

Analyzing Numerical Data: Estimating Large Numbers

I.A Student Activity Sheet 1: Estimating Crowds



Estimating the number of people in a large crowd (for example, watching a parade or attending/marching in a political rally) is quite challenging and often leads to controversies. One method sometimes used is to focus on a small section of the crowd, such as a rectangular area.

1. Make a square measuring 5 feet by 5 feet, and have your friends stand inside it as if they are watching a band at a small club. Count the number of your friends that comfortably fit in the rectangle and find the ratio of this number to the rectangle's area. Explain in your own words what this ratio means.

Analyzing Numerical Data: Estimating Large Numbers

I.A Student Activity Sheet 1: Estimating Crowds

4. **EXTENSION:** Additionally, a method of estimating the number of people participating in a march is to assign someone to a spot on the sidewalk with the job of counting the number of marchers who pass in 1 minute. Then multiply this number by the minutes it takes for all marchers to pass that spot.

Use this method to estimate the students who attend your school (or a certain building on your campus) on a given morning. Count the number of students entering the building in 1 minute and multiply this number by the minutes it takes the entire group to enter.

5. **REFLECTION:** What other methods could be used to estimate the size of a large crowd? Describe how and why the crowd size estimate supplied by event organizers might differ from estimates made by other groups.

Analyzing Numerical Data: Estimating Large Numbers

I.A Student Activity Sheet 2: Filling Your Classroom with Tennis Balls

1. Consider the situation of determining the number of tennis balls needed to fill your classroom. With a partner or small group, take 5 minutes and make an estimate of the number of tennis balls required to fill the room.

Do you think your estimate is too small or too large? Why?

2. Using the materials provided, make a third estimate. This time, find a lower bound and upper bound for the number of tennis balls you think it would take to fill your classroom. Draw sketches as needed.
3. **REFLECTION:** In determining your group's estimate, what mathematical model of a tennis ball did you use? What model of the classroom did you use? Did you make other simplifications or assumptions (for example, assuming there are no desks in the room)?

Analyzing Numerical Data: Estimating Large Numbers

I.A Student Activity Sheet 2: Filling Your Classroom with Tennis Balls

Fermi Questions

“How large is the crowd?” is an example of a *Fermi question*, which requires estimating physical quantities to produce an answer. Enrico Fermi was an Italian physicist known for his ability to take limited information and arrive at a reasonable numerical estimate. He approached a problem by setting a lower estimate—a value that the item in question was more than—and an upper estimate—a value that the item in question was smaller than. Fermi then worked through the problem using this approach and reached a numerical estimate within identified limits. Fermi worked at the University of Chicago, and his classic question was “How many piano tuners are there in the city of Chicago?”

Fermi questions rely on making logical assumptions about a situation to reach an estimate to an order of magnitude (usually a power of 10). That is, the estimate is expressed in hundreds, thousands, millions, and so on—whatever unit is appropriate for the question.

Another example of a Fermi question is:

Can every person in the world fit inside a cubic mile?

What assumptions do you need to make? What information do you need to find?

Analyzing Numerical Data: Estimating Large Numbers

I.A Student Activity Sheet 2: Filling Your Classroom with Tennis Balls

4. **EXTENSION:** Tackle one or more of the following Fermi questions:
- Which is more valuable: a pile of pennies equaling your weight or a stack of quarters equaling your height?
 - What is your estimate for the total amount of time students in your school spent during this past semester studying for final exams?
 - If the land area of Earth was divided equally for each person on the planet, about how much would you get?
 - How many songs are played on your favorite radio station in a given year?
 - What is your estimate for the number of airline flights in a year?
 - What is your estimate for the total number of hairs on your head?
 - What is your estimate for the number of square inches of pizza consumed by all students in your school during the past year?
 - How many gallons of gasoline do cars use each year in the United States?
 - How many hot dogs are eaten at major league baseball games during a season?
 - How many minutes will U.S. high school students spend texting on their cell phones during the next year?
 - How much does it cost to raise a child?

Analyzing Numerical Data: Estimating Large Numbers

I.A Student Activity Sheet 3: Not Enough Numbers

Telephone Numbers

During the past few years, there have been several statements of the form: “It is conceivable that we may run out of area codes and telephone numbers within the next 10 years.” The shortage of phone numbers is due to the proliferating use of cell phones, pagers, and fax machines. In the next few problems, you will estimate the number of possible telephone numbers in North America. The assignment of telephone numbers is coordinated by the North American Numbering Plan Administration (www.nanpa.com). (NANPA actually covers only the United States and its territories, Canada, and the Caribbean. Mexico is not included.)

A telephone number in the form *NYZ-ABC-XXXX* has three sections:

NYZ	ABC	XXXX
area code	exchange code	station code

Before 1995, all area codes had the form *NYZ*, where *N* was any digit from 2 to 9 (2-9), *Y* was 0 or 1, and *Z* was 1-9 if *Y* was 0 or *Z* was 2-9 if *Y* was 1. The restrictions on *N* saved 0 for *call operator* and 1 for long-distance calls. In addition, codes such as 800 and 911 were (and still are) used for special purposes.

The restriction that *Y* = 0 or 1 was removed in 1995 because all possible area codes had been assigned. Today *N* is 2-9, *Y* is 0-8, and *Z* is 0-9; the exception to these rules are codes of the form *37Z* and *96Z*, which are being reserved for future use. Area codes where *Y* = *Z* are called *easily recognizable codes* and are often assigned to special services such as 800 and 877.

1. How many area codes were possible before 1995?

2. According to the post-1995 rules, how many area codes are possible today?

Analyzing Numerical Data: Estimating Large Numbers

I.A Student Activity Sheet 3: Not Enough Numbers

3. The 7-digit numbers in a given area code have the form ***ABC-XXXX***, where ***X***, ***B***, and ***C*** can be any digit 0-9 and ***A*** is restricted to 2-9. There are two other restrictions:
- ***B*** and ***C*** cannot both equal 1 since these values are designated for other purposes such as ***911*** (emergency) and ***411*** (information), and
 - 555-0100 through 555-0199 are reserved for fictional uses such as in television shows or movies.

According to these conditions, how many 7-digit numbers are possible in a single area code?

4. Using your answers to the previous questions, determine how many 10-digit numbers are possible in North America.

Analyzing Numerical Data: Estimating Large Numbers

I.A Student Activity Sheet 3: Not Enough Numbers

5. **EXTENSION:** Some states are running out of license plate numbers. Delaware currently uses six-digit numbers in its license plate numbering system, although recent reports show that its numbers are approaching 1 million (Delaware uses retired numbers for new cars in many cases). The state of Washington recently stated that it needs to explore options to its system of three numerical digits followed by three letters because it is running out of numbers. New Jersey changed its system of three letters followed by two numerical digits and one letter to a new system of one letter, two numerical digits, and three letters. (The last number under the old system was ZZZ 99Z, followed by A10 AAA in the new system.)
- How could you determine how many plate numbers New Jersey was able to assign under the previous system? What assumptions did you make in your calculation?
 - How many additional license plate numbers can New Jersey assign under the new numbering system?
 - Why do you think the first license plate under the new system was not A00 AAA?

Analyzing Numerical Data: Estimating Large Numbers

I.A Student Activity Sheet 3: Not Enough Numbers

d. How do New Jersey's previous and new systems relate to the license plate numbering systems used in Delaware and Washington?

e. What do license plate numbers tell you about the population in the state?

6. **REFLECTION:** In solving real-world problems such as estimating a crowd size, the number of possible telephone numbers, or the answer to a Fermi question, why is it necessary to first make some reasonable assumptions about the situation?

Responses to this Reflection will show students' different approaches to the question.

Analyzing Numerical Data: Using Ratios

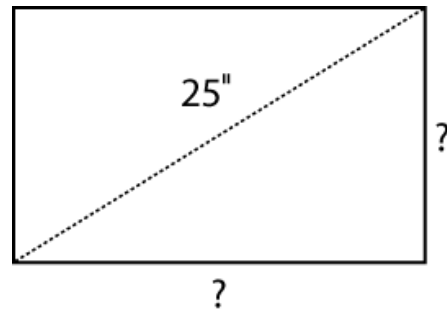
I.B Student Activity Sheet 4: Ratios in the Media

For a rectangular shape such as a display screen, the longer side is called the width (W) and the shorter side is the height (H). The aspect ratio is $W:H$ or W/H .

1. What is the approximate aspect ratio of the screen on your graphing calculator? Consider only the window, not the entire screen.

The size of a television is the length of the diagonal of its screen in inches. The aspect ratio of the screens of older televisions is 4:3, while the aspect ratio of newer wide-screen televisions is 16:9.

2. Find the width and height of an older 25-inch television whose screen has an aspect ratio of 4:3.



Find the area of this screen.

Student: _____ Class: _____ Date: _____

Analyzing Numerical Data: Using Ratios

I.B Student Activity Sheet 4: Ratios in the Media

3. Repeat this process to find the width and height of a newer 48-inch television whose screen has an aspect ratio of 16:9.

Determine the area of the screen of a newer 48-inch television whose screen has an aspect ratio of 16:9.

Analyzing Numerical Data: Using Ratios

I.B Student Activity Sheet 4: Ratios in the Media

When movies that were made in one aspect ratio are shown on televisions that have a different aspect ratio, black bars of equal width cover a portion of the screen. Portions of the screen are not needed to project images that were created with different **aspect ratios**.



Figure 1
4:3 screen displaying a 16:9 image



Figure 2
16:9 screen displaying a 4:3 image

4. Figure 1 shows a letterboxed image with an aspect ratio of 16:9 displayed on a screen with an aspect ratio of 4:3. What percent of the screen's area is occupied by the image? Justify your answer.

Some people do not like seeing the letterboxes when watching a 16:9 image on a 4:3 display, as shown in Figure 1. What would happen to the image if it filled the height of the TV?

Figure 2 shows a pillarboxed 4:3 image displayed on a 16:9 screen. What percent of the screen's area is occupied by the image? Justify your answer.

Analyzing Numerical Data: Using Ratios

I.B Student Activity Sheet 4: Ratios in the Media

People who own a wide-screen television can choose one of three views of a 4:3 image on their display.

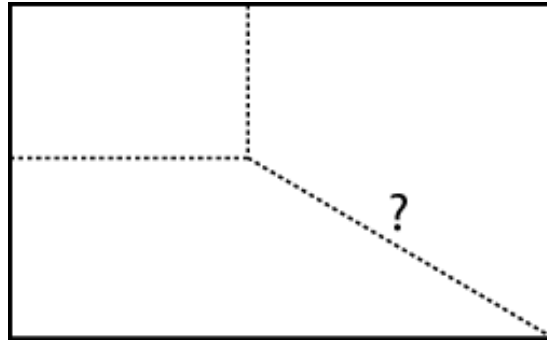
- The normal view shows the pillarboxes, as shown in Figure 2.
- Another option is to stretch the width of the image, keeping the height the same.
- A third option is to zoom in on the image, making the width of the image take the full width of the display.

What affect do these options have on the image?

Analyzing Numerical Data: Using Ratios

I.B Student Activity Sheet 4: Ratios in the Media

5. **EXTENSION:** When changing the size of a rectangular image using a computer-drawing program, how can you maintain the same aspect ratio? Investigate this scenario using a drawing program within a computer application.



Analyzing Numerical Data: Using Ratios

I.B Student Activity Sheet 4: Ratios in the Media

6. **EXTENSION:** Today most movies made for theaters are shot on film with an aspect ratio of 1.85:1 or 2.35:1. When a film with an aspect ratio of 1.85:1 is shown on a newer wide-screen 16:9 television, minor letterboxing often occurs, but the resulting black bars are usually not visible. Research commonly used aspect ratios and the history associated with this concept.

Why are all aspect ratios identified as something to 1? Or perhaps even recorded as a decimal value?

Express 16:9 to a value in ratio to 1.

A movie with an aspect ratio of 1.85:1 is shown as a letterboxed image on a newer 50-inch 16:9 television. Calculate the height of the image, the height of each barely visible black bar at the top and bottom of the screen, and the percent of the screen's area that is occupied by the image. Use a variety of representations to justify your response.

A movie with an aspect ratio of 2.35:1 is shown as a letterboxed image on an older 25-inch 4:3 television. Find the height of the image, the height of each black bar at the top and bottom of the screen, and the percent of the screen's area that is occupied by the image. Use a variety of representations to justify your response.

Based on your analysis of television aspect ratios, why do you think the consumer market has moved having wide-screen TV in the home?

Analyzing Numerical Data: Using Ratios
 I.B Student Activity Sheet 4: Ratios in the Media

7. **EXTENSION:** Using appropriate measuring devices, determine the aspect ratio of several objects.

Item	Length	Width	Aspect Ratio	
Sheet of paper	11 inches	8.5 inches	11:8.5	1.3
Index card				

A first-class piece of mail (letter or postcard) must have an aspect ratio that falls between 1.3 and 2.5, inclusive. What are some typical sizes of envelopes and postcards? Justify your reasoning.

Analyzing Numerical Data: Using Ratios

I.B Student Activity Sheet 4: Ratios in the Media

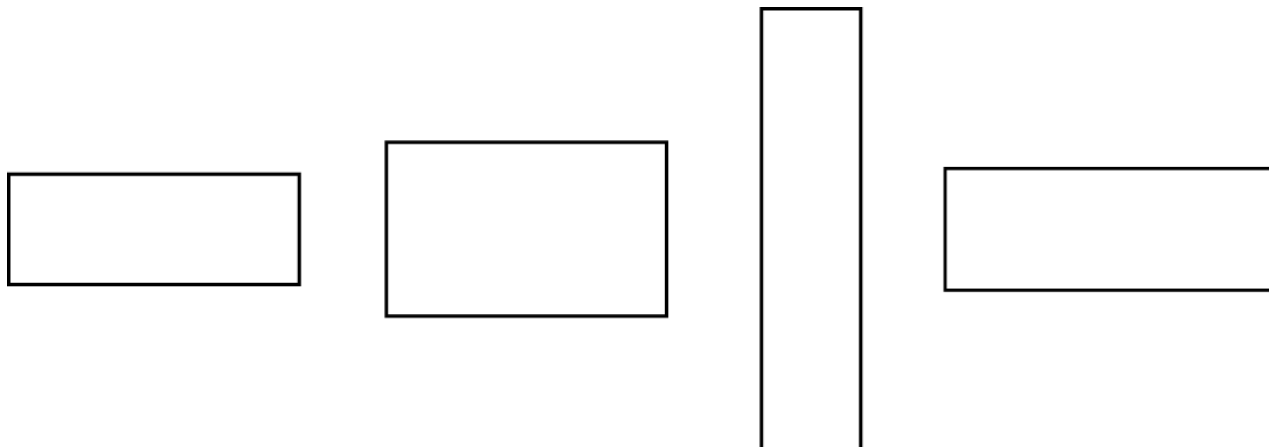
8. **EXTENSION:** The involvement of aspect ratios in taking and printing photographs is essentially the same as it is in showing movies on a television set. Common still photography aspect ratios are 4:3 (most point-and-shoot digital cameras) and 3:2 (35-mm film). Common U.S. print sizes (in inches) are 4 x 5, 4 x 6, 5 x 7, 8 x 10, and 11 x 14. Verify the following statement regarding image and print size using a variety of representations.

If your camera uses a 4:3 aspect ratio, but you want a 4 x 6 (3:2) print, many megapixels are wasted (11%).

Analyzing Numerical Data: Using Ratios

I.B Student Activity Sheet 4: Ratios in the Media

9. **EXTENSION:** Some figures have what is known as a *golden ratio*, which is an aspect ratio that is considered to be visually appealing. Rectangles that have a golden aspect ratio are often called *golden rectangles*. Which of the following rectangles is a golden rectangle?



The aspect ratio of the second rectangle from the left is the golden ratio. Conduct research on how the value of the golden ratio is found and why figures with this aspect ratio are considered to be visually appealing.

Analyzing Numerical Data: Using Ratios

I.B Student Activity Sheet 4: Ratios in the Media

10. **REFLECTION:** How does learning about the various uses of aspect ratio affect your buying habits of products with a visual screen? What types of questions might you ask a sales representative or product help screen online?

Analyzing Numerical Data: Using Ratios

I.B Student Activity Sheet 5: Changing Tires

You have just purchased a new vehicle equipped with factory-installed P245/70R16 tires. You think these tires look too small, so you replace them with P285/75R16 tires. How does this change in tire size affect the accuracy of speedometer and odometer readings? Specifically, your goal is to complete the following statements:

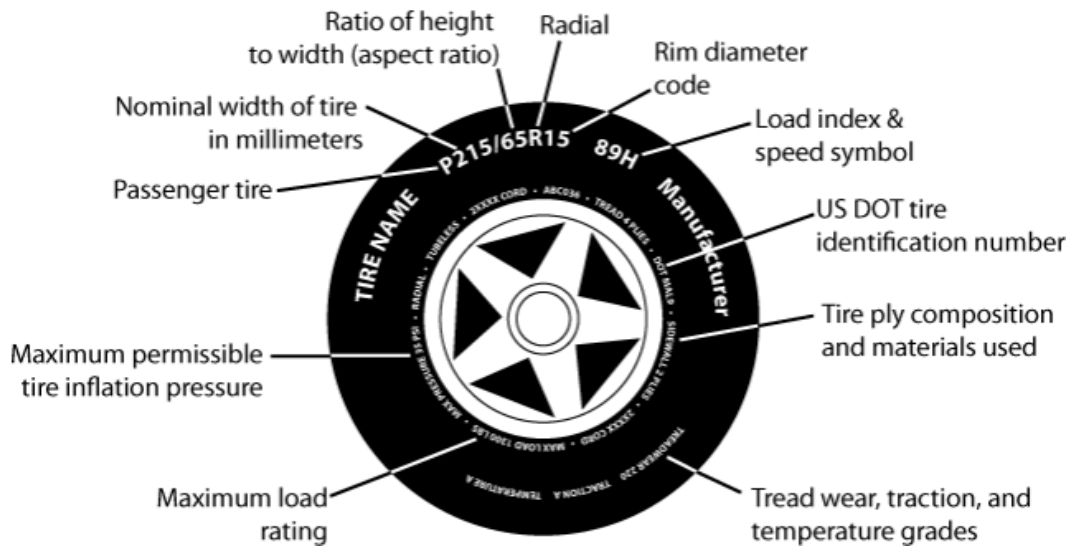
If your odometer reading is 20000, you have actually traveled _____ miles.

If your speedometer reading is 60, your actual speed is _____ miles per hour.



Analyzing Numerical Data: Using Ratios

I.B Student Activity Sheet 5: Changing Tires



The calibration of a vehicle’s speedometer and odometer is based on the circumference of the vehicle’s factory-installed tires. For the P245/70R16 tires,

- *P* means passenger tire;
- *245* specifies the tire’s width in millimeters;
- *70* is the tire’s aspect ratio—that is, the ratio of the tire’s height to its width reported as a percent; and
- *16* is the diameter of the tire’s rim in inches.



Analyzing Numerical Data: Using Ratios

I.B Student Activity Sheet 5: Changing Tires

1. Fill in the missing information for each tire size. Find the circumference of each tire.

Tire	P245/70R16	P285/75R16
Width (mm)		
Aspect ratio (%)		
Height (in.)		
Diameter (in.)		
Circumference (in.)		

Analyzing Numerical Data: Using Ratios

I.B Student Activity Sheet 5: Changing Tires

2. After one rotation of the wheel, how many inches further has the truck with the larger tires traveled than the truck with the factory-installed tires?

3. After one rotation of the wheel, the truck with the larger tires has traveled _____ times further than the truck with the factory-installed tires.

4. Use the results from the table in Question 1 to assist in completing the following statements about the truck after the larger tires have been installed on it.

If the odometer reading is $\boxed{2}\boxed{0}\boxed{0}\boxed{0}\boxed{0}$, you have actually traveled _____ miles.

If the speedometer reading is 60, your actual speed is _____ miles per hour.

The following principles apply when determining actual distance and speed traveled according to tire size:

Actual mileage = $k \cdot$ odometer reading (mileage)

Actual speed = $k \cdot$ speedometer reading (miles per hour)

where $k = \frac{\text{circumference of bigger tire}}{\text{circumference of factory-installed tire}}$

Analyzing Numerical Data: Using Ratios

I.B Student Activity Sheet 5: Changing Tires

5. What is the percent error in the odometer readings? In the speedometer readings?
6. Using the odometer readings in the truck equipped with the larger tires, you determine that the gas mileage is 18 miles per gallon. What is your actual gas mileage in miles per gallon?
7. If you were driving in the truck with the larger tires and the speedometer showed a speed of 65 miles per hour, could you be ticketed for exceeding the 65-mph speed limit by more than 5 mph? More than 10 mph? Justify your answers.
8. **REFLECTION:** What is the relationship between the ratio of an actual distance to an odometer distance of 1 mile and the ratio of the circumference of a current tire to the circumference of a factory-installed tire?

Analyzing Numerical Data: Using Ratios

I.B Student Activity Sheet 5: Changing Tires

9. **EXTENSION:** On your new small car, you replace the factory-installed P185/75R14 tires with slightly larger P205/75R14 tires. Find the missing number in each statement:

If your odometer reading is $\boxed{2}\boxed{0}\boxed{0}\boxed{0}\boxed{0}$ (miles), you have actually traveled _____ miles.

If your speedometer reading is 60, your actual speed is _____ miles per hour.

Analyzing Numerical Data: Using Ratios

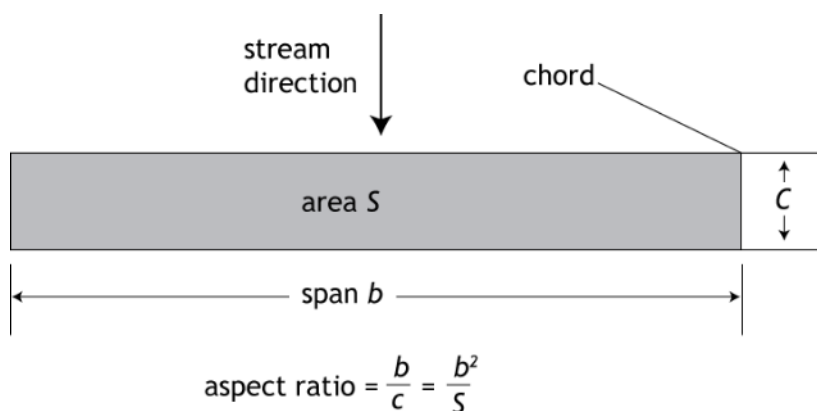
I.B Student Activity Sheet 5: Changing Tires

10. EXTENSION

The ability of birds and airplanes to fly is related to the aspect ratio of their wings.

Since early airplane wings were usually nearly rectangular, the aspect ratio of a wing was the ratio of the span of a wing or airfoil to the chord of a wing, where the span is the maximum cross-stream dimension (longer side) and the chord is the dimension in the stream-wise direction (shorter side), as shown in the figure below. Later, when wings were tapered or had complex platforms, another definition became necessary;

currently, the general definition of aspect ratio is $A = \frac{b^2}{S}$, where A is the aspect ratio, b is the span, and S is the area.



Demonstrate that the current definition of the aspect ratio of a wing coincides with the earlier definition when the wing was rectangular.

Conduct independent research on how the aspect ratio of a wing is related to the flight of airplanes and birds.

Analyzing Numerical Data: Indices Using Weighted Sums and Averages**I.C Student Activity Sheet 6: Final Grade Averages**

When a *weighted average* is applied to a set of numbers, more importance (weight) is placed on some components of the set. Your final average in this class is probably an example of a weighted average.

Consider two grading systems for determining your final class average. Each system is a weighted average of measures that include test grades, final exam grade, homework, and class participation.

Grading System I	Grading System II
Test average—40%	Test average—60%
Final exam grade—25%	Final exam grade—15%
Homework—25%	Homework—15%
Class participation—10%	Class participation—10%

1. If your values are the following, which grading system do you prefer and why?
 - Test average = 84
 - Final exam grade = 68
 - Homework = 90
 - Class participation = 95

2. If you score 10 points higher on the final exam, how does your final grade average change under each system?

Analyzing Numerical Data: Indices Using Weighted Sums and Averages

I.C Student Activity Sheet 6: Final Grade Averages

3. If you score 6 points lower on the final exam, how does your final grade average change under each system? Which system is better for you?
4. Use the following information to find your final course average in each grading system:
- Test grades {80, 74, 82, 88}
 - Final exam grade = 84
 - Homework = 90
 - Class participation = 95

Analyzing Numerical Data: Indices Using Weighted Sums and Averages

I.C Student Activity Sheet 6: Final Grade Averages

5. Your averages and values are the following:

- Test average = 85
- Homework = 90
- Participation = 95

What grade do you need on the final exam to earn a final grade average of at least 87 in each grading system?

6. **REFLECTION:** What weights would you assign to each component to set up a grading system? Each weight must be at least 10%. Why do you think your grading system would be fair and effective?

Analyzing Numerical Data: Indices Using Weighted Sums and Averages**I.C Student Activity Sheet 7: Slugging Averages**

One example of a *weighted average* in sports is a batter's *slugging average* (or *percentage*) in baseball. The slugging average (*SLG*) is calculated using the following equation:

$$SLG = \frac{(1 \cdot S) + (2 \cdot D) + (3 \cdot T) + (4 \cdot HR)}{AB}$$

where *S* = singles, *D* = doubles, *T* = triples, *HR* = home runs, and *AB* = total at-bats.

Each single has a weight of 1, each double a weight of 2, each triple a weight of 3, and each home run a weight of 4. An at-bat without a hit has a weight of 0.

In his first season with the New York Yankees, Babe Ruth set a record for slugging average that stood for more than 80 years. In 1920, Ruth pounded 172 hits in 458 at-bats. His hits consisted of 73 singles, 36 doubles, 9 triples, and 54 home runs, resulting in a total base count of $(73 \cdot 1) + (36 \cdot 2) + (9 \cdot 3) + (54 \cdot 4) = 388$. When his total number of bases (388) is divided by his total at-bats (458), the result is .847, his slugging percentage for the season. This record was broken in 2001 by Barry Bonds, who had 411 total bases in 476 at-bats for a slugging average of .863. (Statistics from www.baseball-almanac.com)

1. Find the slugging average for a player with the following statistics:

$$S = 68$$

$$D = 40$$

$$T = 4$$

$$HR = 16$$

$$AB = 320$$

Analyzing Numerical Data: Indices Using Weighted Sums and Averages

I.C Student Activity Sheet 7: Slugging Averages

2. **REFLECTION:** Is it possible to have a slugging average of more than 1?

Theoretically, what is the highest possible value for the slugging average? Could a player ever achieve this value during a baseball season? Give an example or explain why none exist.

3. A slugging average of .500 or higher is considered a sign of an excellent player. Suppose a player had 4 triples and a batting average of .300 in 400 at-bats (batting average = hits/at-bats). Determine a combination of singles, doubles, and home runs that gives this player a slugging average higher than .500.

Still using 400 at-bats, what is the maximum number of singles this player could have with a slugging average between .500 and .700? Justify your answer.

Analyzing Numerical Data: Indices Using Weighted Sums and Averages

I.C Student Activity Sheet 7: Slugging Averages

4. **EXTENSION:** Research the slugging averages of some current-day professional baseball players. Then research the slugging averages of some college players. Based on your findings, which college players should professional scouts be watching? Support your answer with statistics.

Analyzing Numerical Data: Indices Using Weighted Sums and Averages

I.C Student Activity Sheet 8: Quarterback Ratings

The National Football League (NFL) rates quarterbacks for statistical purposes against a fixed performance standard based on the statistical achievements of all qualified pro passers since 1960. This system allows passing performances to be compared from one season to the next.

The following categories are used to compute the quarterback rating:

- percent of completions per attempt (**%COMP**)
- percent of touchdown passes per attempt (**%TD**)
- percent of interceptions per attempt (**%INT**)
- average yards gained per attempt (**YD**)

(From www.nfl.com/help/quarterbackratingformula)

The following is the formula for compiling the quarterback rating (**QR**):

$$QR = \frac{25 + 10(\%COMP) + 40(\%TD) - 50(\%INT) + 50(YD)}{12}$$

(Note: This formula is subject to a few conditions discussed after Question 2.)

1. For the first two games of the 2008 season, Dallas Cowboys quarterback Tony Romo completed 45 passes in 62 attempts for a total of 632 yards, with 4 touchdowns and 2 interceptions. Verify that Romo's quarterback rating for those games is approximately 113. Round each value to the nearest tenth.

Percent of completions per attempt (**%COMP**) =

Percent of touchdown passes per attempt (**%TD**) =

Percent of interceptions per attempt (**%INT**) =

Average yards gained per attempt (**YD**) =

2. As of 2009, Steve Young has the highest career quarterback rating in NFL history. He completed 2,059 passes in 3,192 attempts for a total of 25,479 yards, with 174 touchdowns and 85 interceptions. Find Young's career quarterback rating.

Analyzing Numerical Data: Indices Using Weighted Sums and Averages**I.C Student Activity Sheet 8: Quarterback Ratings**

The actual formula for the quarterback rating is subject to some conditions, represented below.

$$QR = \left[\frac{5(\%COMP - 30)}{6} \right] + \left[\frac{10(\%TD)}{3} \right] + \left[\frac{25(19 - 2(\%INT))}{12} \right] + \left[\frac{25(YD - 3)}{6} \right]$$

where it is understood that each item in brackets [] is truncated to be no smaller than 0 and no larger than $475/12$ (giving a maximum rating of $475/3$ or 158.3). This rating is rounded to the nearest tenth.

These truncated values mean there is a minimum and maximum value for each component of the rating.

For example, if $\frac{10(\%TD)}{3} \leq \frac{475}{12}$, then $\%TD \leq \frac{475 \cdot 3}{12 \cdot 10} = 11.875\%$.

3. Show that $30\% \leq \%COMP \leq 77.5\%$.

4. If $c \leq YD \leq d$, find the values of c and d .

5. If $0 \leq \%INT \leq e$, find the value of e .

Analyzing Numerical Data: Indices Using Weighted Sums and Averages

I.C Student Activity Sheet 8: Quarterback Ratings

6. Explain why the actual formula and the simplified formula are the same when each component is between its minimum and maximum possible value.

$$\left[\frac{5(\%COMP - 30)}{6} \right] + \left[\frac{10(\%TD)}{3} \right] + \left[\frac{25(19 - 2(\%INT))}{12} \right] + \left[\frac{25(YD - 3)}{6} \right] =$$

$$\frac{10(\%COMP) - 300 + 40(\%TD) + 25 \cdot 19 - 50(\%INT) + 50(YD) - 150}{12}$$

7. **REFLECTION:** An index is a numerical scale. Characteristics of an index can be used for the following:

- to compare variables with one another or a reference number,
- to give information about general trends, and
- to help make comparisons and judgments.

It is often calculated as a weighted sum of various factors resulting in a single summary number. How can the quarterback rating system be thought of as an index?

Analyzing Numerical Data: Indices Using Weighted Sums and Averages

I.C Student Activity Sheet 9: Fan Cost Index

An *index* is a numerical scale. Characteristics of an index can be used for the following:

- to compare variables with one another or a reference number,
- to give information about general trends, and
- to help make comparisons and judgments.

It is often calculated as a weighted sum of various factors resulting in a single summary number.

The Fan Cost Index (FCI), compiled by Team Marketing Report, tracks the cost for a family of four to attend a professional sporting event. The FCI includes the prices of 2 average-price adult tickets, 2 average-price child tickets, 4 small soft drinks, 2 small other drinks, 4 regular-size hot dogs, parking for 1 car, 2 game programs, and 2 least expensive, adult-size adjustable caps. The Average Ticket Price in the following tables represents the average cost of a ticket for each member of the family.

The FCI for each Texas team in professional baseball, basketball, and football for 2006 and 2007 is shown in the following two tables. The Soft Drink column includes the price of 1 drink and its size in ounces.

2006 Fan Cost Index

Team	Avg. Ticket Price	Soft Drink (size in oz)	Other Drink	Hot Dog	Parking	Program	Cap	FCI
Rangers	15.81	2.75 (16)	6.00	2.50	8.00	5.00	10.00	134.24
Astros	26.66	4.00 (21)	7.00	4.00	10.00	4.00	11.00	192.64
Spurs	45.88	3.75 (24)	6.25	4.00	10.00	5.00	18.00	283.02
Mavericks	54.24	2.75 (12)	4.00	4.00	20.00	6.00	15.00	313.96
Rockets	38.64	4.00 (22)	5.75	4.25	15.00	—	20.00	254.06
Cowboys	66.12	3.50 (32)	5.00	3.50	12.00	5.00	10.00	344.48
Texans	56.97	3.25 (21)	6.00	5.00	15.00	5.00	20.00	337.88

(Compiled from www.teammarketing.com)

Analyzing Numerical Data: Indices Using Weighted Sums and Averages
 I.C Student Activity Sheet 9: Fan Cost Index

2007 Fan Cost Index

Team	Avg. Ticket Price	Soft Drink (size in oz)	Other Drink	Hot Dog	Parking	Program	Cap	FCI
Rangers	16.47	3.00 (20)	6.50	2.75	8.00	5.00	11.00	141.88
Astros	26.90	4.00 (21)	7.00	4.25	10.00	4.00	12.00	196.60
Spurs	51.24	2.50 (24)	6.00	4.00	10.00	6.00	21.00	306.96
Mavericks	60.58	2.75 (15)	4.00	4.00	20.00	6.00	15.00	339.32
Rockets	41.98	4.00 (22)	7.00	4.50	15.00	—	20.00	270.92
Cowboys	84.12	3.50 (32)	5.00	3.50	12.00	5.00	10.00	416.48
Texans	62.41	3.25 (21)	6.00	5.00	15.00	5.00	20.00	359.64

(Compiled from www.teammarketing.com)

1. Complete the 2006 table by calculating the missing FCI values.

Analyzing Numerical Data: Indices Using Weighted Sums and Averages

I.C Student Activity Sheet 9: Fan Cost Index

The following table contains the average FCI for all teams in Major League Baseball (MLB), the National Basketball Association (NBA), and the National Football League (NFL) for 2002 through 2007.

FCI Average

	2002	2003	2004	2005	2006	2007
MLB	145.26	148.66	155.52	164.43	171.19	176.55
NBA	254.86	261.26	263.44	267.37	274.67	281.90
NFL	290.41	301.75	321.62	329.91	346.16	367.31

This table contains the rates of inflation from 2003 through 2007.

Year	2003	2004	2005	2006	2007
Inflation Rate (%)	2.27	2.68	3.39	3.24	2.85

Use this information and additional calculations to answer the following questions.

- Did the rate of increase of each sport's average FCI exceed the rate of inflation for all these years? Find the sport, year, and rate of increase for all cases in which the average FCI rate increase did not exceed the inflation rate.
- Find the sport and year in which the average rate of increase exceeded the inflation rate by the greatest amount.

Analyzing Numerical Data: Indices Using Weighted Sums and Averages

I.C Student Activity Sheet 9: Fan Cost Index

7. **REFLECTION:** Why do you think the FCI is highest for the NFL and lowest for MLB?

8. **EXTENSION:** There are several indices related to consumer spending. Two of them are the Consumer Price Index (CPI) and the Consumer Confidence Index (CCI). Find out how they are calculated and how they are used.

Analyzing Numerical Data: Indices Using Weighted Sums and Averages

I.C Student Activity Sheet 10: Readability Indices

An *index* is a numerical scale. Characteristics of an index can be used for the following:

- to compare variables with one another or a reference number,
- to give information about general trends, and
- to help make comparisons and judgments.

It is often calculated as a weighted sum of various factors resulting in a single summary number.

Fog Index

The Fog Index, which was developed by Robert Gunning in 1952, measures the readability of a piece of writing. Calculation of the index produces a number that indicates the years of education a person needs to have to easily understand the text. For example, a passage with a Fog Index of 8 has an eighth-grade reading level. The following table provides more detail on the Fog Index.

Use	Writers can use the Fog Index to gauge whether or not their material is suitable for the targeted audience. The ideal Fog Index number for a writing piece for the general population is 7 or 8. Technical material usually has a Fog Index of 10 to 15.
Calculation	<ol style="list-style-type: none"> 1. Select a block of text that is approximately 100 words. 2. Compute the average sentence length of the selected passage, dividing the number of words by the number of sentences. 3. Count words with three or more syllables (that is, complex words). Exclude proper nouns, compound words, words ending in suffixes such as <i>-es</i>, <i>-ed</i>, or <i>-ing</i> that result in a third syllable, terminology familiar to the audience, and hyphenated words. Divide the number of complex words by the total number of words. 4. Add the average sentence length and the percent of complex words. 5. Multiply the result by 0.4.
Formula	$FI = 0.4 \left(\frac{\text{number of words}}{\text{number of sentences}} + 100 \cdot \frac{\text{number of complex words}}{\text{number of words}} \right)$
Limitations	Considered an accurate indication of readability, the Fog Index has limitations. For example, some multisyllabic words are not considered to be complex by the general population, but the index does not take this into account.

Analyzing Numerical Data: Indices Using Weighted Sums and AveragesI.C Student Activity Sheet 10: Readability Indices

Flesch-Kincaid Grade Level Index

Another widely used measure is the Flesch/Flesch-Kincaid Readability Tests, which consist of the Flesch Reading Ease (developed by Rudolph Flesch in 1948) and the Flesch-Kincaid Grade Level (developed by John P. Kincaid in 1976). The Flesch-Kincaid Grade Level Index converts the Flesch Reading Ease into a number that indicates the years of education a person needs to have to easily understand the text.

Calculate the Flesch-Kincaid Grade Level Index (**FKGLI**) using the following formula:

$$FKGLI = 0.39 \left(\frac{\text{number of words}}{\text{number of sentences}} \right) + 11.8 \left(\frac{\text{number of syllables}}{\text{number of words}} \right) - 15.59$$

Analyzing Numerical Data: Indices Using Weighted Sums and Averages

I.C Student Activity Sheet 10: Readability Indices

Declaration of Independence (excerpt)

We hold these truths to be self-evident, that all men are created equal, that they are endowed by their Creator with certain unalienable Rights, that among these are Life, Liberty and the pursuit of Happiness. – That to secure these rights, Governments are instituted among Men, deriving their just powers from the consent of the governed. – That whenever any Form of Government becomes destructive of these ends, it is the Right of the People to alter or to abolish it, and to institute new Government, laying its foundation on such principles and organizing its powers in such form, as to them shall seem most likely to effect their Safety and Happiness.

1. Find the Fog Index and the Flesch-Kincaid Grade Level Index of the excerpt of the *Declaration of Independence*.

Check your answer for the Flesch-Kincaid Grade Level Index using MS Word. (Go to the Help section and look up Readability to find how to access this score.)

Analyzing Numerical Data: Indices Using Weighted Sums and Averages

I.C Student Activity Sheet 10: Readability Indices

One of the most famous speeches in U.S. history is the *Gettysburg Address*—a 275-word speech that required only a few minutes to deliver.

Gettysburg Address

Four score and seven years ago our fathers brought forth on this continent a new nation, conceived in Liberty, and dedicated to the proposition that all men are created equal.

Now we are engaged in a great civil war, testing whether that nation, or any nation, so conceived and so dedicated, can long endure. We are met on a great battlefield of that war. We have come to dedicate a portion of that field, as a final resting place for those who here gave their lives that that nation might live. It is altogether fitting and proper that we should do this.

But, in a larger sense, we cannot dedicate—we cannot consecrate — we cannot hallow — this ground. The brave men, living and dead, who struggled here, have consecrated it, far above our poor power to add or detract. The world will little note, nor long remember what we say here, but it can never forget what they did here. It is for us the living, rather, to be dedicated here to the unfinished work which they who fought here have thus far so nobly advanced. It is rather for us to be here dedicated to the great task remaining before us — that from these honored dead we take increased devotion to that cause for which they gave the last full measure of devotion — that we here highly resolve that these dead shall not have died in vain — that this nation, under God, shall have a new birth of freedom — and that government of the people, by the people, for the people, shall not perish from the earth.

2. Find the Fog Index and the Flesch-Kincaid Grade Level Index for the *Gettysburg Address*.

Check your answer for the Flesch-Kincaid Grade Level Index using MS Word.

Analyzing Numerical Data: Indices Using Weighted Sums and Averages

I.C Student Activity Sheet 10: Readability Indices

3. Find the Fog Index and the Flesch-Kincaid Grade Level Index for a passage of around 100 words from one of your favorite books or short stories.

4. Find the Fog Index and the Flesch-Kincaid Grade Level Index for a passage of around 100 words from your science textbook.

5. Write the copy for a radio commercial advertising a real or imagined product. The potential buyers of the product are high school students requiring a Fog Index and a Flesch-Kincaid Grade Level Index of between 10 and 12.

Analyzing Numerical Data: Indices Using Weighted Sums and Averages

I.C Student Activity Sheet 10: Readability Indices

6. **EXTENSION:** There are several indices related to consumer spending. Two of them are the Consumer Price Index (CPI) and the Consumer Confidence Index (CCI). Find out how they are calculated and how they are used.

7. **REFLECTION:** When do you think it would be important to pay attention to the reading level of a passage you write? Think of careers that would be affected by reading levels.

Analyzing Numerical Data: Indices Using Weighted Sums and Averages

I.C Student Activity Sheet 11: Simpson's Paradox

Sabermetrics is the statistical analysis of baseball data. This field of study takes its name from the acronym of the Society for American Baseball Research (SABR). A sabermetrician uses statistics to answer objective questions about baseball players.

1. The Smith brothers (Jo, Bo, and Mo) are talented baseball players. In fact, each one had a .300 batting average going into the final day of the season. During the final game, Jo went 3 for 5, Bo went 4 for 5, and Mo went 5 for 5. Assuming Jo, Bo, and Mo had 240, 420, and 600 official at-bats, respectively, before the final game, who ended the season with the highest batting average? Justify your answer.

2. The previous problem is an example of Simpson's Paradox—averaging the averages of different populations does not necessarily give the average of the combined population. In this case, each player's batting average for the entire season does not equal the combination of his batting average for all but the last day and his batting average on the last day. This paradox is named for Edward H. Simpson, who described the phenomenon in his 1951 paper.

Find out more about Simpson's Paradox and other real-world situations where it occurred. Present your findings to the class.

3. **REFLECTION:** Have you ever experienced a paradox? Describe the situation.

Analyzing Numerical Data: Validating Identification Numbers

I.D Student Activity Sheet 12: Universal Product Codes

Identification numbers are present everywhere in society. Today's identification numbers are more sophisticated than those introduced years earlier (for example, Social Security numbers). Today's numbers have a check digit to partially ensure that they have been correctly scanned or entered into a computer.

Universal Product Codes (UPCs), typically in the form of barcodes, identify retail products.



The 12-digit UPC barcode consists of three parts:

- manufacturer number,
- product number, and
- check digit.

For example, the manufacturer number for the Dr. Pepper Company is 078000 and appears in the first six digits of all of the company's product UPC barcodes. GS1, formerly the Uniform Code Council, issues a company this six-digit number. Every item sold by a company requires a different five-digit product number. This includes specific products, their different sizes, their array of colors, their variety of flavors, and other distinguishing features. The last number is the **check digit**, which guards against entry errors and fraud.

The check digit in a UPC number (that is, the twelfth digit) is determined in the following manner:

- Multiply the first digit by 3.
- Add the second digit.
- Multiply the third digit by 3.
- Add the fourth digit.
- Continue this alternating process for the Digits 5 to 12.

Analyzing Numerical Data: Validating Identification Numbers

I.D Student Activity Sheet 12: Universal Product Codes

The check digit is chosen so that the calculation described previously totals a number whose final digit is 0. In the UPC number $a_1a_2a_3a_4a_5a_6a_7a_8a_9a_{10}a_{11}d$, the check digit is d , for which the sum

$$3a_1 + a_2 + 3a_3 + a_4 + 3a_5 + a_6 + 3a_7 + a_8 + 3a_9 + a_{10} + 3a_{11} + d$$

ends in 0. In this weighted sum, the weights are: {3, 1, 3, 1, 3, 1, 3, 1, 3, 1, 3, 1}.

When entering a code number, the *single-digit error* is most common (for example, keying in 8 instead of 3). Another common error is the transposition error, where the order of two adjacent digits is reversed (for example, writing 83 instead of 38). Systems have been established to detect and correct (when possible) these and other errors almost immediately.

1. Show that 0-58200-48826-5 is a valid UPC number.

2. Show that 0-52200-48826-5 is an invalid UPC number.

If someone made a single-digit error when entering this invalid number, can you tell which digit is incorrect? Why or why not?

Change one digit in this invalid number so the resulting number is valid.

Analyzing Numerical Data: Validating Identification Numbers

I.D Student Activity Sheet 12: Universal Product Codes

3. Determine the check digit (d) for the UPC number 38137009213 d .

4. Suppose you entered 8 instead of 9 when recording the UPC number 1 55210 02149 6. Explain why the UPC method will detect this error.

Do you think the UPC method will detect all such single-digit errors? Either give several examples of numbers with an error that will not be detected or explain why you think all such single-digit errors will be detected.

Analyzing Numerical Data: Validating Identification Numbers

I.D Student Activity Sheet 12: Universal Product Codes

5. Suppose you use the weights $\{7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1\}$ instead of $\{3, 1, 3, 1, 3, 1, 3, 1, 3, 1, 3, 1\}$ in the UPC method. Do you think this altered method will detect single-digit errors? Either give several examples of numbers with an error that will not be detected or explain why you think all such single-digit errors will be detected.

6. Suppose you use the weights $\{2, 1, 2, 1, 2, 1, 2, 1, 2, 1, 2, 1\}$ instead of $\{3, 1, 3, 1, 3, 1, 3, 1, 3, 1, 3, 1\}$ in the UPC method. Do you think this altered method will detect single-digit errors? Either give several examples of numbers with an error that will not be detected or explain why you think all such single-digit errors will be detected.

7. When entering UPC numbers, will the error of transposing 2 and 7 be detected? That is, ... 72 ... is recorded as ... 27 How about 2 and 6? Make a conjecture about the values of a and b for which the error of transposing ab and ba is detected by the UPC method. Explain your reasoning in making your conjecture.

Analyzing Numerical Data: Validating Identification Numbers

I.D Student Activity Sheet 12: Universal Product Codes

8. **REFLECTION:** In what ways is the UPC check digit calculation like the activities you explored in Section C, “Indices Using Weighted Sums and Averages”?

9. EXTENSION

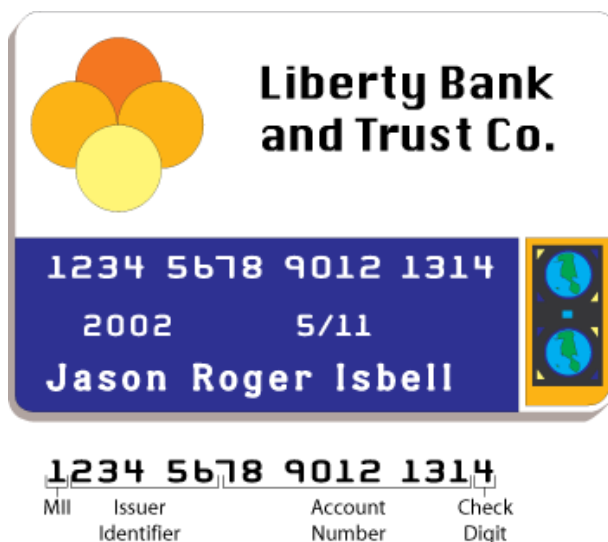
- The European Article Numbering Code (EAN 13) is the worldwide standard. Conduct research on how the EAN 13 number is calculated and what types of errors this method detects. Prepare a report and presentation about your findings.
- The International Standard Book Numbers (ISBNs) used to identify books have a check-digit scheme that detects all single-digit and transposition errors. Conduct research on how the ISBN number is calculated and why this method detects all single-digit and transposition errors. Prepare a report and presentation about your findings.
- In a vehicle identification number (VIN), the check digit comes in the middle of the number instead of at the end. Conduct research on how the VIN is calculated and what types of errors this method detects. Prepare a report and presentation about your findings.
- Conduct research on the meaning of various digits in credit card numbers and how check digits are determined. Prepare a report and presentation about your findings.
- Imagine that you are employed by a company that does not have an identification code number for its employees and products. Prepare a report discussing various numbering alternatives and make a recommendation. Include a check digit in your code number. Indicate how successful your method of determining the check digit will be in detecting single-digit and transposition errors. The company can be your school.

Analyzing Numerical Data: Validating Identification Numbers

I.D Student Activity Sheet 13: Credit Card Numbers

Identification numbers are present everywhere in society. Today’s identification numbers are more sophisticated than those introduced years earlier (for example, Social Security numbers). Today’s numbers have a check digit to partially ensure that they have been correctly scanned or entered into a computer.

Credit cards have 16-digit numbers, of which the first 15 digits identify the credit card and the sixteenth digit is the check digit. The following figure shows the significance of the digits:



MII stands for *major industry identifier*; 4 and 5 indicate “Banking and Financial.” VISA cards begin with 4 and MasterCard cards with 5.

1. MasterCard numbers begin with 51, 52, 53, 54, or 55. What is the maximum number of credit cards that MasterCard can issue?

Analyzing Numerical Data: Validating Identification Numbers

I.D Student Activity Sheet 13: Credit Card Numbers

A **check digit** is used to help validate credit card numbers. The credit card companies use the Codabar method to determine the check digit. This method consists of the following steps:

- Add the digits in the odd-numbered positions and double this total.
- Add the *number* of odd-position digits that are more than 4 to the total.
- Add the even-position digits.
- Choose a check digit that makes this calculation total a number whose final digit is 0.

Libraries, shipping/receiving companies, and blood banks also use the Codabar method.

2. Show that the check digit (d) for the VISA card 4162 0012 3456 789 d is 3.

3. What is the check digit (d) for the MasterCard number 5424 9813 2720 008 d ?

Analyzing Numerical Data: Validating Identification Numbers

I.D Student Activity Sheet 13: Credit Card Numbers

4. Show that 4128 0012 4389 0110 is an invalid VISA credit card number.

If someone made a single-digit error when entering this invalid number, can you tell which digit is incorrect? Why or why not?

Change one digit in this invalid number so the resulting number is valid.

5. The following is another way to explain the Codabar method:

- If a digit is in an even-numbered position, add it to the total.
- If a digit is in an odd-numbered position, multiply it by 2. If the product is equal to or greater than 10, subtract 9 from the product. Add this difference to the total.
- After the first 15 digits have been processed, choose the check digit so that the sum of the 16 digits ends in 0.

Explain why both methods yield the same check digit.

Analyzing Numerical Data: Validating Identification Numbers

I.D Student Activity Sheet 13: Credit Card Numbers

6. Suppose you entered 8 instead of 9 when recording the credit card number 4001 2560 0196 4310. Explain why the Codabar method will detect this error.

Do you think the Codabar method will detect all such single-digit errors using a method similar to the one for showing all such errors are detected for UPC numbers? Either write *Yes* or give several examples of numbers with an error that will not be detected.

7. The Codabar method detects 98% of all transposition errors. There is one pair of numbers, however, that when transposed will not be detected by the Codabar method. Find this pair of numbers and explain why this transposition error will go undetected. (*Hint:* There are several cases to consider depending on the positions of the transposed digits.)

Analyzing Numerical Data: Validating Identification Numbers

I.D Student Activity Sheet 13: Credit Card Numbers

8. **REFLECTION:** In what ways is the Codabar method like the activities you explored in Section C, “Indices Using Weighted Sums and Averages”?

9. EXTENSION

- The European Article Numbering Code (EAN 13) is the worldwide standard. Conduct research on how the EAN 13 number is calculated and what types of errors this method detects. Prepare a report and presentation about your findings.
- The International Standard Book Numbers (ISBNs) used to identify books have a check-digit scheme that detects all single-digit and transposition errors. Conduct research on how the ISBN number is calculated and why this method detects all single-digit and transposition errors. Prepare a report and presentation about your findings.
- In a vehicle identification number (VIN), the check digit comes in the middle of the number instead of at the end. Conduct research on how the VIN is calculated and what types of errors this method detects. Prepare a report and presentation about your findings.
- Retailers use Universal Product Codes (UPC) to identify products. Conduct research on the meaning of various digits in the UPC number and how check digits are determined. Prepare a report and presentation about your findings.
- Imagine that you are employed by a company that does not have an identification code number for its employees and products. Prepare a report discussing various numbering alternatives and make a recommendation. Include a check digit in your code number. Indicate how successful your method of determining the check digit will be in detecting single-digit and transposition errors. The company can be your school.