordinate grid below, the horizontal axis represents n, the number of DVD rentals, and the vertical axis represents B_n, the total number of DVD points earned from rentals at Bopper's.

a. Plot the sequence $\{B_n\}$ for $n = 1, 2, 3, \dots 8$



b. List as many properties of the graph of $\{B_n\}$ as possible.

5. How many rentals are needed to obtain a free DVD rental from Bopper's? Show the work that leads to your answer.

TRY THESE **A**

a. List the first six multiples of 4.

b. If $a_1 = 4$ in the above sequence, write an algebraic expression for a_n in terms of n, the term number.

c. Write an algebraic expression for the sequence $\{2, 5, 8, \ldots, 20\}$ in terms of n, the term number.

d. If $a_n = 20$ in the sequence in Part c, then solve for n.

A new video rental business, Fantastik Flicks, opens near Bopper's. The new store offers its own DVD point program to attract customers.

Fantastic Flicks DVD Store

Earn more DVD Points with each rental at Flicks!
Earn a free DVD rental faster at Flicks!

Earn 3 DVD Points with the first rental.
Earn 5 DVD Points with the second rental.
Earn 7 DVD Points with the third rental, and so on . . .
Free DVD Rental for every 100 Points earned at Flicks!

6. Complete the following table to indicate the number of DVD points earned with the n^{th} rental and the total accumulated number of DVD points after each Fantastik Flicks video rental.

Fantastik Flicks DVD Store						
DVD Rentals, n	1	2	3	4	5	6
DVD Points Earned with n^{th} Rental, P_n						
Total DVD Points After n Rentals, F_n						

7. The table in Item 6 indicates that there are two sequences associated with Fantastik Flicks DVD points, P_n and F_n.

a. At Fantastik Flicks, the number of points earned increases with each rental of a DVD. Write an algebraic expression for P_n, the number of points earned with the n^{th} rental, in terms of n.

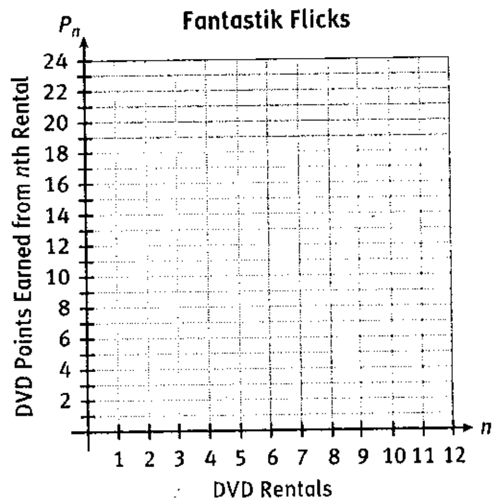
b. Explain how the values in the third row of the table in Item 6 are obtained from the values in the second row.

c. Find the value of the sum $P_1 + P_2 + P_3 + P_4 + P_5 + P_6 + P_7$.

d. What is the value of F_8? Show your work.

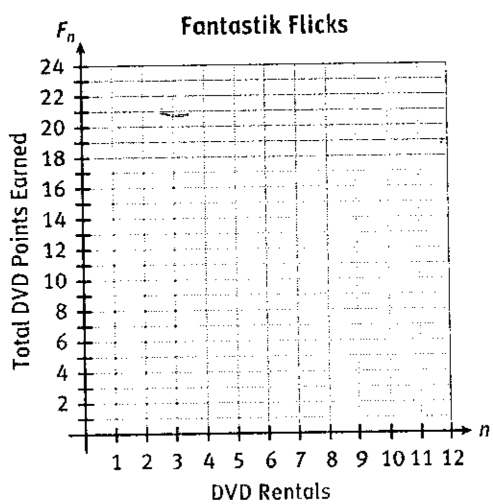
e. Write an equation that expresses F_{n+1} in terms of F_n and P_{n+1}.

8. On the coordinate grid below, the horizontal axis represents n, the number of DVD rentals at Fantastik Flicks, and the vertical axis represents P_n, the number of points earned per DVD rental. Use the grid to plot the sequence P_n for $n = 1, 2, ..., 8$.



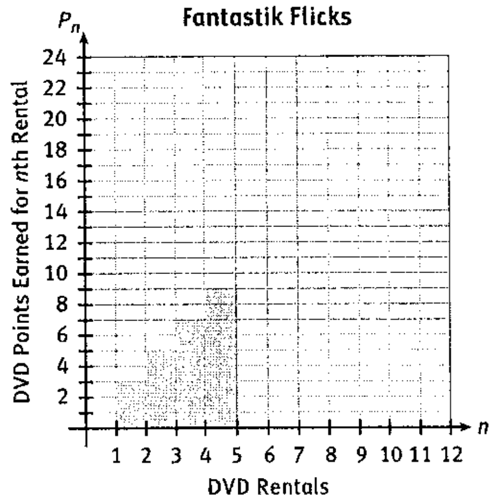
9. On the coordinate grid below, the vertical axis represents F_n, the total number of accumulated points earned after n rentals.

a. Plot the sequence F_n for $n = 1, 2, 3$, and 4.



b. Does the graph of F_n appear to be linear? Explain your answer.

10. The graph below shows the first four terms of the sequence $\{P_n\}$ as a shaded area.



a. Find the area of this shaded region. Explain how this area relates to the *Flicks'* DVD rental point promotion.

b. Extend the plot of the sequence $\{P_n\}$ to $n = 5, 6, 7,$ and 8. Find the area of the region under each of these plotted values and the new total area for each.

c. Investigate the relationship of the area under the plotted points of P_n to the accumulated number of Flicks' DVD points earned after n rentals for $n = 5, 6, 7$ and 8. Does the pattern discovered in Part a also hold for $n = 5, 6, 7,$ and 8? Explain.

Fantastik Flicks DVD points accumulate according to the sum $P_1 + P_2 + \ldots + P_{n-1} + P_n$ for the first n DVD rentals. **Sigma notation** can be used to streamline the writing of such sums. Using sigma notation, the sum $P_1 + P_2 + \ldots + P_{n-1} + P_n$ is written as

$$\sum_{j=1}^{n} P_j = P_1 + P_2 + P_3 + \ldots + P_{n-1} + P_n.$$

* Σ IS THE UPPER CASE LETTER SIGMA. IN MATH IT MEANS SUM.
* IN THIS CASE THE LETTER j IS THE "INDEX OF SUMMATION." 1 IS THE LOWER LIMIT OF SUMMATION WHICH TELLS US WHERE TO START ADDING. n IS THE UPPER LIMIT OF SUMMATION WHICH TELLS US WHERE TO STOP ADDING.

EXAMPLES

(1) $\sum_{r=1}^{4} r$

(2) $\sum_{i=0}^{3} i$

(3) $\sum_{r=2}^{7} r$

(4) $\sum_{K=5}^{11} K$

SIGMA NOTATION CAN BECOME MORE COMPLICATED THAN EXPRESSIONS WITH ONLY ONE VARIABLE.

EXAMPLES

(5) $\sum_{r=1}^{4} r^2$

(6) $\sum_{k=2}^{4} k(k-1)$

(7) $\sum_{i=1}^{5} (i^3+2)$

TEAM WORK EVALUATE THE FOLLOWING:

(A) $\sum_{r=11}^{14} r^2$

(B) $\sum_{i=1}^{5} 3i$

(C) $\sum_{n=2}^{6} \frac{1}{2}n(n+3)$

(D) $\sum_{k=0}^{4} \frac{k^4}{2}$

SIGMA NOTATION CAN ALSO BE USED IN GENERAL CASES:

EXAMPLES

(8) $\sum_{r=1}^{n} r^4$

(9) $\sum_{k=5}^{n} \frac{1}{k}$

(10) $\sum_{w=0}^{n} \frac{w}{w+1}$

TEAM WORK EVALUATE THE FOLLOWING:

(A) $\sum_{r=3}^{n} r^2(r-1)$

(B) $\sum_{K=6}^{n} (2K+4)$

(C) $\sum_{K=1}^{n} \frac{2}{K+1}$

(D) $\sum_{K=1}^{n} (2^K-1)$

(E) $\sum_{r=1}^{n} (-1)^r r^2$

FURTHER EXAMPLES

(11) $\sum_{i=1}^{3} X_i$

(12) $\sum_{K=1}^{n} Y_K$

IF $X_1 = 2$, $X_2 = 5$, $X_3 = 1$, $X_4 = 3$, $Y_1 = 4$, $Y_2 = 1$, $Y_3 = 6$, $Y_4 = 2$,
EVALUATE THE FOLLOWING:

(A) $\sum_{i=1}^{4} X_i$

(B) $\sum_{i=1}^{3} X_i^2$

(C) $\sum_{K=1}^{4} X_K Y_K$

(D) $\sum_{i=1}^{3} Y_{i+1}$

Rewrite each series as a sum.

1) $\sum_{m=1}^{5} (4m^2 + 4)$

2) $\sum_{k=1}^{5} (30 - k^2)$

3) $\sum_{n=1}^{5} n$

4) $\sum_{m=1}^{6} (50 - m)$

5) $\sum_{a=1}^{6} (3a^2 - 2)$

6) $\sum_{m=1}^{5} (100 - m)$

7) $\sum_{m=1}^{4} (5m^2 + 4)$

8) $\sum_{a=4}^{9} (20 - a^2)$

9) $\sum_{m=1}^{6} \frac{m^2 + 1}{m}$

10) $\sum_{n=4}^{9} (100 - n)$

11) $\sum_{m=0}^{5} m(m + 2)$

12) $\sum_{k=0}^{4} (100 - k)$

Evaluate each series.

13) $\sum_{n=1}^{7} (40 - n^2)$

14) $\sum_{k=1}^{5} 3k$

15) $\sum_{a=1}^{7} (500 - a)$

16) $\sum_{k=1}^{7} (30 - k)$

17) $\sum_{a=0}^{5} a$

18) $\sum_{k=0}^{4} 2k$

19) $\sum_{k=1}^{6} k^2$

20) $\sum_{m=1}^{5} 3m$

Rewrite each series using sigma notation.

21) $1 + 2 + 3 + 4$

22) $3 + 9 + 27 + 81 + 243$

23) $3 + 9 + 27 + 81$

24) $1 + 4 + 9 + 16 + 25$

25) $4 + 8 + 12 + 16$

26) $\frac{1}{2} + \frac{2}{3} + \frac{3}{4} + \frac{4}{5} + \frac{5}{6}$

27) $1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6}$

28) $\frac{1}{2} + \frac{2}{3} + \frac{3}{4} + \frac{4}{5} + \frac{5}{6} + \frac{6}{7}$

Critical thinking questions:

29) Are these equal? Why or why not?

$\sum_{x=1}^{50} \frac{1}{x}$ and $\sum_{x=21}^{70} \frac{1}{x - 20}$

30) Rewrite the following so that it starts at $x = 0$

$\sum_{x=7}^{10} x(x + 1)$

1.12

12. In Item 10, the number of *Flicks'* DVD points that accumulate with successive rentals can be modeled by the area that accumulates under the graph of the successive terms of P_n. Use the figures shown below to answer the following.

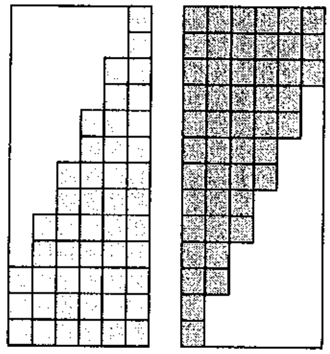


Figure 1 Figure 2

a. If Figure 1 represents a given number of DVDs rented from Fantastik Flicks and the total number of DVD points from these rentals, how many rentals are represented by this figure and how many DVD points have been accumulated from these rentals?

Number of rentals: ______ Number of accumulated points: ______

b. How does Figure 2 relate to Figure 1? Explain.

c. Figure 3 has a width that is equal to a given number of video DVD rentals from Fantastik Flicks. How is the height of Figure 3 related to the first and last terms of the sum $\sum_{j=1}^{6} P_j$?

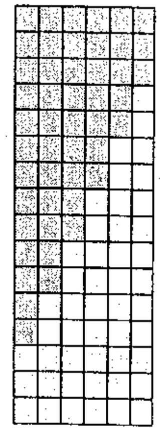


Figure 3

d. Find the area of the rectangle in terms of 6, P_1, and P_6.

e. Find the area of each of the two regions in Figure 3 in terms of 6, P_1, and P_6.

13. Use the results in Item 12 to answer the following.

a. Write an expression for the sum $\sum_{j=1}^{6} P_j$ in terms of 6, P_1, and P_6.

b. Express the sum $\sum_{j=1}^{n} P_j$ as a formula in terms of n, P_1, and P_n.

c. Complete the following table to verify that the formula in Part b is true for other values of n besides $n = 6$.

DVD Rentals, n	Points Earned with n^{th} Rental, P_n	Total Points after n Rentals, F_n	Formula from Part b
1			
2			
3			
4			
5			
6			

14. How many Fantastik Flicks DVD rentals are needed to obtain a free DVD rental? Show your work.

TRY THESE C

Bopper's is designing a new DVD display. 15 DVDs will be placed on the top row of the display, and the number of DVDs in each successive row will increase by 4 from the top row to the bottom row of the display.

a. Let r_n denote the number of DVDs in the n^{th} row of the display. Write an expression for r_n in terms of n, the term number.

b. How many DVDs will there be in the 10^{th} row from the top of the display?

c. Let T_n denote the total number of DVDs in the display for n rows. Write a formula for T_n in terms of n.

d. How many DVDs will be on display if there are 10 rows?

The Fantastik Flicks' DVD point program provides an example of an arithmetic sequence $\{P_n\}$. An **arithmetic sequence** has the general algebraic form $a_n = a_1 + (n - 1)d$, where a_1 is the first term and d is the constant difference between consecutive terms.

15. Use the table in Item 6 and examine the values for P_n and F_n.

a. Explain why $\{P_n\}$ is an arithmetic sequence.

b. Explain why $\{F_n\}$ is a sequence but *not* an arithmetic sequence.

The sum of all the terms of an arithmetic sequence forms an **arithmetic series.** The sequence $\{F_n\}$ is called a **sequence of partial sums** of an arithmetic series, because each term of the sequence $\{F_n\}$ is a sum of the first n terms of an arithmetic sequence.

The following illustration shows the connection between an arithmetic sequence, an arithmetic series, and the partial sums.

Arithmetic Sequence	Arithmetic Series	Sequence of Partial Sums
1, 2, 3, 4, ...	$1 + 2 + 3 + 4 + \ldots$	$S_1 = 1$ $S_2 = 1 + 2$ $S_3 = 1 + 2 + 3$ $S_4 = 1 + 2 + 3 + 4$ ...

TRY THESE D

Given the sequence $\{20 - 4n\}$, where $n = 1, 2, 3, \ldots$

a. List the first five terms of this sequence.

b. Write the associated arithmetic series using the first five terms.

c. Write the sequence of the first five partial sums.

d. Given an arithmetic sequence, find a_1 if $a_{15} = 63$ and $d = 5$.

Graphical representations of P_n and the area under the graph of the terms P_n were used to develop the formula $F_n = \frac{1}{2}n(P_1 + P_n)$. The fact that this formula holds true for values of n greater than six could be verified by extending the table created in Item 13c for larger values of n. However, extending the table would not prove that the formula is true for *all* positive integers n unless the table went on forever.

1.16

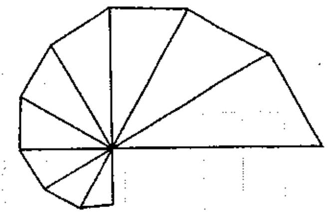
CHECK YOUR UNDERSTANDING

Write your answers on notebook paper or grid paper. Show your work.

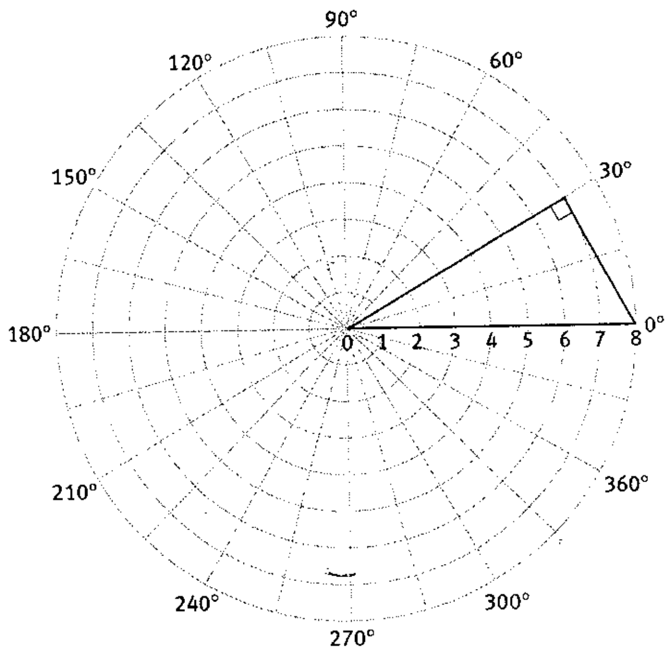
1. Given the sequence 1, 8, 15, . . .
   a. Write an expression for a_n in terms of n, the term number.
   b. Calculate a_{30}
   c. Given $a_n = 148$, solve for n.
2. Evaluate each of the following.
   a. $\sum_{j=1}^{5}(3j + 4)$
   b. $\sum_{k=0}^{5}\left(\frac{1}{2}\right)^k$
3. Rewrite the following series using sigma notation: 27 + 22 + 17 + 12 + 7 + 2
4. Find the sum of the first 20 terms of the arithmetic sequence 27, 22, 17, . . .
5. The MIU theater has 26 rows of seats. The first row has 18 seats, the second row has 20 seats, the third row has 22 seats, and this pattern continues. Write an expression for the total number of seats in the first n rows and use this expression to calculate the seating capacity of the theater.
6. Trip is the head cheerleader. Each time his team scores, he does push-ups, one push-up for each point his team has on the scoreboard. At the Homecoming game, Trip's team scored a field goal so he did 3 push-ups. Trip's team then scored 7 points so he did 10 push-ups. That day, Trip's team went on to score 7 points five more times, and his team won, 45–17. How many push-ups did Trip do at the Homecoming game?
7. How many terms of the arithmetic sequence −12, −3, 6, ... must be added to arrive at a sum of 363?
8. Which of the following sequences are arithmetic? For those sequences that are arithmetic, identify d and write an expression for a_n in terms of n, the term number.
   I. 100, 98, 96, 94, ...
   II. 64, 32, 16, 8, ...
   III. $2^1 + 1, 2^2 + 1, 2^3 + 1, 2^4 + 1, \ldots$
   IV. $\frac{3}{4}, \frac{13}{12}, \frac{17}{12}, \frac{7}{4}, \ldots$
9. For each of the arithmetic sequences in the previous problem,
   a. write the associated arithmetic series using the first 6 terms, and express that sum using sigma notation.
   b. write the sequence of the first 6 partial sums.

1.17

Shy Shelly Sellers sells sea shells in her *Fourth North Shore Store*. She is planning a sign for the storefront. She wants a large neon spiral reminiscent of the cross section of a shell. Each triangle in the design is a 30°-60°-90° triangle, and the length of the longest segment in the figure is 8 feet.



1. Recreate Shelly's design on the polar grid below. The first triangle is already drawn with the longest hypotenuse from the origin along the positive *x*-axis. Each successive hypotenuse is angled 30° counterclockwise from the previous one. As you calculate each hypotenuse length, record it in the table below the graph.



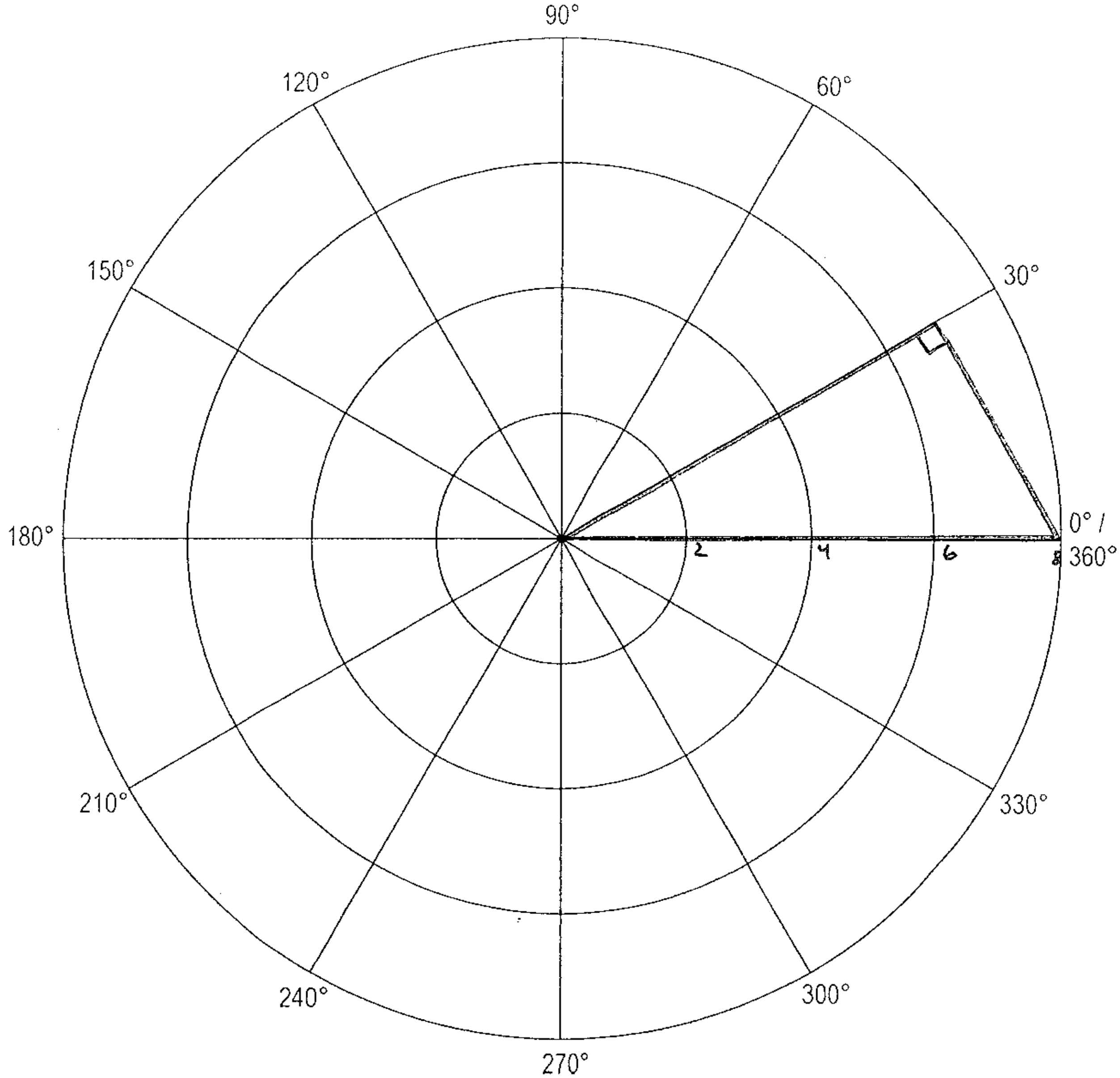
MATH TIP

The length of the hypotenuse in a 30°-60°-90° triangle is twice the length of the shorter leg and the length of the longer leg is $\sqrt{3}$ times the length of the shorter leg.

MATH TERMS

A **polar grid** is made up of concentric circles, the center of which is the **pole.** Coordinates for points on this grid are given in form (r, θ) where r represents the distance from the pole and θ represents an angle measured from the positive x-axis as shown.

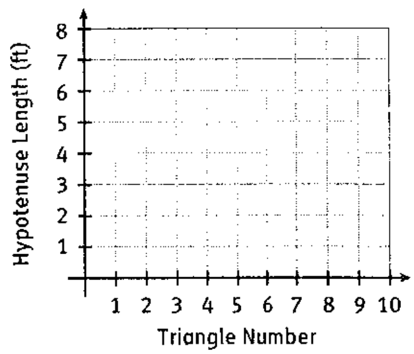
Triangle Number	1	2	3	4	5	6	7	8	9
Exact hypotenuse length (ft)	8								
Decimal approximation to nearest hundredth									



Triangle Number	1	2	3	4	5	6	7	8	9
Exact hypotenuse length	8								
Decimal approx. to nearest 100th	8.00								

2. a. List the exact values of the hypotenuse lengths from the table in Item 1 as a sequence with, $a_1 = 8$.

b. Plot the approximations on the grid.



3. Do the lengths of the hypotenuses in Shelly's design form an arithmetic sequence? Explain your reasoning.

A **geometric sequence** is a sequence for which the ratio, r, of a term to its preceding term is a constant. The constant r is known as the **common ratio**.

4. Complete each of the following ratios for a geometric sequence with common ratio, r.

$r = \frac{a_2}{\square}$ $\quad$ $r = \frac{\square}{a_5}$ $\quad$ $r = \frac{a_n}{\square}$

5. Explain why the sequence in Item 2 is a geometric sequence.

6. Each term in a geometric sequence can be written as a product of the first term and powers of the common ratio.

a. Complete the table for the terms in the sequence in Item 2.

n	a_n	a_n Written as a Product of a_1 and a Power of the Ratio
1	8	$8\left(\frac{\sqrt{3}}{2}\right)^0$
2		
3		
4		
5		
6		

b. Write an expression for a_n in terms of 8 and a power of $\frac{\sqrt{3}}{2}$.

7. A function of the form $f(x) = ab^x$ is an *exponential function*. Show why the expression that you wrote in Item 6 for a_n is an exponential function by identifying a and b.

8. Use the expression for a_n in Item 6 to verify the length of the 9th hypotenuse that you found in Item 1. (Do not use a_8.)

9. Suppose the spiral in Shelly's design is continued beyond the 9th triangle. Use the expression for a_n in Item 6 to calculate a_{15}.

10. Given another geometric sequence with $a_1 = \frac{1}{2}$ and $a_2 = \frac{-3}{2}$, calculate r and a_{10}.

11. Write an expression for a_n in terms of a_1, r, and n that can be used for any geometric sequence.

TRY THESE **A**

a. Identify which of the sequences below are geometric. If the sequence is a geometric sequence, identify a_1 and r, write an expression for a_n and calculate a_{15}.

I. $-5, -15, -45, -135, -405$

II. $-5, -1, 3, 7, 11$

III. $1, 4, 9, 16, 25$

IV. $6, -9, 13.5, -20.25, 30.375$

b. Determine the first term of a geometric sequence with $r = 1.4$ and $a_5 = 76.832$.

c. Calculate n for a geometric sequence with $a_1 = \frac{1}{32}$, $r = 2$, and $a_n = 4$.

d. Do the lengths of the short legs in the right triangles in Shelly's design form a geometric sequence? Make a prediction, calculate the actual lengths, and confirm or change your response.

12 a. For each partial sum in the table, write an expression for $\frac{\sqrt{3}}{2} S_n$ in expanded form. Use exact values.

n	S_n	$\frac{\sqrt{3}}{2} S_n$
1	8	
2	$8 + 4\sqrt{3}$	
3	$8 + 4\sqrt{3} + 6$	
4	$8 + 4\sqrt{3} + 6 + 3\sqrt{3}$	
5	$8 + 4\sqrt{3} + 6 + 3\sqrt{3} + \frac{9}{2}$	

MATH TERMS

The sum of the terms of a sequence is often used in applications. These sums are known as a **series**. The sum of the first n terms of a sequence is called the ***n*****th partial sum.**

b. Use your table from Part a to complete the table below.

- Complete the second column by writing an expression for $S_n - \frac{\sqrt{3}}{2} S_n$ as the difference of two terms.
- Complete the third column by factoring a_1 from each term in the second column.
- Complete the fourth column by expressing q in the third column as a power of r.

n	$S_n - \frac{\sqrt{3}}{2} S_n$	$a_1(1-q)$	$a_1(1-r^n)$
1	$8 - 4\sqrt{3}$	$8\left(1 - \frac{\sqrt{3}}{2}\right)$	$8\left(1 - \left(\frac{\sqrt{3}}{2}\right)^1\right)$
2			
3			
4			
5			

13. Use the results of the table in Item 12 to complete the following equation in terms of a_1, r, and n:

$S_n - rS_n =$

14. Solve the equation in Item 13 for S_n to find the formula for the sum of a finite geometric series.

15. Use the equation that you wrote in Item 14 to calculate the sum of the first 9 hypotenuses in Shelly's design.

MATH TERMS

Recall that a *series* is the sum of the terms in a sequence. A **geometric series** is the sum of the terms of a geometric sequence.

TRY THESE **B**

a. Find the sum of the first 8 terms of the geometric sequence whose first term is -2.5 and ratio is 2.

b. Evaluate $\sum_{k=0}^{9} 6(1.5)^k$

c. Express the sum in Part a using sigma notation.

d. Find the sum of the first 5 terms of the geometric sequence if $a_2 = 8$ and $a_3 = 10$. Show your work.

e. Find the sum of the areas of the first 10 right triangles in Shelly's design. Show your work.

16. As n increases what is happening to the length of each successive hypotenuse in the triangles in Shelly's design?

ACADEMIC VOCABULARY

converge
diverge

If the terms of an infinite sequence approach some number L as n increases without bound, the sequence is said to **converge**. If the sequence does not converge, it **diverges**.

CONNECT TO AP

The value to which a sequence converges is called the limit of the sequence. Later in this course and in AP Calculus, you will learn about and use limits in a variety of ways.

17. Does the sequence whose terms are the lengths of the hypotenuses in Shelly's design converge or diverge? If the sequence converges, what is the value that the terms appear to approach as n increases?

18. Use the infinite geometric sequences below.

I. $0.025, 0.25, 2.5, 25, 250, \ldots$
II. $100, 50, 25, 12.5, 6.25, \ldots$
III. $-4.2, 4.2, -4.2, 4.2, -4.2, \ldots$
IV. $\frac{1}{9}, -\frac{1}{3}, 1, -3, 9, \ldots$
V. $25, \frac{25}{\sqrt{5}}, \frac{25}{5}, \frac{25}{5\sqrt{5}}, \frac{25}{25}, \ldots$
VI. $32000, 320, 3.2, 0.032, 0.00032, \ldots$
VII. $1, \sqrt{2}, 2, 2\sqrt{2}, 4, \ldots$

a. Determine the common ratio for each sequence.

b. Which of the sequences converge and which diverge? For each sequence that converges, determine, if possible, the value to which the terms are approaching.

19. Create two infinite geometric sequences of your own, one that converges and one that diverges.

20. How can the ratio of a geometric sequence be used to determine whether a sequence converges or diverges, and if a sequence converges, what can be said about the value to which the terms are drawing near?

21. Calculate the first five partial sums for sequences I, II and III in Item 18 and the sequences you created in Item19. Which sequences of partial sums appear to converge and which appear to diverge?

22. If you calculated the 100^{th} partial sum for each of the sequences in Item 21, would any of your responses change? Explain your reasoning.

23. Suppose the spiral in Shelly's shell design is continued and allowed to overlap itself. As n increases, is there a length to which the hypotenuses are drawing near? Is there a value to which the sum of the lengths of the hypotenuses is drawing near?

24. What must be true for r, the common ratio of a geometric sequence, in order to have the sequence of the partial sums converge?

An **infinite sequence** is a sequence with an infinite number of terms. An **infinite series** is the sum of the terms of an infinite sequence.

ACADEMIC VOCABULARY

infinite sequence

25. For some series, an infinite number of terms can be added to get a finite sum! Recall the formula for the sum of the first n terms in a geometric sequence from Item 14: $S_n = \frac{a_1(1 - r^n)}{(1 - r)}$. Let $|r| < 1$. As n increases and gets very large, what happens to each of the following expressions?

a. r^n

b. $1 - r^n$

c. $\frac{a_1(1 - r^n)}{1 - r}$

26. Shelly wants to know the total length of material she will need for the display. Find the sum of the lengths of the hypotenuses in the triangles in her design if $a_1 = 8$ and the number of triangles in her design increases without bound.

CONNECT TO AP

If a sequence diverges, then the corresponding series also diverges. However, if a sequence converges, the corresponding series may or may not converge. Determining whether or not an infinite series converges or diverges is a topic you will study in AP Calculus.

TRY THESE C

a. For each infinite series in Item 18 that converges, find its sum.

b. Evaluate $\sum_{k=1}^{\infty} 10(0.8)^k$

c. The repeating decimal 0.363636... can be written as an infinite geometric series: $0.36(.01)^0 + 0.36(.01)^1 + 0.36(.01)^2 + \ldots$ Express the repeating decimal as a fraction by finding the sum of the corresponding infinite series.

CHECK YOUR UNDERSTANDING

Write your answers on notebook paper or grid paper. Show your work.

1. Determine which sequences below are geometric. For each geometric sequence, calculate r, find an expression for a_n and calculate a_{12}.
   I. 100, 98, 96, 94, ...
   II. $2^0, 2^1, 2^2, 2^3, \ldots$
   III. $3, 3\sqrt{6}, 18, 18\sqrt{6}, \ldots$
   IV. $2^0 + 1, 2^1 + 2, 2^2 + 3, 2^3 + 4, \ldots$
2. Find a_1 for a geometric sequence if $r = 5$ and $a_6 = 1000$.
3. Each month, the balance in Freda's bank account is 1.003 times as large as the previous month due to interest. If Freda does not deposit or withdraw any money from this account and she begins with \$2500, find the balance in this account at the beginning of the 36th month.
4. Complete the following analogy and explain your response. Arithmetic sequences are to linear functions as geometric sequences are to ____________.
5. Find the sum of the first 10 terms in each of the geometric sequences in Problems 1 and 2.
6. As a reward for inventing chess, Ja'qubi asked the Shah of Persia for 1 grain of wheat for the first of the 64 chessboard squares, 2 grains for the second, 4 grains for the third, 8 grains for the fourth and so on for all 64 squares. Calculate the number of wagons needed to transport the wheat if there are 20 million grains of wheat per ton and each wagon can carry 5 tons of grain.
7. Given the geometric sequence with $a_1 = 0.56$ and $r = 10$. Write the series representing the sum of the first 6 terms of the sequence and express this sum using sigma notation.
8. Evaluate $\sum_{k=0}^{4} 9\left(\frac{-1}{3}\right)^k$.
9. Calculate the sum for each of the following infinite geometric series that converge.
   I. $4 + 2 + 1 + \frac{1}{2} + \ldots$
   II. $4^2 + 2^2 + 1^2 + \left(\frac{1}{2}\right)^2 + \ldots$
   III. $\frac{-16}{18} + \frac{-8}{27} + \frac{-4}{9} + \frac{-2}{3} + \ldots$
   IV. $\frac{\sqrt{10}}{4} + \frac{10}{16} + \frac{10\sqrt{10}}{64} + \frac{100}{256} + \ldots$
   V. $\frac{8}{25} + \frac{4\sqrt{5}}{25} + \frac{2}{5} + \frac{\sqrt{5}}{5} + \ldots$
10. Express the repeating decimal 0.757575... as an infinite series and write it as a fraction.
11. **MATHEMATICAL REASONING** Not all infinite geometric series can be calculated. What must be true about an infinite geometric series if that series can be calculated?

SECTION 1.3 - RECURSIVE RELATIONSHIPS

* **EXAMPLE:** GIVEN THE SEQUENCE

$100, 50, 25, 12.5, \ldots$

WRITE AN EXPRESSION FOR a_n, THE nth TERM OF THE SEQUENCE.

THIS FORM IS CALLED AN "EXPLICIT EXPRESSION." FOR ANY GIVEN TERM OF THE SEQUENCE, THE EXPLICIT EXPRESSION CAN BE USED BY MERELY SUBSTITUTING THE TERM NUMBER, n, IN ORDER TO DETERMINE THE nth TERM.

* WHAT IS THE VALUE OF THE 10th TERM?

SEQUENCES CAN ALSO BE EXPRESSED "RECURSIVELY." A RECURSIVE EXPRESSION BASICALLY GIVES THE VALUE OF A TERM IN THE SEQUENCE, a_n, BASED ON THE VALUE OF THE PREVIOUS TERM, a_{n-1}. IT SHOULD BE NOTED THAT WHEN USING THIS FORM, THE FIRST TERM (a_1) MUST BE GIVEN OR WE WILL NOT KNOW WHERE TO BEGIN.

EXAMPLE: FOR THE SAMPLE SEQUENCE ABOVE,

(A) WRITE THE FIRST TERM: $a_1 =$

(B) WHAT IS BEING DONE TO THE FIRST TERM TO OBTAIN THE SECOND TERM?

(C) USE THIS TO WRITE AN EXPRESSION FOR THE nth TERM, a_n, BASED UPON THE PREVIOUS TERM, a_{n-1}.

$a_n =$

∴ THE RECURSIVE FORM OF THIS SEQUENCE IS

$a_1 =$

$a_n =$

TEAM WORK – FOR EACH OF THE FOLLOWING SEQUENCES, WRITE THE EXPLICIT AND RECURSIVE FORMS.

① $1, 2, 4, 8, \ldots$

② $2, 2\sqrt{3}, 6, 6\sqrt{3}, 18, \ldots$

③ $3, 6, 12, 24, \ldots$

④ $25, 22, 19, 16, \ldots$

GEOMETRIC SEQUENCES	ARITHMETIC SEQUENCES

APPLICATION - A LAKE INITIALLY CONTAINS 5200 FISH. EACH YEAR THE POPULATION DECLINES 30% DUE TO FISHING AND OTHER CAUSES, AND THE LAKE IS RESTOCKED WITH 400 FISH.

(A) WRITE A RECURSIVE EXPRESSION FOR THE NUMBER OF FISH, a_n, AT THE BEGINNING OF THE nth YEAR.

(B) HOW MANY FISH ARE THERE AT THE BEGINNING OF THE 5th YEAR?

(C) WHAT HAPPENS TO THE FISH POPULATION IN THE LAKE OVER TIME?

PRACTICE

① GIVE THE FIRST 5 TERMS FOR EACH SEQUENCE.

(A) $a_1 = 1$
$a_n = a_{n-1} + 4$

(B) $a_1 = 10$
$a_n = 3a_{n-1}$

(C) $a_1 = 2$
$a_n = (a_{n-1})^2 + 2$

(D) $a_1 = 5$
$a_n = n^2 - a_{n-1}$

(E) $a_1 = 48$
$a_n = \frac{1}{2}a_{n-1} + 2$

(F) $a_1 = 1$
$a_2 = 3$
$a_n = a_{n-1} \cdot a_{n-2}$

(2) WRITE AN EXPLICIT AND A RECURSIVE EXPRESSION FOR EACH SEQUENCE.

(A) $a_1 = 2$
$r = 10$

(B) $a_1 = 3$
$d = 10$

(C) $a_1 = -1$
$d = -\frac{3}{2}$

(D) $a_1 = \frac{1}{2}$
$r = 4$

(3) WRITE A RECURSIVE EXPRESSION FOR EACH SEQUENCE. THEY MAY BE GEOMETRIC, ARITHMETIC, OR NEITHER.

(A) 1, 7, 13, 19, . . .

(B) 66, 33, 16.5, 8.25, . . .

(C) 41, 32, 23, 14, . . .

(D) 6, $6\sqrt{2}$, 12, $12\sqrt{2}$, . . .

(E) 2, 5, 10, 50, 500, . . .

(4) SUPPOSE A TREE FARM INITIALLY HAS 9000 TREES. EACH YEAR 10% OF THE TREES ARE HARVESTED AND 800 SEEDLINGS ARE PLANTED.

(A) WRITE A RECURSIVE EXPRESSION FOR THE NUMBER OF TREES AT THE BEGINNING OF THE nth YEAR.

(B) HOW MANY TREES REMAIN AT THE START OF THE 4th YEAR?

(C) WHAT HAPPENS TO THE NUMBER OF TREES AFTER AN EXTENDED PERIOD OF TIME?

(5) A PERSON TAKES 20 mg OF A PRESCRIBED DRUG EVERY 4 HOURS. 30% OF THE DRUG IS REMOVED FROM THE BLOODSTREAM EVERY 4 HOURS.

(A) WRITE A RECURSIVE EXPRESSION FOR THE AMOUNT OF THE DRUG IN THE BLOODSTREAM AFTER n DOSES.

(B) WHAT VALUE DOES THE DRUG LEVEL IN THE PERSON'S BODY APPROACH AFTER AN EXTENDED PERIOD OF TIME? THIS VALUE IS CALLED THE MAINTENANCE LEVEL.

(C) SUPPOSE THE FIRST DOSE IS DOUBLED (TO 40 mg), BUT THE NORMAL DOSAGE IS TAKEN THEREAFTER. DOES THE MAINTENANCE LEVEL CHANGE?

Track-and-field events have a long history. When people first developed the ability to time events accurately, the record times of running events became part of the sport's tradition and added a new dimension to competition. Runners were not only interested in winning, but also in breaking an existing record. As the record time for the mile run dropped, an excitement was created as runners aspired to break a barrier once thought to be unreachable, the 4-minute mile.

The table below shows a list of the mile run records from 1911 to 1945. By 1945, the record for the mile run was 4:01.4 minutes. People in track and field around the world became fascinated by the prospect of the mile being run in four minutes or less. Who would break the 4-minute barrier? When would the record-breaking time occur? A sportswriter in 1945 would have quite a scoop if he or she could predict when the 4-minute mark would be passed.

Table I
Evolution of the Record for the Mile Run

Year	Time (seconds)	Time (minutes)	Runner (country)
1911	255.4	4:15.4	John Paul Jones (United States)
1913	254.6	4:14.6	John Paul Jones (United States)
1915	252.6	4:12.6	Norman Taber (United States)
1923	250.4	4:10.4	Paavo Nurmi (Finland)
1931	249.2	4:09.2	Jules Ladoumegue (France)
1933	247.6	4:07.6	Jack Lovelock (New Zealand)
1934	246.8	4:06.8	Glenn Cunningham (United States)
1937	246.4	4:06.4	Sidney Wooderson (Great Britain)
1942	246.2	4:06.2	Gunder Haegg (Sweden)
1942	246.2	4:06.2	Arne Anderson (Sweden)
1942	244.6	4:04.6	Gunder Haegg (Sweden)
1943	242.6	4:02.6	Arne Anderson (Sweden)
1944	241.6	4:01.6	Arne Anderson (Sweden)
1945	241.4	4:01.4	Gunder Haegg (Sweden)

1. Use the mile run records in the previous table to plot the data points on an entire sheet of coordinate graph paper.
   - Label the horizontal axis "Years Since 1900." and label the vertical axis, "Time (seconds)."
   - Narrow the range to [220, 270].Choose scales that will show the scatter plot on as much of the paper as you can.

2. Use the scatter plot from Item 1 for each of the following.
   - Locate a line that is a good model for the data points that you plotted.
   - Compare your line in Part a with the lines chosen by the rest of your group. Choose the line that your group believes best models the data. Provide a rationale for your choice.
   - Identify two points on the line chosen by your group and draw that line on your scatter plot.
   - Write the coordinates of the two points that you identified on your line.

3. The points chosen in Item 2 can be used to find an algebraic representation for the line.
   - **a.** Write the equation of this linear model.
   - **b.** What does the independent variable in your linear model denote? Include units of measure in your answer.
   - **c.** What does the dependent variable in your linear model denote? Include units of measure in your answer.
   - **d.** What is the slope of your linear model? Explain what information the slope gives about the mile run record data.
   - **e.** What is the y-intercept of your linear model? Explain what information the y-intercept gives about the mile run record data.

4. Imagine that you are a sportswriter in 1945. Use your linear model to predict the year that the 4-minute mile would be run. Show your work.

5. Compare your group's linear model for the mile run scatter plot data with linear models obtained by other groups in the class.
   a. Describe any differences among the linear models in terms of comparing the slope and the y-intercept.

   b. Explain why some of these differences are expected to occur.

To make a future prediction outside the data set, called **extrapolation**, a criterion for determining a best linear model can only be made based upon the available data. *Residuals* become important in determining a best linear model. A **residual** is the difference between an actual value and a predicted value.

For example, if the point (5, 8) is in a data set modeled by the function $y = 2x$, the predicted value of y is $y = 2(5) = 10$, so the residual is: actual y − predicted $y = 8 - 10 = -2$.

For the Mile Run Record data, the residual for a particular value of the independent variable equals the actual time for that value minus the predicted time for that value.

Residuals should be found for all ordered pairs in the data set.

ACADEMIC VOCABULARY

extrapolation
residual

6. Suppose a sportswriter uses the linear model $y = -0.35t + 259$ for the scatter plot of the mile run data. The table below is based on the sportswriter's model.

a. Use a calculator to verify the entries below and calculate the missing entries to complete this table.

Sportswriter's Linear Model: $y = -0.35t + 259$				
Years since 1900	**Actual Record Time (seconds)**	**Predicted Record Time (seconds)**	**Residual (Actual – Predicted) (seconds)**	**Squared Residual (seconds)**
11	255.4	255.15	0.25	0.0625
13	254.6	254.45	0.15	0.0225
15	252.6	253.75	−1.15	1.3225
23	250.4	250.95		
31	249.2	248.15		
33	247.6	247.45		
34	246.8			
37	246.4			
42	246.2			
42	246.2			
42	244.6			
43				
44				
45				
		Sum:	−1.15	19.6025

b. Some of the residuals in the table are positive and some are negative. Explain what the sign on each residual indicates about the linear model's fit to the actual scatter plot data.

c. Suppose a residual has a value of zero. What does this tell you about the relationship of the predicted value versus the actual value? Explain.

d. What reasons might there be to support squaring the residuals, as is done in the last column of the table?

You now have two possible criteria for judging a best linear model for a set of data.

- The line that minimizes the sum of the residuals.
- The line that minimizes the sum of the squared residuals.

7. In order to judge your group's linear model found in Item 3, complete the table below.

Your Group's Linear Model: $y =$				
Years since 1900	**Actual Record Time (seconds)**	**Predicted Record Time (seconds)**	**Residual (Actual – Predicted) (seconds)**	**Residual Squared (seconds)**
11	255.4			
13	254.6			
15	252.6			
23	250.4			
31	249.2			
33	247.6			
34	246.8			
37	246.4			
42	246.2			
42	246.2			
42	244.6			
43	242.6			
44	241.6			
45	241.4			
		Sum:		

CONNECT TO AP

An understanding of outliers, residuals, correlation coefficients, and written communication skills is an essential part of AP Statistics.

8. Compare the sum of the residuals and the sum of the squared residuals for linear models in the class.

a. Of the linear models constructed for the scatter plot data, including the sportswriter's model in Item 6, which minimizes the sum of the residuals? Write that linear model.

b. Of the linear models constructed for the scatter plot data, including the sportswriter's model in Item 6, which minimizes the sum of the squared residuals? Write that linear model.

c. Based on the results in Parts a and b, which of these two criteria, the sum of the residuals or the sum of the squared residuals, should be used for measuring the best linear model? Explain your reasoning.

It is no coincidence that the sportswriter's linear model in Item 6 does a better job of minimizing the sum of the squared residuals than any of the linear models in class. The sportswriter's linear model is the unique linear model that minimizes the sum of the squared residuals. This special linear model is known as the **line of least squares**, or the **regression line**.

9. Your linear model was found by guessing and by group consensus. Some calculators provide an option for creating a linear model that is based on the criterion of minimizing the sum of the squared residuals for the scatter plot data.

a. Use a calculator to find the equation for the regression line for the data in the table at the beginning of this activity and record it below to verify that it is the same as the sportswriter's linear model.

1911–1945 regression line: ______________________

b. What is the slope of the regression line?

c. What does the regression line suggest about the rate at which the world record in the mile run is changing based on the 1911–1945 data?

d. Using a model to estimate an outcome within a data set is called **interpolation**. What does the linear regression model estimate the mile run record to have been had the record been broken in 1927?

e. The 4-minute mile was a mythical barrier in the 1940s. What year does the regression line predict that the 4-minute mile would be run for the first time? Show the work that justifies your response.

10. Discuss, first in your group, and then prepare a report individually, concerning the extrapolation of the mile record for years outside those for which data is given. To start your group thinking, use your calculator's regression line model to predict what the record for the mile run would have been when Julius Caesar was assassinated (44 BC) and what the record would be in the year 3000. Use these extrapolations in your statement.

Table II presents additional data for the world record times in the mile run.

TABLE II
Evolution of the Record for the Mile Run

Year	Time (seconds)	Time (minutes)	Runner (Country)
1954	239.4	3:59.4	Roger Bannister (Great Britain)
1954	238.0	3:58.0	John Landy (Australia)
1957	237.2	3:57.2	Derek Ibbotson (Great Britain)
1958	234.5	3:54.5	Herb Elliot (Australia)
1962	234.4	3:54.4	Peter Snell (New Zealand)
1964	234.1	3:54.1	Peter Snell (New Zealand)
1965	233.6	3:53.6	Michael Jazy (France)
1966	231.3	3:51.3	Jim Ryun (United States)
1967	231.1	3:51.1	Jim Ryun (United States)
1975	231.0	3:51.0	Filbert Bayi (Tanzania)
1975	229.4	3:49.4	John Walker (New Zealand)
1979	229.0	3:49.0	Sebastian Coe (Great Britain)
1980	228.8	3:48.8	Steve Ovett (Great Britain)
1981	228.53	3:48.53	Sebastian Coe (Great Britain)
1981	228.4	3:48.4	Steve Ovett (Great Britain)
1981	227.33	3:47.33	Sebastian Coe (Great Britain)
1985	226.32	3:46.32	Steve Cram (Great Britain)
1993	224.39	3:44.39	Noureddine Morceli (Algeria)
1999	223.13	3:43.13	Hicham El Guerrouj (Morocco)

TECHNOLOGY TIP

Add the data for 1954–1999 to L1 and L2 for the 1911–1999 regression line. Then delete the data from 1911–1945 for the 1954–1999 regression line.

11. Record the equation of the regression line from Item 9 in the appropriate space below. Use a graphing calculator and the data in Tables I and II to find equations for the regression lines for the years 1911–1999 and for the years 1954–1999.

1911–1945 regression line: ____________________

1911–1999 regression line: ____________________

1954–1999 regression line: ____________________

12. For each of the following years, identify the appropriate regression equation in Item 11 and predict the mile run record in that year. Explain your reasoning.

a. 1970

b. 1950

c. 1920

13. Use what you have learned from this activity and any data that you feel necessary to construct new linear models to make a prediction of the mile run record for the years stated below. Give the years of the data used to create the equation of the regression line and explain why you chose those years.

a. 1901

b. 2020

14. Use a linear model to predict when the next milestone for the mile-run record, namely the 3.5-minute mile, will be surpassed. Give the equation of the linear model you chose and justify your choice.

CHECK YOUR UNDERSTANDING

Write your answers on notebook paper or grid paper. Show your work.

1. Marla and Tomas needed to model the data shown. They each chose a different linear equation. Marla chose $y = \frac{3}{4}x + 10$, and Tomas chose $y = \frac{2}{3}x + 15$. Calculate the sum of the squared residuals for each model to determine which is a better fit for this data.

x	y
10	20
20	14
30	22
40	40
40	71
60	55
65	50
80	70
85	90
90	79
100	85
120	80

2. Marcille is filling her swimming pool. Every 30 minutes, she measures the depth of the water.

Time (hours)	Depth (inches)
0	0
0.5	7
1	16
1.5	25
2	36
2.5	43
3	48
3.5	50.5
4	53
4.5	55

   a. Use the data and a graphing calculator to find the equation of the regression line that can be used to determine the depth of the water in the pool. Use this equation to estimate the time it will take to fill the pool to 6 feet.

   b. Due to the odd shape of her pool, Marcille decides that two models should be used, one for the first 2.5 hours and one 2.5 hours or more. Find the equations of the two regression lines.

   c. Use the appropriate model to estimate the depth of the water after 1.25 hours.

   d. Use the appropriate model to estimate the time that it takes to fill the pool to 6 feet where Marcille is taking her readings.

3. **MATHEMATICAL REFLECTION** How can you tell from a scatter plot of data that multiple regression lines would provide a better fit for the data than one regression line? How can you tell from the data in a table that it would be best modeled by using more than one regression line?

SECTION 1.5 - LINEAR RELATIONSHIPS

Mr. Mokher loves data. Last week he asked all of his students to write the number of minutes they spent exercising that week so that the data could be analyzed for any association between time exercised and fitness percentage score.

Here is the data collected from his physical education students.

Student	Exercise (min)	Fitness Score	
1	25	55	
2	0	57	
3	80	93	
4	75	85	
5	65	80	
6	10	95	
7	45	65	
8	60	73	
9	55	75	
10	35	55	
11	40	62	
12	55	67	
13	65	75	
14	120	100	
15	65	83	
16	90	62	
17	70	88	
18	80	90	
19	85	92	
20	90	95	

1. Make a scatter plot on graph paper comparing each student's fitness score to the amount of time exercised in minutes.

2. How did you determine the scale for the axes?

MATH TERMS

Points where the value of the y-variable is significantly greater or less than expected are called **outliers**.

3. Describe the association between the amount of time exercised and the fitness percentage.

4. Some students' data appear to be **outliers**, points that do not fit the overall linear pattern. Identify these students and give an interpretation for each student's score.

ACADEMIC VOCABULARY

The **correlation coefficient** gives an indication of the strength of the linear relationship that exists between two numerical variables. Note that it is dimensionless; there are no units attached to it.

One way to describe how closely two numerical variables, like time exercised and fitness percentage scores, are associated is through the **correlation coefficient** (r). A correlation coefficient is a number between -1 and $+1$ that quantifies without units the linearity of two data sets.

5. For the following sets of data, use technology to create a scatter plot. Then find the correlation coefficient. If you use a graphing calculator, you may need to find the regression line to do this.

a.

x	y
1	2
2	4
3	6
4	8
5	10

b.

x	y
1	2
2	2
3	4
4	6
5	11

c.

x	y
1	3
2	2
2	4
3	1
3	5

d.

x	y
1	10
2	7
3	7
5	5
9	4

6. Based on your results, what do the correlation coefficients $+1$, -1, and 0 indicate?

7. Find the correlation coefficient for Mr. Mokher's class's data. What does this number tell you about the linear relationship between the minutes exercised and the student's fitness percentage?

8. Write a linear regression model for this data set.

READING MATH

The appropriate notation for a linear regression model is

$$\hat{y} = ax + b.$$

$\hat{y}$ is read as "y-hat," which represents the value of y generated by the model.

9. Graph the linear model you found on the scatter plot. Explain the procedure you used.

10. Explain the meaning of the slope and y-intercept of the model in terms of the contextual situation.

11. Remove the outliers you identified in Mr. Mokher's class data and recalculate the correlation coefficient.

   a. What is the new correlation coefficient and what does it indicate about the data set without the outliers?

   b. Could you have predicted the effect that the omission of the outliers would have on the correlation coefficient? Explain.

12. How did removing the outliers affect the slope and y-intercept of the regression line? Could you have predicted this? Explain.

Recall from Activity 1.5 that a **residual** is the predicted value, $\hat{y}$, subtracted from the actual value, and that a residual may be positive or negative.

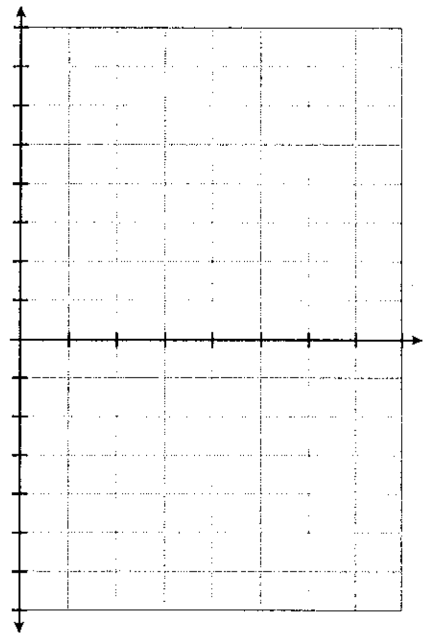
13. Calculate the residuals for each student and record them in the fourth column of the original table of data. Label the fourth column as Residuals.

14. What do you notice about the residuals for the outliers you identified? Is this an expected result? Explain.

15. Find the sum of the residuals and record it. What do you notice? Is this an unexpected result? Explain.

A **residual plot** provides a picture of how much each data point deviates from the regression line. In order to make a residual plot, the x-value of the data point is plotted along with the residual value for that data point. The line $y = 0$ should be plotted to use as a reference for the regression line.

16. Make a residual plot for the class data that Mr. Mokher collected.



In a residual plot, a random scattering of residuals above and below the x-axis indicates that a linear regression may be an appropriate model for the data.

17. How did you determine the scale for the axes of the residual plot?

18. What patterns do you notice in the residual plot? Do they support the idea that the regression model for this data should be linear?

The data below show the United States population from 1890 to 1990, according to the U.S. Census Bureau.

Year	Population
1890	62,979,766
1900	76,212,168
1910	92,228,496
1920	106,021,537
1930	123,202,624
1940	132,164,569
1950	151,325,798
1960	179,323,175
1970	203,302,031
1980	226,542,199
1990	248,709,873

19. Use a piece of graph paper to make a scatter plot of this data. Explain how you chose the scale for each axis of your graph.

20. Describe the overall association between the year and the United States population.

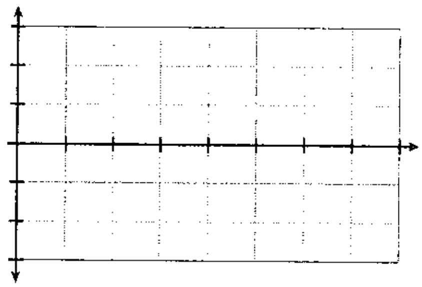
21. Find the correlation coefficient for the census data. Explain what this measure indicates.

22. Find the linear regression model for this data.

23. Explain the meaning of the slope and y-intercept of your regression line in the context of this data.

24. Draw the regression line on the scatter plot.

25. Create and label the residual plot for this data.



26. Write a summary of this linear regression model that uses the United States population data. Be sure to include:

- which points appear to be outliers, if any, and explain why you chose them and why they may have occurred
- whether the residuals exhibit a pattern which supports the use of a linear regression model or not
- whether or not the residual plot supports the value of the correlation coefficient that you found in Item 22.

CHECK YOUR UNDERSTANDING

Write your answers on notebook paper or grid paper. Show your work.

Mr. Mokher and his Pre-Calculus students went to the local elementary school to collect data. First, his students made a sample composed of a random selection from Grade 0 (Kindergarten) through Grade 6. They collected data on each child's grade level and height in inches.

Student	Grade Level	Height (in.)
1	0	36
2	0	38
3	0	38
4	1	39
5	2	40
6	2	42
7	3	45
8	4	47
9	5	50
10	5	55
11	5	66
12	6	60

1. Make a scatter plot of the grade versus height.
2. Describe the overall association between the grade level and the student's height in inches.
3. Give the correlation coefficient for the grade level and student's height in inches. What does this value indicate about the relationship between the grade level and student's height in inches?
4. Write the regression equation.
5. Interpret the slope and y-intercept in the context of this problem.
6. Graph the regression line on your scatter plot. Explain your process.
7. Identify any students who do not fit the pattern and provide a possible plausible explanation for this data point.
8. Remove the possible outlier(s) from the data set. Calculate the new correlation coefficient. Does the correlation coefficient indicate a stronger or weaker linear relationship? Explain.
9. How did removing the outliers affect the slope and y-intercept of the regression equation?
10. Create a residual plot of the original data.
11. Does the residual plot support the use of a linear regression model? Explain.
12. MATHEMATICAL PRACTICES What graphical and numerical results can be used to determine whether the use of a linear model is appropriate? Explain.

ADDITIONAL PROBLEMS

ACTIVITY 1.1

1. Given the sequence $-3\sqrt{2}, -\sqrt{2}, \sqrt{2}, \ldots$

a. Write an expression for a_n in terms of n, the term number.

b. Calculate a_{25}

c. Solve for n if $a_n = 19\sqrt{2}$

2. Evaluate.

a. $\sum_{k=1}^{6}(k^2 - 3)$

b. $\sum_{i=2}^{7} 12i$

3. Rewrite the following series using sigma notation: $\frac{-5}{2} + \frac{-3}{2} + \frac{-1}{2} + \frac{1}{2} + \frac{3}{2}$

4. The new MIU arena has 24 rows of seats in the upper level. The first row has 180 seats, the second row has 184 seats, the third row has 188 seats and this pattern continues. Calculate the seating capacity of the upper level.

5. How many terms of the arithmetic sequence $-5, -1, 3, \ldots$ must be added to reach the sum of 400?

6. Given the arithmetic sequence with $a_1 = 7$ and $d = -2.5$.

a. Write the first 5 terms of this sequence.

b. Write the associated arithmetic series using the first 5 terms and express this sum using sigma notation.

c. Write the sequence of the first 5 partial sums.

7. Verify the following formulas using mathematical induction.

a. $1 + 5 + 9 + \ldots + (4n - 3) = n(2n - 1)$

b. $\frac{1}{3} + \frac{1}{15} + \frac{1}{35} + \ldots + \frac{1}{4n^2 - 1} = \frac{n}{2n + 1}$

ACTIVITY 1.2

8. Given the geometric sequence with $a_1 = 90$ and $r = \frac{2}{3}$.

a. Find an explicit expression for a_n.

b. Calculate a_8.

9. Find a_1 for the geometric sequence with $a_4 = 54$ and $a_5 = 16.2$.

10. On the average, a certain car will depreciate 15% in value each year. Find the value of one of these cars five years after it is purchased if the purchase price is $22,000.

11. Evaluate $\sum_{j=1}^{5} 24\left(\frac{1}{2}\right)^j$

12. A bungee jumper drops 160 ft, bounces back up and drops 120 ft and then 90 ft and so on until he comes to rest. If each drop is three fourths the length of the previous drop, find the total distance that the bungee jumper falls.

13. Express the repeating decimal 0.454545. . . as a fraction.

ACTIVITY 1.3

14. Eddie opens a money market account with $4500. The account pays 3% interest, compounded monthly. Each month $248 is taken out for a car payment.

a. Write a recursive expression for a_n.

b. Write an explicit expression for a_n.

c. Calculate a_{12}

d. Explain the meaning of a_{12} in this problem.

e. Eddie does not want the balance in his account to go below $500. How many months can Eddie go before he needs to add money to his account?

15. For each of the sequences below, write a recursive expression and an explicit expression for a_n.

a. $400, $375, $350, . . .

b. $400, $400(0.95) + $30, $410(0.95) + $30, . . .

c. $400, $200, $100, . . .

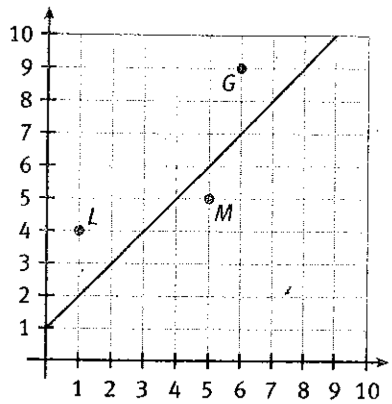
d. $400, $400(1.003) + $10, $411.20(1.003) + $10, . . .

e. $400, $480, $576, . . .

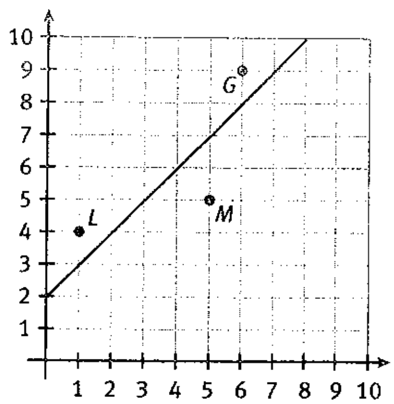
f. $400, $418, $436, . . .

16. Which is the correct median-median line for the 3 medians L, M, and G?

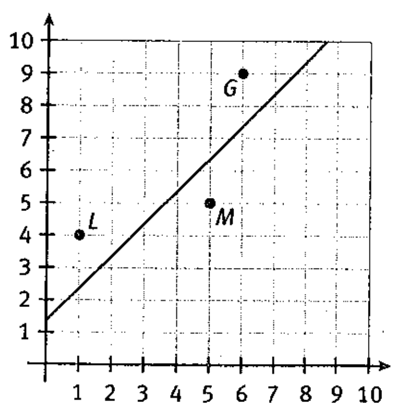
a.



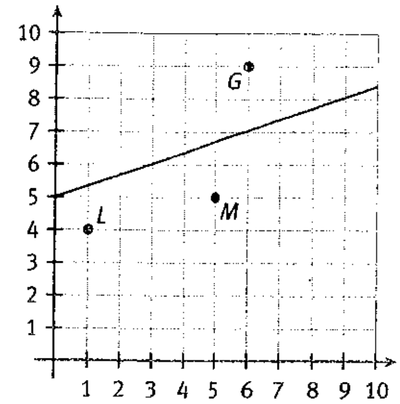
b.



c.



d.



17. Selma House is a local real estate agent, and based on recent sales in her neighborhood, she wants a linear model to use when pricing properties. The table below lists the square footage for each of the last 15 houses sold and the selling price for that house.

Square Feet	Selling Price	Square Feet	Selling Price
1000	$122,000	2200	$194,000
1200	$140,000	2400	$239,000
1400	$132,000	2500	$240,000
1500	$158,000	2750	$280,000
1600	$169,000	3000	$330,000
1850	$160,000	3200	$325,000
2000	$185,000	3500	$360,000
2150	$208,000		

a. Write the equation for the median-median line for Selma House's data.

b. Explain the meaning of the slope of the median-median line in the context of Selma's data.

c. Selma is listing a house with 1700 sq. ft. Use the median-median line to determine a likely selling price for this house.

d. Use the median-median line to estimate the square footage of a house that sells for $250,000.

ACTIVITY 1.5

18. Will Purchase, one of Selma's clients from 17, talks with Selma about the selling prices for the houses in her neighborhood. He believes an accurate linear model for these prices, given the square footage is $y = 90x + 23{,}000$ where x represents the square footage of the house and y represents the selling price.

Use the values from the table in **17** to calculate the sum of the squared residuals for Will's model.

19. Upon closer inspection of the list of the selling price for the houses in her neighborhood, Selma decides that there should be two different linear models: one for houses with less than 2400 square feet of space and those houses with 2400 square feet or more.

a. Use a graphing calculator to find equations for the two regression lines.

b. Use the appropriate model to estimate the selling price for a house with 1700 square feet of space.

c. One of Selma's clients is looking to spend between \$200,000 and \$260,000 on a house in her neighborhood. What size houses (in square feet) should Selma show this client?

ACTIVITY 1.6

Mr. Mokher and his Pre-Calculus students went to the local elementary school to collect data. First his students made a sample composed of a random selection from the Kindergarten (grade 0) through sixth (grade 6). They collected each child's grade level and weight in pounds.

Student	Grade Level	Weight (lb)
1	0	39
2	0	40
3	0	42
4	1	45
5	2	47
6	2	52
7	3	57
8	4	95
9	5	68
10	5	85
11	5	85
12	6	100

20. Make a scatter plot of the grade versus weight.

21. Describe the association between the grade level and the student weight in pounds.

22. Find the correlation coefficient for the grade level and student weight. What does this value indicate about the linear relationship between grade and weight?

23. Write the regression equation.

24. Interpret the slope and y-intercept in the context of this problem.

25. Graph the regression line on your scatter plot. Explain your process.

26. Identify any students who do not fit the pattern and provide a possible explanation.

27. Circle and remove the possible outlier(s) from the data set.

a. Give the new correlation coefficient.

b. Does the correlation coefficient indicate a stronger or weaker linear relationship? Explain.

28. How did removing the outlier(s) affect the slope and y-intercept of the regression equation?

29. Make a residual plot of the original data.

30. Does the residual plot support the use of a linear regression model? Explain.