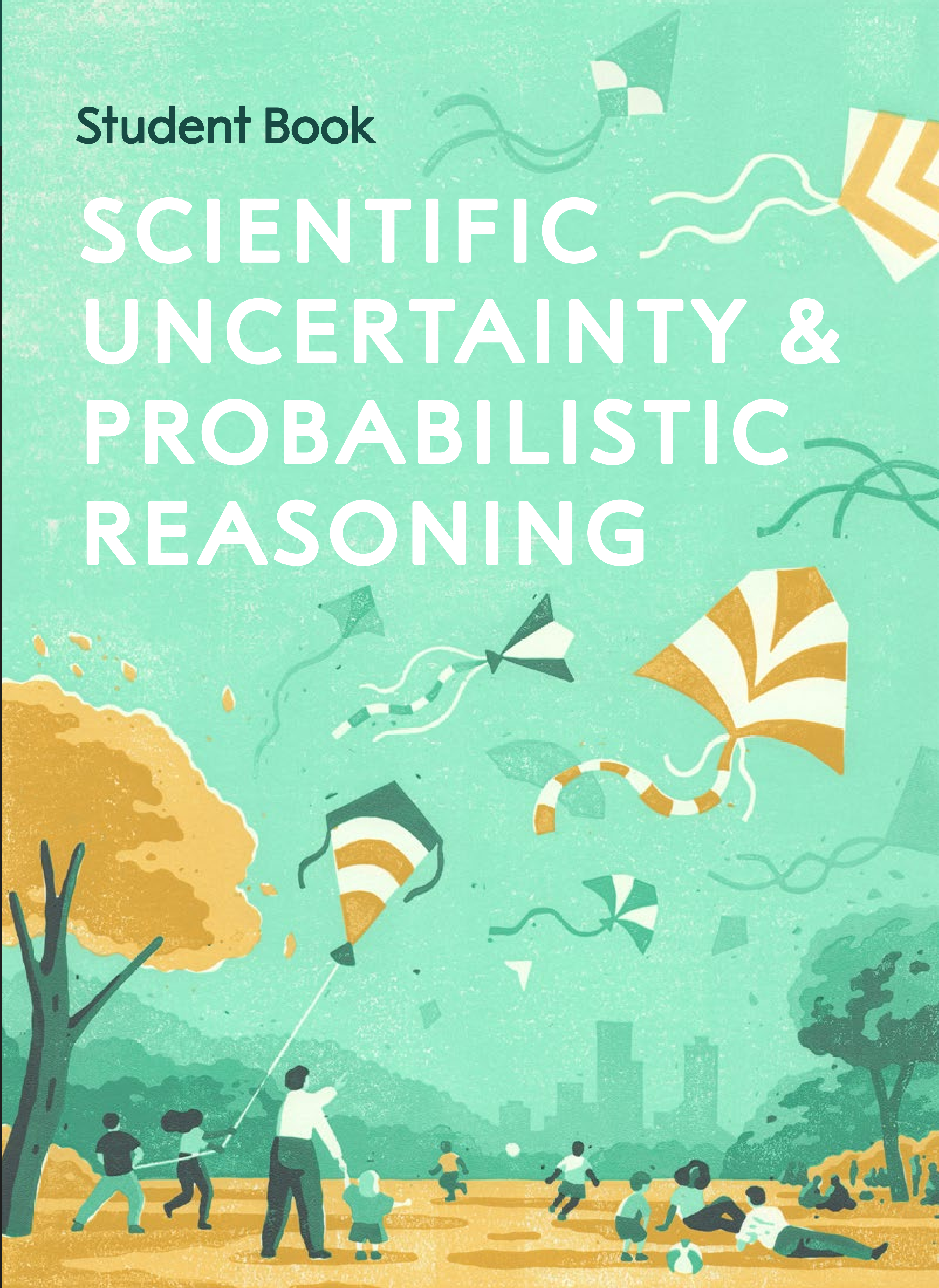
UNIT 3

v 1.0

Student Book

SCIENTIFIC UNCERTAINTY & PROBABILISTIC REASONING



SCIENTIFIC THINKING FOR ALL : A TOOLKIT

This book is part of the *Scientific Thinking for All: A Toolkit* curriculum that is a high school adaptation of the University of California, Berkeley, "Big Ideas" course titled Sense and Sensibility and Science https://sensibility.berkeley.edu/. It was developed by professors Saul Perlmutter, John Campbell, and Robert MacCoun and represents a collaboration among physics, philosophy, and psychology. *Scientific Thinking for All: A Toolkit* was developed by curriculum developers and researchers at The Lawrence Hall of Science, University of California. The initiative is a cooperation between Nobel Prize Outreach (NPO) and Saul Perlmutter. This work is supported by a consortium of funders including Kenneth C. Griffin, the William and Flora Hewlett Foundation, the John D. and Catherine T. MacArthur Foundation, the Gordon and Betty Moore Foundation, and The Rockefeller Foundation.

SCIENTIFIC THINKING FOR ALL TEAM

PRINCIPAL INVESTIGATOR
Saul Perlmutter

PROJECT DIRECTOR
Ben Koo

PROJECT LEAD
Janet Bellantoni

UNIT 3 AUTHORS
Manisha Hariani
Kristina Duncan
David House
Emlen Metz

OTHER CONTRIBUTORS
Janet Bellantoni
Maia Binding
Eric Greenwald
Kelly Grindstaff
Sara Kolar
Ben Koo

FIELD TEST COORDINATOR
Kelly Grindstaff

RESEARCH
Eric Greenwald
Kelly Grindstaff
Crystal Menzies

SCIENTIFIC REVIEW
Lynsey K. Parker
Master of Science in Civil and Environmental Engineering, University of California, Berkeley

PRODUCTION

EDITING
Trudihope Schlomowitz

DESIGN
otherwise

TOOL ICONOGRAPHY
DOT Stockholm

COVER ILLUSTRATION
Merijn Jansen

Scientific Thinking For All : A Toolkit

UNIT 3
Scientific Uncertainty & Probabilistic Reasoning
v1.0 AUG 2025

STUDENT BOOK

THE LAWRENCE HALL OF SCIENCE
University of California, Berkeley
Berkeley, CA 94720-5200

DEAR STUDENT,

Have you ever read a news article and wondered if the information it contained was true? Or had a friend make a claim and wondered if they were right? This curriculum will equip you with ideas and techniques from science that can be applied to everyday life. Your conceptual toolkit will include strategies to help you evaluate information, reflect on your thinking, and make more informed decisions. You will use these tools to ask questions, brainstorm ideas, interpret data, manage trade-offs, and develop solutions.

Scientific tools and techniques can lead to a better understanding of how the natural world works and provide approaches to solving the problems facing individuals, communities, and the environment. Each unit will provide you with additional conceptual tools for your toolkit, and you'll practice applying them to personal and societal issues such as human health, environmental pollution, and energy use. For example, you may consider questions such as: *How much sleep do I need? Is my community's drinking water safe?*

Science offers many useful strategies for learning about the world, including:

1. working together to share observations, questions, and ideas;
2. techniques for making sense of observations and data; and
3. the iteration of ideas by modifying them as new information becomes available.

Since it's difficult for anyone to catch their own mistakes, you'll collaborate with your classmates to share your thinking and learn from one another. It is our hope that this science toolkit will empower you to think more clearly about the things you care about, to provide you with strategies for addressing problems, and to help you achieve your personal goals.

Sincerely,

Scientific Thinking for All Program Team

TABLE OF CONTENTS

Scientific Uncertainty & Probabilistic Reasoning

UNIT 3

SCIENTIFIC UNCERTAINTY & PROBABILISTIC REASONING

Scientific investigations try to answer questions about the natural world, such as identifying components of Earth's atmosphere and how air quality affects health. This can involve a certain amount of scientific uncertainty. Uncertainty in science can come from incomplete information or scientific errors. In this unit, you will explore air quality data to learn how to identify uncertainty and errors in science. By recognizing these uncertainties and errors, you will learn how to reduce them and become more sure of your findings. You will use probabilistic reasoning to make predictions, such as the likelihood of wildfire spread. You will also look for meaningful signals in data and understand how false positives and false negatives can influence decisions. Throughout the unit, you will apply these conceptual tools to air quality issues at both the local and global levels.

UNIT DRIVING QUESTION

How do you address scientific uncertainty when investigating claims about air quality?

CONCEPTUAL TOOLS

In this course, conceptual tools refer to scientific ideas and approaches that can be applied to real-world situations. Each conceptual tool is further explained in the activity in which it is introduced. The conceptual tools found in this unit are shown here and in each activity in which they appear.

Probabilistic Reasoning

Signal & Noise

Systematic & Random Errors

False Positives & False Negatives

ACTIVITY 1

Investigating Probabilistic Reasoning

CARD-BASED INVESTIGATION

Exposure to poor air quality can result in symptoms similar to those of other illnesses.

1 : INVESTIGATING PROBABILISTIC REASONING

GUIDING QUESTION

How do you make predictions with incomplete information?

INTRODUCTION

Since the middle of the 1900s, many countries have been working to improve air quality. In the United States, the first Clean Air Act was passed in 1963. This law helped reduce air pollutants such as ground-level ozone and sulfur dioxide. Another air pollutant is particulate matter (PM), microscopic particles suspended in the air that are so small that they can be inhaled. Sources of particulate matter include car emissions, oil and gas to heat homes, manufacturing processes, and power generation. Wildfire smoke is another source. Wildfire smoke particles vary in size, but about 90% are small enough to be called PM2.5—particulate matter in the air that has a diameter of 2.5 micrometers (μm) or less, also known as fine particles.

Studies show a strong link between breathing in particulate matter, especially PM2.5, and health risks. For many people, the symptoms of air pollution are similar to allergies, a cold, the flu, or COVID-19. So how can you figure out the likely cause of someone's symptoms? Probabilistic reasoning is a way of making predictions or drawing conclusions based on how likely something is to happen, especially when there is not enough clear data. In this activity, you will use probabilistic reasoning to figure out the most likely cause of a person's respiratory symptoms.

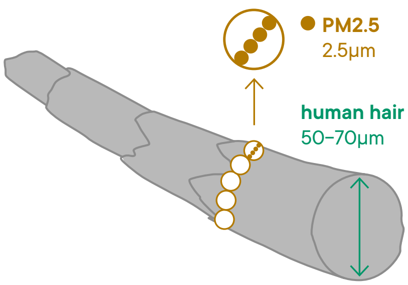


FIGURE 1.1
The size of PM2.5 particles is much smaller than a single human hair, which is typically 50–70 micrometers (μm) in diameter.

If you need to review the concept of probability, you will find a Science Review at the end of this activity.

CONCEPTUAL TOOLS

MATERIALS LIST

FOR EACH GROUP OF FOUR STUDENTS

- SET OF 4 STUDENT HEALTH CARDS
- SET OF 4 STUDENT FOLLOW-UP CARDS

FOR EACH PAIR OF STUDENTS

- SET OF COLORED PENCILS (2 DIFFERENT COLORS)

FOR EACH STUDENT

- 2 STUDENT SHEETS 1.1 "Analyzing Symptoms"
- STUDENT SHEET 1.2 "Symptom Chart"

PROCEDURE

1 With your group of four, read the following factual article.

INCREASE IN WILDFIRES MAY AFFECT RESPIRATORY HEALTH

Extreme wildfires are increasingly becoming a major cause of poor air quality in populated areas. Some wildfires occur naturally and create new habitats, stimulate new plant growth, and cycle nutrients. However, around the world, wildfire patterns are changing. Less than 10% of wildfires are now causing more than 90% of the total area burned each year. These events are known as extreme wildfires. In the last few years, extreme wildfires have occurred in many countries around the world, including Australia, Brazil, Canada, Russia, and the United States.

Children, teens, and the elderly are particularly vulnerable to poor air quality. One concern is the effect of particulate matter on those with asthma, a noninfectious condition that affects the lungs. Asthma can be controlled by taking medicine and avoiding triggers that can cause an attack such as air pollution, smoke, pet hair, and mold. Doctors suggest that the best thing to do when outdoor air quality is poor is to reduce your outdoor exposure.

2 Apply probabilistic reasoning to determine the most likely cause(s) of a person's respiratory symptoms. Your group will receive a set of 4 Student Health cards, and each pair will analyze 2 cards. Pick 1 card and read it aloud to your partner.

3 Work with your partner to identify the student's symptoms and record them on Student Sheet 1.1, "Analyzing Symptoms." As you work, your partner will also fill in their copy of Student Sheet 1.1.

4 Compare the student's symptoms to the symptoms on Student Sheet 1.2, "Symptom Chart."

a Identify the student's symptoms on Student Sheet 1.2, using a colored pencil to circle and/or underline symptoms. (You will use the other colored pencil for a second fictional student.)

b Determine the student's two most probable causes and explain your reasoning on Student Sheet 1.1. While you do not need to agree with your partner, remember to listen to and consider the ideas of others. If you disagree, explain why you disagree.

5 Brainstorm two questions to ask the student about their health to improve your diagnoses. Record those questions on Student Sheet 1.1.

6 Based on your diagnoses, work with your partner to recommend one or more actions for the student from the following list. Record your recommendations on Student Sheet 1.1.

- Rest and wait to see if symptoms improve.
- Ignore the symptoms and maintain routine activities.
- Stay indoors.
- Avoid strenuous outdoor activities such as sports.
- Isolate from other people as much as possible.
- Wear a face mask.
- Go to an emergency room as soon as possible.
- Make an appointment to see a doctor in a few days.
- Take an over-the-counter allergy medicine.
- See a doctor about getting an inhaler prescribed for asthma.
- Take an at-home COVID-19 test.
- Take an over-the-counter headache medicine.
- Take an over-the-counter cough suppressant.
- Take an over-the-counter fever reducer.
- Another course of action not listed here.

7 Collect a Student Follow-Up card for this fictional student. On Student Sheet 1.1, record what happened to your student.

8 Discuss with your partner whether the information from the Student Follow-Up card supported your recommended course(s) of action. Record your ideas on Student Sheet 1.1.

9 Revisit your diagnoses to determine if you would like to revise it. If you make revisions, record them on Student Sheet 1.1, along with your reasoning for the likelihood of each illness.

10 Scientific uncertainty is an understanding that there are limits to data and conclusions about the natural world, and additional data and/or investigations can lead to increased surety or new questions. Discuss and record how sure you are of your diagnosis, using a scale of 0–100%, where:

0% = there is no chance your diagnosis is correct
50% = your diagnosis is just as likely to be wrong as it is to be correct
100% = you are absolutely sure that your diagnosis is correct

11 Choose another Student Health card and repeat Steps 2–10. This time, have your partner read aloud the card. Each of you should use a different colored pencil to mark the symptoms on your copies of Student Sheet 1.2 for this second fictional student.

12 Share your findings with the other half of your group. Discuss how sure you were of your diagnoses and what additional information you would have liked to have known.

BUILD UNDERSTANDING

The Build Understanding and Connections to Everyday Life items are intended to guide your understanding. Some of these items may be discussed with a partner, be part of a class discussion, or require an individual written response. Your teacher will guide you as to how these items will be used in your class.

1. Doctors, like other scientists, try to identify and reduce sources of uncertainty in science.
   a What factors caused you to be uncertain about your diagnoses?
   b What factors caused you to reduce uncertainty in your diagnoses?
2. Examine the data in the following table.

TABLE 1.1
Evaluation of Over-the-Counter COVID-19 Tests, 2021

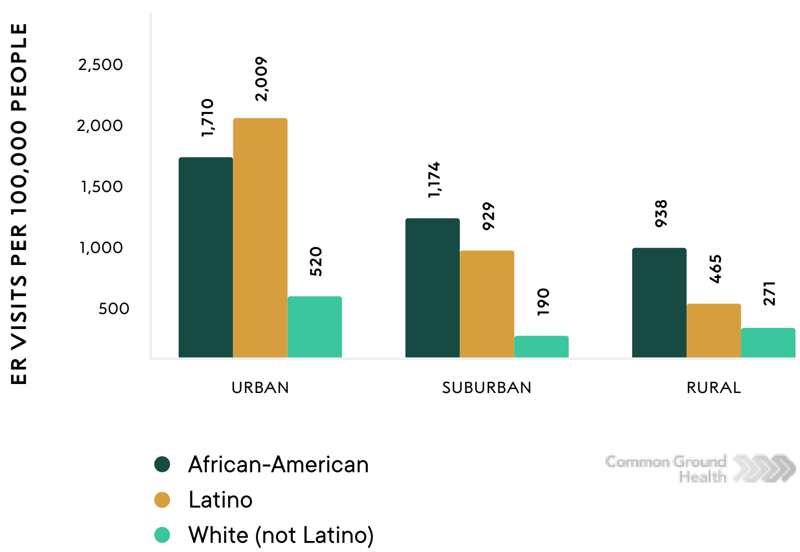
BRAND	PERCENT OF POSITIVE COVID-19 CASES CORRECTLY IDENTIFIED	PERCENT OF NEGATIVE COVID-19 CASES CORRECTLY IDENTIFIED
1	49.4%	100%
2	44.6%	100%
3	45.8%	97%
4	54.9%	100%

a A false positive is a type of error when something is incorrectly identified as present—for example, a positive COVID-19 test result when someone does not have COVID-19. A false negative is a type of error when something is incorrectly identified as absent but is actually present—for example, a negative COVID-19 test result when someone does have COVID-19. Based on Table 1.1, are over-the-counter COVID-19 tests more likely to result in a false positive or a false negative? Explain your reasoning.

b Imagine that you were feeling unwell and had symptoms similar to those of COVID-19. You take an over-the-counter COVID-19 test, and the test result is negative. Use probabilistic reasoning to explain whether or not you should go to a friend's birthday party.

3. Examine the graph in Figure 1.2, which provides data about emergency room visits for a three-year period.

FIGURE 1.2
Emergency Room (ER) Visits for Asthma
in New York Finger Lakes Region, 2014–2016



a What can you conclude about the likelihood of emergency room visits for asthma?

b Imagine your state has the funds to build one urgent care asthma center. Would you recommend they build it in an urban, suburban, or rural area? Support your answer with evidence and identify the trade-offs of your decision. A trade-off is a desirable outcome given up to gain another desirable outcome.

CONNECTIONS TO EVERYDAY LIFE

4. Which of the following are examples of probabilistic reasoning? Explain.
   - **a** estimating the chance of getting stuck in a traffic jam based on the time of day
   - **b** deciding on where to have dinner based on your favorite food and the cost of the meal
   - **c** a basketball player calculating the odds of making a shot based on their past performance and the current situation on the court
   - **d** a doctor considering the likelihood of a specific disease based on a patient's symptoms and test results
   - **e** selecting a concert to attend based on which concert venue is the closest to where you live
   - **f** figuring out your chances of getting a job offer based on your qualifications and the competition for the position
5. Choose one example of probabilistic reasoning from item 4 and describe one action you could take to reduce scientific uncertainty in that example.

EXTENSION

Learn how scientists handle uncertainty in science and use probabilistic reasoning by watching a 9-minute video narrated by Saul Perlmutter, who won the Nobel Prize in Physics in 2011. Consider the advantages and disadvantages of using probabilistic reasoning, as explained in the video.

KEY SCIENTIFIC TERMS

false negative
false positive
particulate matter (PM)
PM2.5
probabilistic reasoning
scientific uncertainty
trade-off

SCIENCE REVIEW

Probability

In this unit, you will be learning about probabilistic reasoning. You may also be familiar with the concept of probability, a mathematical measure of the likelihood that a specific event will happen. Probability can be expressed as a percentage. For example, a 25% chance of rain represents a 1 out of 4 (¼) chance that it will rain. The probability of an event depends on the number of successful outcomes divided by the number of total possible outcomes:

$$\text{probability of an event} = \frac{\text{number of successful outcomes}}{\text{total number of possible outcomes}}$$

Probability can be converted between fractions and percentages by using formulas:

FRACTION TO %

$$\frac{A}{B} \times 100 = C\%$$

EXAMPLE

$$\frac{4}{5} \times 100 = 80\%$$

% TO FRACTION

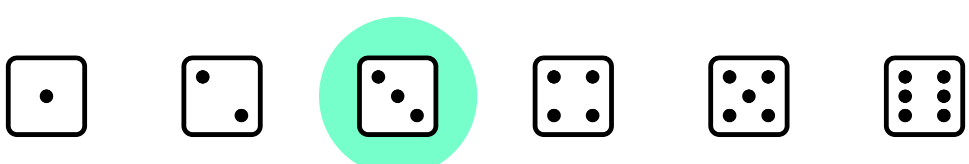
$$C\% = \frac{C}{100}$$

EXAMPLE

$$80\% = \frac{80}{100} = \frac{4}{5}$$

Consider a number cube. When you roll a number cube, there are 6 possible outcomes:

If you wanted to calculate the probability of rolling a 3, only one of the rolls would be a successful outcome:

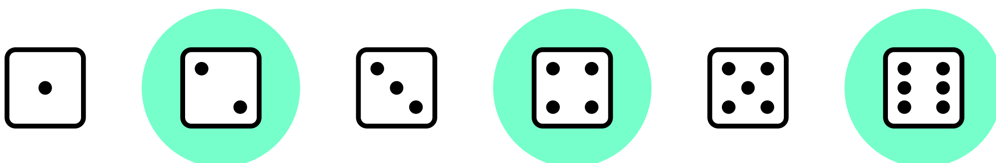


The calculation would be as follows:

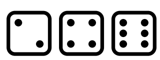
probability of rolling a  $= \dfrac{\text{1 successful outcome}}{\text{6 possible outcomes}} = \dfrac{1}{6}$

Expressed as a fraction, the chance of rolling a 3 is 1 out of 6 times, or 17%. But this does not guarantee that 1 out of every 6 rolls will be a 3! The chance of any given event occurring is independent of the previous events. So every roll is independent of the rolls you did before. It is sometimes only after many, many events that you will start to see the probability—for example, after calculating the average of many rolls.

Probability can be determined when there is more than one possible event that counts as a successful outcome. The following example shows the probability of rolling an even number on the cube. In this case, there are three possible events that count as successful outcomes:



So, the probability of rolling an even number is:

probability of rolling a  $= \dfrac{\text{3 successful outcome}}{\text{6 possible outcomes}} = \dfrac{3}{6} = \dfrac{1}{2}$

On average, you will roll an even number 1 out of 2 times, or 50%.

ACTIVITY 2

Signal and Noise

DATA ANALYSIS

Imagine trying to spot a single moth on the bark of a large tree. How could you find this signal among the environmental noise?

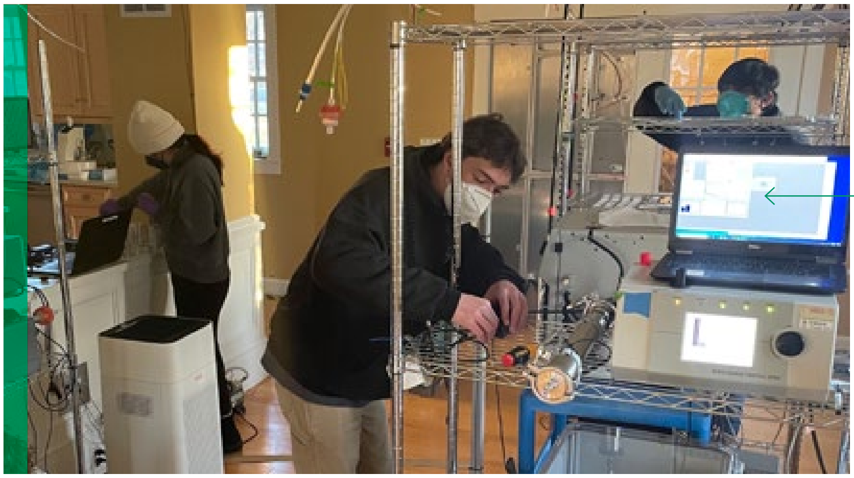
2 : SIGNAL AND NOISE

GUIDING QUESTION

How do you identify the meaningful data in a data set?

INTRODUCTION

What effect do wildfires have on indoor air quality? One group studying this is atmospheric chemists who research the chemistry of the air. Atmospheric chemists, like other scientists, try to separate the meaningful information about the phenomenon that is being investigated—the signal—from other factors that might interfere with it. These interfering factors are called noise. Noise is any information that hides, distracts from, or falsely resembles the meaningful information that is being investigated. For example, think of a search-and-find book where you need to find a cookie jar in a kitchen scene full of items. The cookie jar is the signal, and the other items are noise. While noise can increase uncertainty in science, probabilistic reasoning can help identify the signal being investigated. In this activity, you will act as an atmospheric chemist trying to separate the signal from the noise while studying data on wildfires and indoor air quality.



Researchers conduct experiments on indoor air quality at the Net Zero Energy Residential Test Facility in Gaithersburg, Maryland. The Chemical Assessment of Surface and Air (CASA) project aims to learn more about the chemical reactions affecting air and surfaces inside homes.

CONCEPTUAL TOOLS

SIGNAL & NOISE

MATERIALS LIST

FOR EACH GROUP OF FOUR STUDENTS

- SET OF COLORED PENCILS (3 DIFFERENT COLORS)

FOR EACH STUDENT

- STUDENT SHEET 2.1 “Outdoor vs. Indoor Air Quality Measurements”

PROCEDURE

1 With your group of four, read the following article.

IS COOKING WORSE THAN WILDFIRES FOR INDOOR AIR QUALITY?

When an extreme wildfire occurs, most people head indoors to avoid poor air quality outdoors. They may close windows, run an air purifier, or wear a face mask to avoid exposure to particulate matter and pollutants in the air. These actions may not be enough to prevent negative health effects such as asthma, dry eyes, and even lung disease. Yet most people are concerned about the air quality only when it’s bad outside. In many places, people spend more than 90% of their time indoors.

Scientists are now investigating factors that affect indoor air quality, including wildfires. They are also concerned that indoor air quality can become worse when people cook large meals or do a lot of frying.

2 In your group of four, examine the two graphs on Student Sheet 2.1, "Outdoor vs. Indoor Air Quality Measurements," and discuss what conclusions you can make about outdoor air quality vs. indoor air quality during a wildfire. The graphs show air quality measurements taken by both an outdoor and an indoor air sensor placed on opposite sides of the same wall at a home where wildfires are occurring nearby. Remember to listen to and consider the ideas of group members. If you disagree with other group members, explain why you disagree.

3 Use Student Sheet 2.1 to evaluate whether outside air quality during a wildfire affects indoor air quality.

   a Draw boxes vertically around sections of both graphs where you think poor indoor air quality could have been caused by poor outdoor air quality.

   b Make a hypothesis about whether outside air quality during a wildfire affects indoor air quality. Identify at least three dates (and approximate times) that provide evidence to support your hypothesis. Record your hypothesis and evidence in your science notebook.

4 Share your hypothesis and evidence with your group.

5 As an atmospheric scientist, you contact the residents of the home and ask questions about factors that could affect their indoor air quality. You find out that the following events occurred:

- A house fan was running near the indoor sensor on the evening of September 15.
- A neighbor was smoking outside near the outdoor sensor in the middle of the day on September 13 and on the afternoon of September 14.
- The family was cooking large meals indoors at midday September 15, the mornings of September 16 and September 18, and late at night on September 19.

On Student Sheet 2.1, use three different colored pencils (one for each type of event) to highlight the dates and approximate times on which these three different events occurred. Label each colored area with the event that occurred at that time.

6 Discuss with your group what effect, if any, each event appeared to have on indoor air quality.

7 Work with your group to identify the signal and the noise in determining whether outdoor air quality is affecting indoor air quality during a wildfire. Do this by labeling each vertical box on Student Sheet 2.1 as either signal or noise.

8 Revisit your hypothesis and make any needed changes.

9 WIth your group, brainstorm actions you could take to improve indoor air quality in your home.

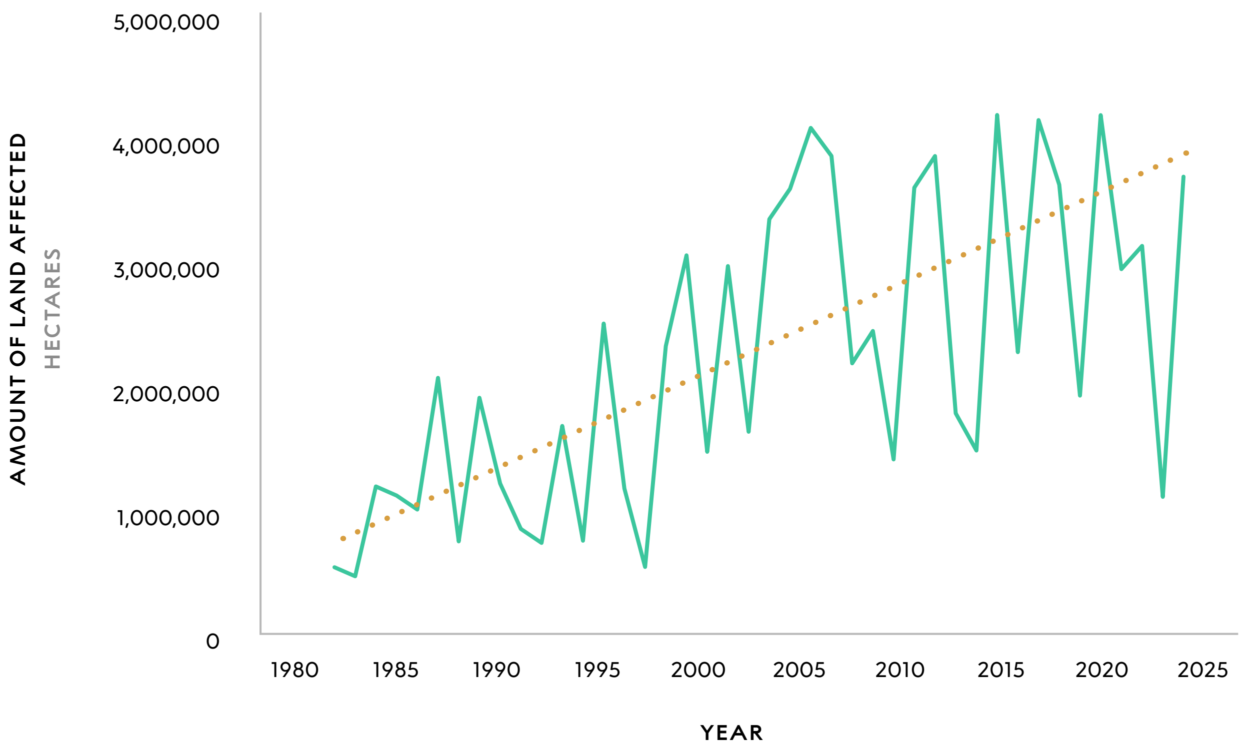
10 Share your ideas with the class.

BUILD UNDERSTANDING

1. Would you prefer to stay indoors or be outdoors on a day when there are extreme wildfires in your area and your family is cooking a large meal? Explain your reasoning and support your answer with data from this activity.

2. Figure 2.1 is a graph that shows the land area affected by wildfires in the United States since 1983. A hectare is a metric unit for measuring land area, equal to 2.47 acres or 10,000 square meters.

FIGURE 2.1
Land Area Affected by Wildfires in the United States, 1983–2024



a. What is the signal?

b. Describe what the noise looks like in the graphed data.

c. Based on the signal, what can you conclude about land area in the United States burned by wildfires over time?

d. How might this pattern be similar or different from the area in which you live?

CONNECTIONS TO EVERYDAY LIFE

3. You are watching a movie with friends. Another friend is joining you later.

   a Your dog is barking, making it difficult for you to hear what's happening in the movie. What is the signal and what is the noise?

   b You move your dog to another room. He later barks to let you know that someone is at the door. Is your dog's bark a signal or a noise?

   c Explain how identifying the signal vs. the noise depends on the information you are interested in.

4. You may have heard warnings about contaminated food, such as to avoid eating lettuce that has been recalled. In many cases, scientists gather data from individuals who fall sick to determine the common cause of their illness. The U.S. Centers for Disease Control and Prevention (CDC) uses this information and has estimated the annual number of foodborne illnesses in the United States to be 47.8 million cases. In the case of detecting foodborne illness:

   a What is the signal?

   b What is a possible source of noise in the data? (What other explanation[s] could there be for the data?)

   c For its estimate, the CDC accounted for scientific uncertainty and identified that the range of actual cases of foodborne illness in the United States per year may be as low as 28.7 million and as high as 71.1 million cases. Would you expect that reducing noise in this data would increase or decrease the estimated range of the data? Explain your reasoning.

KEY SCIENTIFIC TERMS

noise
signal

ACTIVITY 3

Scientific Uncertainty in Data

COMPUTER INVESTIGATION

Crowdsourced air sensors can easily be placed around homes and neighborhoods.

3 : SCIENTIFIC UNCERTAINTY IN DATA

GUIDING QUESTION

What are some sources of scientific uncertainty in data?

INTRODUCTION

In 2015, Utah resident Adrian Dybwad noticed wind blowing dust from a nearby gravel pit into the air. Curious about the effect this had on air quality, he bought air quality sensors. However, the sensors gave very different readings, leaving him unsure about the air quality. In science, the term *error* doesn't mean a mistake, as it does in everyday language. Scientific error is the difference between a measured or observed value and the true value of a quantity. The true value is the actual number that would be found if the measurement could be made without error.

Dybwad, an engineer, started experimenting with air quality sensors to reduce scientific error. He created more reliable sensors than those typically available for purchase and began sharing them with others. Dybwad went on to found PurpleAir, a website that tracks particulate matter data from sensors that people can buy and place in their communities. This type of data collection (by many people in a community) is called *crowdsourcing*. The data collected can help people make informed decisions about their exposure to particulate matter in the air.

In this activity, you will examine air quality data to better understand how close you can get to the true value and to look for sources of uncertainty in the measurements.

CONCEPTUAL TOOLS

PROBABILISTIC REASONING

SYSTEMATIC & RANDOM ERROR

FALSE POSITIVES & FALSE NEGATIVES

MATERIALS LIST

FOR EACH PAIR OF STUDENTS

- COMPUTER WITH INTERNET ACCESS

FOR EACH STUDENT

- STUDENT SHEET 3.1 "Analyzing Crowdsourced Air Quality Data"

PROCEDURE

PART A: CROWDSOURCED SENSOR DATA

1 Review the following Air Quality Index, which is a general guide to air quality based on several air pollutants, including particulate matter, ozone, nitrogen dioxide, sulfur dioxide, and carbon monoxide.

FIGURE 3.1
Air Quality Index (AQI)

AQI CATEGORY (COLOR)	INDEX VALUE	DESCRIPTION OF AIR QUALITY
Good (green)	0–50	Air quality is satisfactory, and air pollution poses little or no risk.
Moderate (yellow)	51–100	Air quality is acceptable. However, there may be a risk for some people, particularly those who are unusually sensitive* to air pollution.
Unhealthy for Sensitive* Groups (orange)	101–150	Members of sensitive* groups may experience health effects. The general public is less likely to be affected.
Unhealthy (red)	151–200	Some members of the general public may experience health effects. Members of sensitive* groups may experience more serious health effects.
Very Unhealthy (purple)	201–300	Health alert: The risk of health effects is increased for everyone.
Hazardous (maroon)	301 AND HIGHER	Health warning of emergency conditions: Everyone is more likely to be affected.

* According to the American Lung Association, sensitive groups include children under 18, adults over 65, people with chronic heart or lung disease, people who are pregnant, and people with diabetes. Adults who are active outdoors, including outdoor workers and frequent outdoor exercisers, can be considered sensitive because of prolonged exposure to outside air.

2 With your partner, explore the air quality of your state today by using more widely available crowdsourced air quality data, such as from the PurpleAir website.

3 On Student Sheet 3.1, “Analyzing Crowdsourced Air Quality Data,” record today’s date and the name of your state in the appropriate places in Table 1. Then work with your partner to complete the first row of Table 1.

a Record the range of PM2.5 measurements in your state today.

b Use this air quality data and the Air Quality Index to determine the general air quality category in your state today and then record today’s air quality and how you made your determination.

c Work with your partner to describe possible reason(s) for scientific uncertainty in the data and then record your ideas.

HINT: Think about reasons that the air quality data might not represent the true value and/or possible sources of scientific error in the data.

4 Work with your partner to identify an area of your state where there are *few* sensors and then complete the second row of Table 1.

a Record the name of the area and the available data.

b Discuss what you can conclude about the air quality in that area, using the Air Quality Index, and record today’s air quality and how you made your determination.

c Work with your partner to describe possible reason(s) for scientific uncertainty in the data and record your ideas.

5 Work with your partner to identify an area of your state that has *many* sensors and then complete the third row of Table 1.

a Record the name of the area and the range of data available in this area.

b Discuss what you can conclude about the air quality in that area, using the Air Quality Index, and record today’s air quality and explain how you made your determination.

c Work with your partner to describe possible reason(s) for scientific uncertainty in the data and record your ideas.

d Discuss whether you are more sure of your determination of air quality in the area with fewer sensors or more sensors. Explain your reasoning.

6 Work with your partner to gather more data from an area with many sensors and then complete the "Crowdsourced Data" column of Table 2 on your student sheet.

a Select and record five measurements.

b Identify and record any number(s) that differs significantly among the five measurements. With your partner, discuss possible explanations why this data point(s) is significantly different from the others.

c Calculate and record the average of the five measurements.

d Use the Air Quality Index (from Procedure Step 1) to determine the air quality for this area and record it.

e Compare your air quality findings based on the average of five data points with your determination from many sensors (Procedure Step 5b). Describe to your partner how your results are similar or different, which air quality determination is more likely to be closer to the true value, and why.

PART B: HIGHER-QUALITY SENSOR DATA

7 a Go to a site that provides data from higher-quality sensors, such as the Fire and Smoke Map at *AirNow*. *AirNow* is a partnership of federal, state, local, and tribal air quality agencies in the United States. It shows data from permanent and temporary air monitoring stations (which have higher-quality sensors) and crowdsourced sensors such as PurpleAir.

b If you are using *AirNow*, remove the data from sensors such as PurpleAir by clicking on the MAP SETTINGS button on the top right of the screen and then turning off the AIR SENSORS toggle (seen under the LAYERS bar).

Environmental Protection Agency (EPA) scientist Andrea Clemens (standing) examines an air quality sensor at an EPA research site in North Carolina.

8 Find the area on the map with many sensors that you investigated in Procedure Steps 5 and 6. Work with your partner to gather more data to complete the "Higher-Quality Data" column of Table 2 on your student sheet.

a Select up to five measurements from the available higher-quality sensors and record them. (NOTE: There may not be five high-quality sensors in the area you selected. If not, use the data available in the area.)

b Identify and record any number(s) that differs significantly among the five measurements. If there are no differing data points, record "none."

c Calculate and record the average of these measurements.

d Use this air quality data and the Air Quality Index (from Procedure Step 1) to determine the air quality category for this area and record it.

9 Compare your air quality findings from high-quality sensors such as *AirNow* with sensors such as PurpleAir. Describe to your partner:

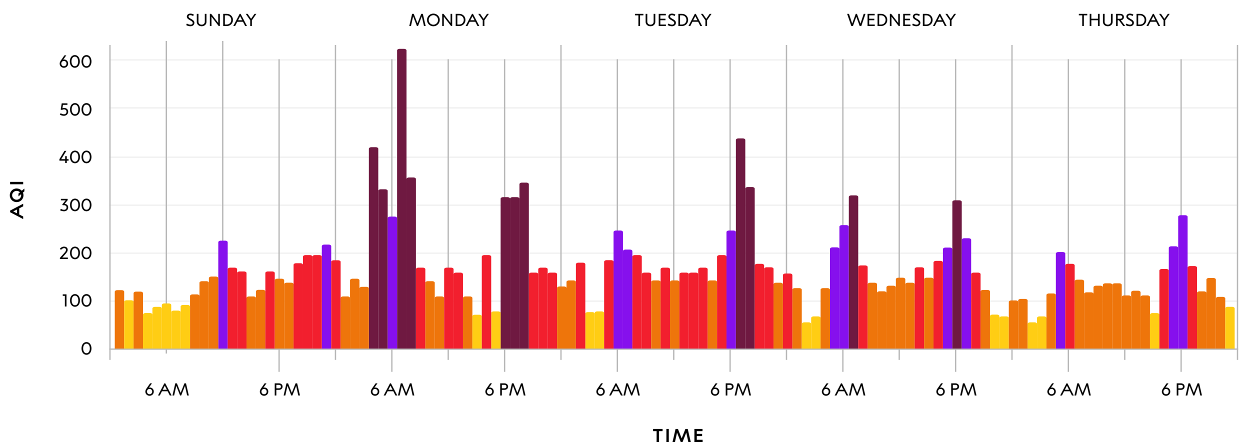
- how the data are similar or different.
- which data set is likely to be closer to the true value (have less scientific error) and why.
- how probabilistic reasoning was involved in making your determination of air quality.

BUILD UNDERSTANDING

1. Review your analysis of local air quality data from both sites.
    - **a** List two possible sources of scientific uncertainty in air quality data.
    - **b** Brainstorm how each source of scientific uncertainty could be reduced.

2. The graph in Figure 3.2 shows several days of hourly AQI measurements for the capital city of N'Djamena in the country of Chad. Explain what conclusions you can make based on the data in the graph. Refer to Figure 3.1, "Air Quality Index (AQI)" in Procedure Step 1 as needed. In your explanation, be sure to include the following:
    - **a** Describe what patterns you observe in the air quality over time.
    - **b** Explain what conclusions you can make about local air quality.
    - **c** Explain at least two possible sources of scientific uncertainty, including possible scientific errors, that may have affected the data.

FIGURE 3.2
Air Quality Index (AQI) Data for Five Days in N'Djamena, Chad



3. Based on the data in Figure 3.2, which day would be better to be outdoors: Monday or Tuesday? Support your answer with evidence.

CONNECTIONS TO EVERYDAY LIFE

4. The local Air Quality Index (AQI) on the day of your team's soccer semifinals is reported as 135. A sensor near your home shows an AQI of 100. The best player on your soccer team has asthma. Would you recommend that she play in the semifinal game? Support your answer with evidence and identify the trade-offs of your decision.

5. People rely on data such as air quality from scientific tools and technology to make decisions. Yet data from such technology can sometimes be inaccurate. In the case of an air quality sensor, is it worse to have a false positive or a false negative? Explain your reasoning.

EXTENSION

Use an online air quality map (such as one from PurpleAir) to look at data in other parts of the world over a period of several days. Look for any patterns in the data and develop a tentative explanation for what you observe. Brainstorm how you could further investigate your ideas while reducing scientific uncertainty.

KEY SCIENTIFIC TERMS

scientific error
true value

ACTIVITY 4

Reducing Error in Experimental Design

LABORATORY

Many factors affect air quality. Which factors do you think have the greatest impact on the air quality in your community?

4 : REDUCING ERROR IN EXPERIMENTAL DESIGN

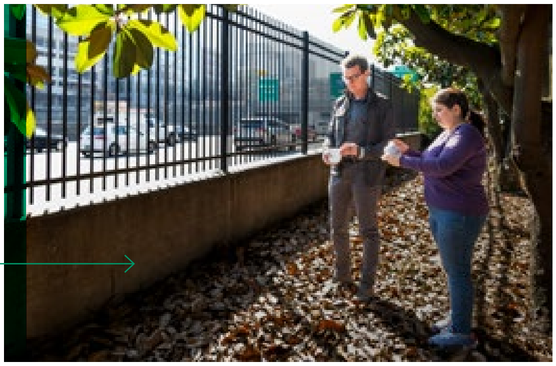
GUIDING QUESTION

How do you design a study to reduce scientific error?

INTRODUCTION

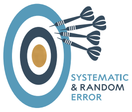
A 2024 study by Georgia State University (GSU) researchers found that planting trees and bushes near highways can help reduce air pollution from cars and trucks. Professor Roby Greenwald, one of the researchers, said, “Trees and bushes near roadways don’t solve the problem of air pollution, but they can help reduce the severity of the problem.” The study looked at air quality at five places along highways in Atlanta and compared them to similar places without plants. They found that areas with plants had 37% less soot and 7% fewer particles less than 0.1 micrometers in size, which are much smaller than what PM2.5 air sensors measure.

To reduce scientific error, researchers look at other factors that could affect the data. One important factor in the GSU study was the direction of the wind. The plants had the biggest effect on air quality when the wind was blowing from the highway toward the plants. In this activity, you will design an experiment to test air quality in your local area. You will also think about possible sources of scientific error and how to reduce them.



Professor Roby Greenwald (left) and graduate student Nadia Alhadidi (right) demonstrate air quality monitoring equipment. It was used to assess the role that trees and bushes play in reducing air pollution near highways.

CONCEPTUAL TOOLS



MATERIALS LIST

FOR EACH GROUP OF FOUR STUDENTS

- MICROSCOPE (OR STEREOSCOPE)
- 9 PETRI DISHES
- PETROLEUM JELLY
- TAPE
- 9 INDEX CARDS
- PERMANENT MARKER
- GRAPH PAPER

SAFETY NOTE

Label all materials clearly and monitor them over the course of the experiment to ensure that they do not pose a hazard to others.

PROCEDURE

1 You will collect data by placing petri dishes coated with petroleum jelly at a collection site. Particulate matter in the air will fall onto the open petri dish and stick to the petroleum jelly. After collecting your dishes, you will use a microscope to observe the particulate matter in the petri dish. Set up your experiment by:

  a washing your hands.

  b creating a control by having one person in your group use a clean finger to smear a very thin layer of petroleum jelly on the bottom of one petri dish.

  c setting up your own 2 petri dishes by using your finger to smear a very thin layer of petroleum jelly on the bottom of 2 petri dishes.

  d using a permanent marker to label the top of each dish, including the control, with the location where it will be placed and your group name.

  e taping the dishes closed for transport.

  f writing the following on an index card for each dish: *Science Experiment—Please do not move!*

FIGURE 4.1 Spreading petroleum jelly on a petri dish

2 Work with your group to design an experiment to test the levels of particulate matter in your local air. Remember to listen to and consider the ideas of other group members. If you disagree with others in your group, explain why you disagree.

Begin by discussing the places where you could set your petri dishes and what question you would investigate. For example, you may want to investigate potential differences in air quality near and far from a highway or major road, indoor vs. outdoor, homes with or without pets, distance from a heating/air conditioning vent, street vs. backyard, distance from trees and plants, and so on. Be sure to consider the weather forecast when planning outdoor data collection.

3 When designing your investigation, be sure to address the following questions:

- What is your hypothesis?
- How and where will you place your petri dishes?
- How and where will you establish your control?
- How will the data you collect help you to make a conclusion?

4 Work with your group to brainstorm aspects of your experimental design and data collection that could result in scientific error. Revise your experimental design to include any ideas you have to reduce these errors.

5 Record your hypothesis and your experimental design in your science notebook.

6 Obtain your teacher's approval for your investigation.

7 Conduct your investigation for three days by:

a taking your dishes and index cards to the data collection site(s).

b uncovering petri dishes when placed at the data collection site(s) and securing the index-card labels nearby.

8 After three days, tape the petri dishes closed and bring them back to your classroom.

9 In your science notebook, make a data table to record *quantitative* (number of particulates) and *qualitative* data (particulate color, any identifiable particulates such as dog hair or other particulates) for each petri dish, including your control.

10 Use a microscope to examine your petri dishes by first placing a petri dish on graph paper and tracing a circle around it (as shown in Figure 4.2). Place your petri dish with the graph paper underneath on the microscope to help you track and record the number and type of particulates in the dish. (If you have too many particulates to count them individually, see Figure 4.3 for additional guidance.)

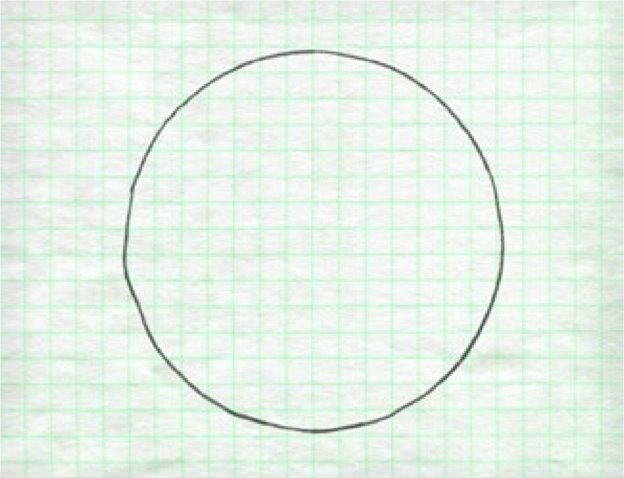


FIGURE 4.2
Tracing a Petri Dish on Graph Paper

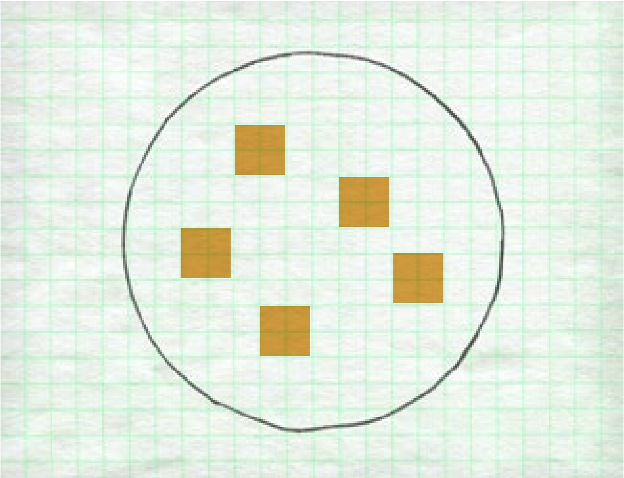


FIGURE 4.3a
Medium Number of Particulates

If you have too many particulates to count individually, count a smaller area of the petri dish to make an estimate of the total number of particulates. Do this by creating a template on the graph paper by marking a few squares on the graph paper (similar to the examples in Figures 4.3a and 4.3b). Then, place your petri dish with your template on the microscope and count only particulates in the squares.

NOTE: Use only one template (either 4.3a or 4.3b) to count the particulates in every petri dish of your group, including the control.

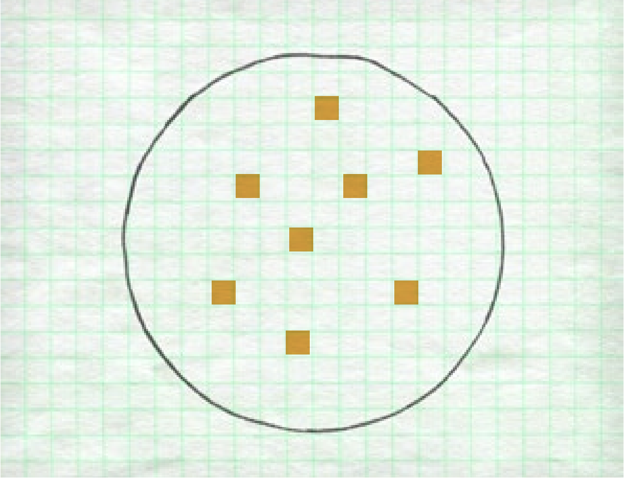


FIGURE 4.3b
High Number of Particulates

11 With your group, share and record data from all the petri dishes, including the control.

12 With your group, discuss your results and the conclusions you can make from your data. In your analysis, consider how your data does or does not support your hypothesis. Record all your findings in your science notebook.

13 Think about the role of scientific error in your experiment.

  a Discuss any possible sources of scientific error with your group and how those may have affected your results.

  b Brainstorm ways in which your experiment could have been improved by identifying ways to reduce random errors and/or systematic errors in your experimental design.

    - Random error is a difference between an observed and true value that has no consistent pattern and is caused by chance and/or unpredictable factors.
    - Systematic error is a difference between an observed and true value in a consistent direction, often caused by experimental equipment or design.

14 Share your results with your class.

BUILD UNDERSTANDING

① Review your experimental design and your results.

  a What were possible sources of systematic error in your experiment?

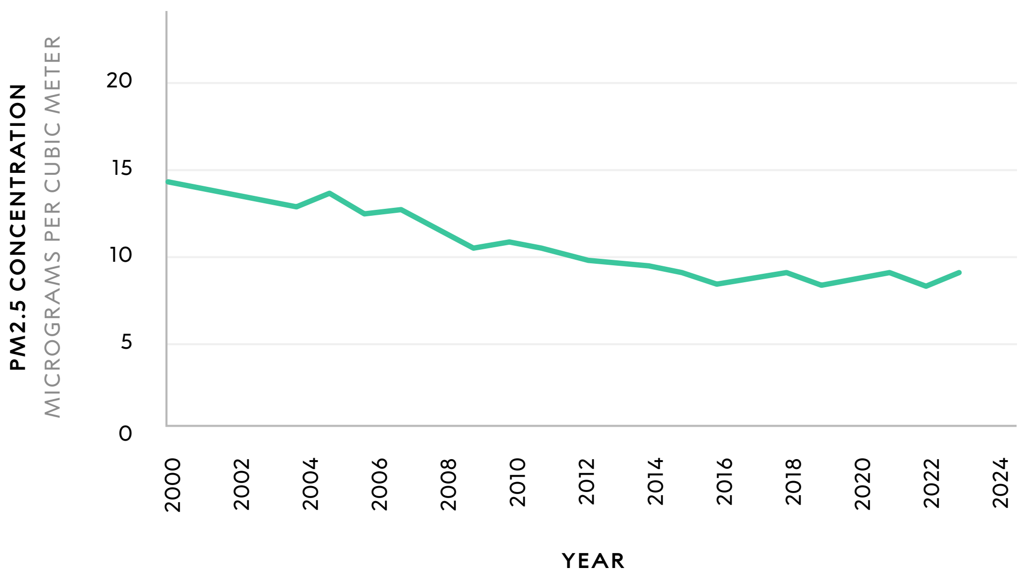
  b How did you address these possible sources of systematic error in your experimental design?

  c What possible sources of systematic error did you **not** address in your experimental design? Given more time and resources, how could you address them?

You can work with your partner to reduce potential sources of error when examining your samples.

② The graph in Figure 4.4 shows annual average PM2.5 air quality in the United States over a 24-year period. PM2.5 levels have decreased by 37% during that time. A person living near an area of frequent wildfires complains that their PM2.5 levels have increased during the same period.

FIGURE 4.4
PM2.5 Air Quality in the United States, 2000–2023



a Explain how this could be true.

b Figure 4.4 is from the U.S. Environmental Protection Agency (EPA), which addresses systematic and random error in its data. Explain one possible systematic error that could affect such data and how it could be addressed to reduce scientific uncertainty.

CONNECTIONS TO EVERYDAY LIFE

3. Like a scientific procedure, a recipe provides a list of steps to follow in order to produce an intended outcome, such as a batch of cookies. Imagine you are in a cooking class where eight groups each baked a batch of chocolate chip cookies using the same recipe. Some of the cookies came out chewy and thin, while others were thick and dry. What are some possible sources of (a) random error and (b) systematic error?

4. Janeen wants to improve her running speed, so she decides to experiment by running the same 1-mile route every day for a week to see if she gets faster. By the end of the week, her mile time had improved by 30 seconds. The following things happened during her experiment. Explain whether they are related to random error or systematic error and how each might have affected her results.

   a. Janeen starts her stopwatch a little bit early because she needs to secure her phone in her pocket before she starts running.

   b. Midway through the week, Janeen starts drinking an electrolyte drink 20 minutes before her run.

   c. Each time she runs, her speed varies slightly due to factors such as how much energy she has, random distractions (like a car honking), or even slight changes in the weather. Some days, she feels tired and runs a little slower.

5. How might you redesign Janeen's experiment from item 4 to reduce sources of scientific uncertainty in her data?

EXTENSION

One way to reduce scientific error is to gather data by using different experimental techniques. Look online for different experimental approaches for measuring air quality. For example, you could evaluate the visibility of a local city, as shown in the following images. Design and conduct a different experiment to measure your local air quality. Compare the results with the results from this activity. Explain which approach is less likely to have random errors and/or systematic errors and why.

The view of Lahore, Pakistan, on a good visibility day (left) and a poor visibility day (right) is affected by the local air quality.

KEY SCIENTIFIC TERMS

random error

systematic error

ACTIVITY 5

Addressing Uncertainty in Science

READING

Environmental Protection Agency researcher Megan MacDonald checks an air quality monitoring system in an industrial district in Louisville, Kentucky. Her research focuses on understanding how industrial sites contribute to air pollution emissions.

5 : ADDRESSING UNCERTAINTY IN SCIENCE

GUIDING QUESTION

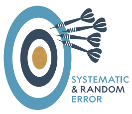
How do scientists reduce uncertainty in science?

INTRODUCTION

There can be scientific uncertainty when studying how wildfires affect health. Many studies average air quality over time (such as a year) or focus on short-term effects (such as how a week of wildfire smoke affects the number of emergency room visits). In 2024, researchers took a new approach to study the long-term effects of wildfire smoke. They looked at factors such as wildfire intensity, how often wildfires happen, and how long the smoke lasts. Rachel Morello-Frosch, a professor at the University of California at Berkeley, said, “Now that wildfires are...increasing in frequency and intensity, we can’t look at them one at a time.” The researchers found that Indigenous communities in California were exposed to 1.7 times more wildfire smoke from 2006 to 2020 than expected. This information can be used to develop new studies or to inform public policy. In this activity, you will read about one of the first large studies on the health effects of poor air quality.

The smoke from a wildfire can travel hundreds or even thousands of kilometers, depending on wind and other factors.

CONCEPTUAL TOOLS



MATERIALS LIST

FOR EACH STUDENT

- STUDENT SHEET 5.1 "Anticipation Guide: Scientific Uncertainty"
- STUDENT SHEET 5.2 "DART: Examples of Scientific Uncertainty"

PROCEDURE

1. Complete the "Before" column on Student Sheet 5.1, "Anticipation Guide: Scientific Uncertainty," to prepare for the following reading.
2. Read the following text to learn how scientists use different methods to manage uncertainty in science.
3. As you read, fill out Student Sheet 5.2, " DART: Examples of Scientific Uncertainty." After reading, compare your ideas with another student's ideas.
4. Complete the "After" column on Student Sheet 5.1. Be sure to think about information from the reading.

READING

Air Quality and Human Health: The Harvard Six Cities Study

In the 1970s, a brown haze was a familiar sight over many cities around the world. In some places, the air quality was so poor that people had to drive with their headlights on during the daytime. Scientists began investigating the effects of air pollution on human health. They started to find links between poor air quality and increased rates of asthma, lung cancer, cardiovascular disease, and diabetes.

Steubenville, Ohio, was one of the cities studied in the Harvard Six Cities Study.

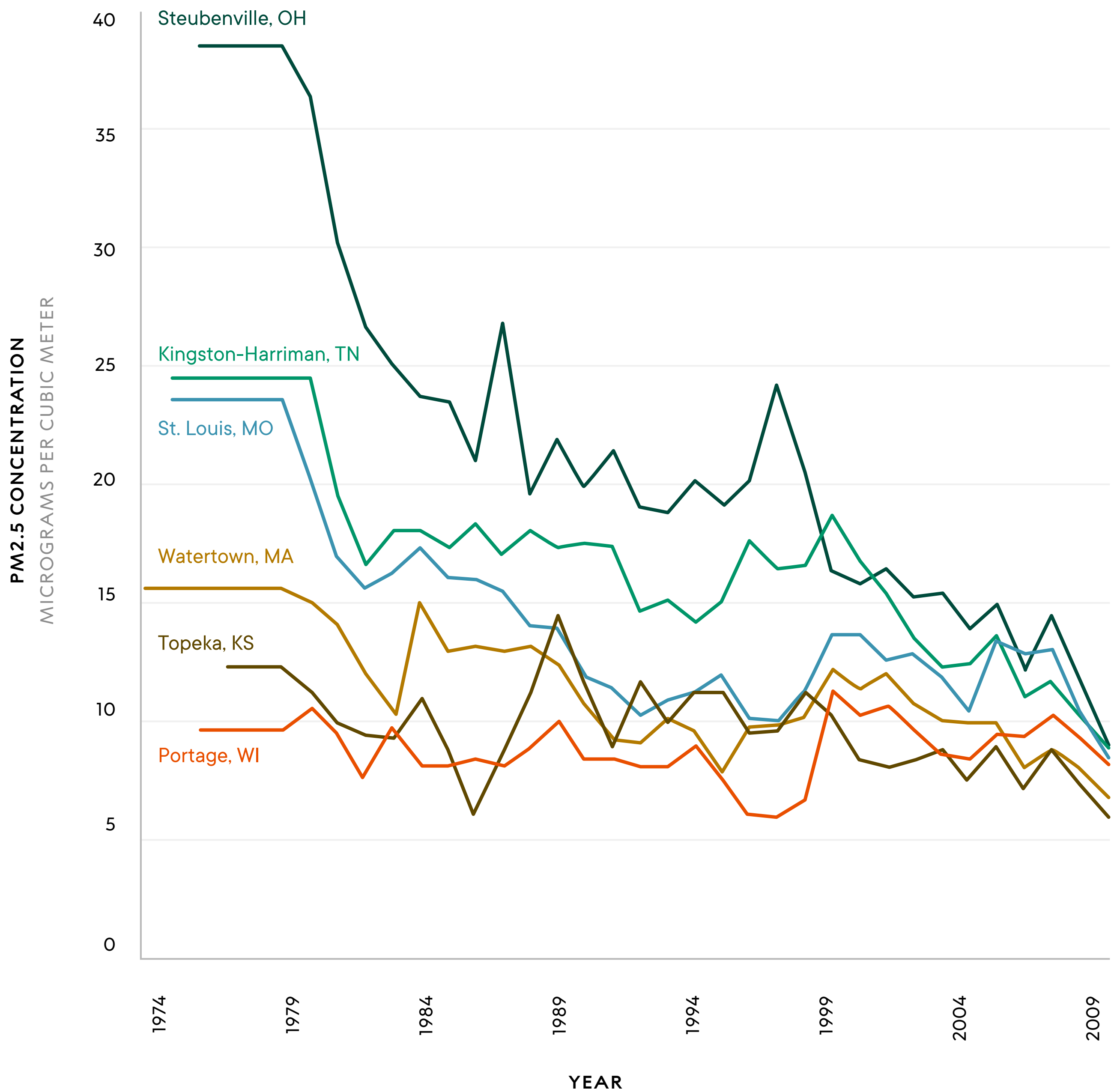
Dr. Douglas Dockery of Harvard University and his team wanted to know if poor air quality was shortening people's lives. Their study, called the Harvard Six Cities study, was published in 1993 and became one of the most important environmental health studies at the time. Between 1974 and 1989, the team collected health and lifestyle data from over 8,000 adults in six U.S. cities. They also gathered air quality data from each city. Since studying the effects of air pollution on health was challenging, the researchers needed to identify and address as many sources of scientific uncertainty as possible. This would allow them to reduce uncertainty in their conclusions once their data collection was complete.

Random Error

Most scientific research is complicated because data can have many sources of scientific uncertainty. Scientists try to identify and reduce these uncertainties, but they can never fully eliminate them. In the Six Cities study, researchers worked to minimize errors in their experimental design and measurements. For instance, air pollution levels can change from hour to hour or day to day because they are influenced by factors such as weather, wind, and industrial activity. These changes can cause random error. For example, someone using a leaf blower might carry particulate matter away from a sensor, leading to air quality readings that seem better than they really are. To reduce this kind of error, the Six Cities study averaged pollution levels over an entire year.

Scientific uncertainty from random error can also be caused by not having enough data. For instance, refer to Figure 5.1 and imagine trying to determine air quality in Steubenville, Ohio, based on just one year, such as 1996. How certain could you be that Steubenville had the highest pollution levels compared to the other cities from just one year of data? In the Six Cities study, researchers collected data from all six cities over many years. Collecting more data reduces uncertainty and makes conclusions more reliable because it increases the chance that the average of the data reflects the true value. By taking many measurements and averaging them, scientists can reduce the effects of random error. This lets them have less uncertainty in their conclusions.

FIGURE 5.1
Average Air Pollution Levels Per Year in the Six Cities , 1974–2009



Systematic Error

The Six Cities researchers also had to consider systematic error. Systematic errors can lead to inaccurate results, no matter how much data is collected or averaged. The only way to reduce a systematic error is to eliminate the source of the error. For example, if air quality sensors always measure PM2.5 levels lower than the true value, the average air pollution readings would also be too low. This would be a systematic error. Scientists regularly test their equipment and compare results from different measurement tools to reduce this type of systematic error.

Another possible source of systematic error was the placement of air quality sensors. For example, if sensors were placed in areas with cleaner air than where participants lived, the study would underestimate pollution exposure. Scientists plan for these types of errors during their study design. Sometimes they even need to redesign their experiments to avoid systematic errors.

Today, the air quality in the six cities, such as St. Louis, Missouri (pictured), has improved.

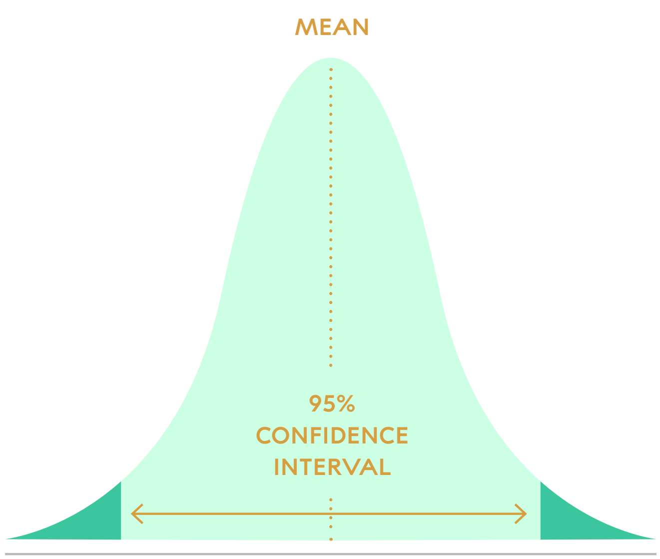
When researchers first analyzed data from the Harvard Six Cities study, they found that people in the least polluted city had a 51% higher survival rate compared to those in the most polluted city. They suspected that one of the biggest sources of systematic error came from confounds related to the health of participants. A confound is a factor that can distort or hide the relationship between two variables being investigated in a study. For example, differences in smoking habits or diet could have a stronger effect on health than air pollution. Researchers adjusted for these confounds in their data analysis. They still found a 26% difference in survival rates between people living in the least and most polluted cities. This led them to conclude that air pollution has a significant impact on life expectancy.

Communicating Scientific Uncertainty

In this unit, you have been communicating how sure you are of your data by selecting a percentage between 0 and 100. One way scientists calculate and communicate uncertainty in data is by using confidence levels and confidence intervals. In everyday life, the term *confidence* refers to feeling capable of handling a situation and achieving a goal; in science, the term *confidence* refers to a way of communicating scientific uncertainty. Scientists use a confidence level—which is a statistical measure of the probability that the true value is within a specified range. A confidence level tells you how often, in many repeated studies, the range will include the true value.

The confidence interval is calculated based on the confidence level a scientist wants to test for. A confidence interval is the range of data expected to contain the true value. Confidence intervals are probabilistic and predict how accurately your findings from a small sample apply to the larger data set. Scientists use it to estimate things such as the average height of students or the average test score in a school. Refer to Figure 5.2, which shows a confidence interval on a graph. As the sample size increases, the range of interval values narrows, meaning that you know the average with much more accuracy than you do from a smaller sample. The Six Cities study had a 95% confidence interval. That means that if the researchers were to conduct the same study 100 times, they would expect the true value to fall within this confidence interval 95 out of those 100 times.

FIGURE 5.2
A 95% Confidence Interval



Applying Science to Decision-Making

When the Six Cities study results were released, the government quickly pushed for stronger air quality standards to protect public health. However, many industries that contributed to air pollution were skeptical of the findings. They questioned whether tougher regulations were really needed. In response, many researchers reanalyzed the data and carried out additional studies to test the results. Over time, other studies confirmed the findings of the Six Cities study, including one involving over 60 million people. This additional research helped reduce scientific uncertainty in the study's conclusions.

Reducing scientific uncertainty has enabled scientists to provide clearer evidence for policymakers. The Six Cities study and the research that followed led to stricter PM2.5 air pollution laws in many countries. This helped to cut global death rates from air pollution nearly in half since 1990. Science continues to play an important role in providing data about air quality.

Face masks can be worn to reduce inhalation of particulate matter.

BUILD UNDERSTANDING

1. Researchers want to determine if reducing indoor air pollution can affect the severity of flu symptoms. The first 100 volunteers to sign up are assigned to a control group and given a device that circulates indoor air. The next 100 volunteers are assigned to an experimental group and given a device that filters indoor air pollutants before circulating the air. The researchers track the severity of flu symptoms in each group and record the number of days it takes each person to recover. At the end of the study, the researchers measure the overall health of each patient.

   Identify at least two scientific errors in the design of this study and explain how to improve the design to correct for these errors.

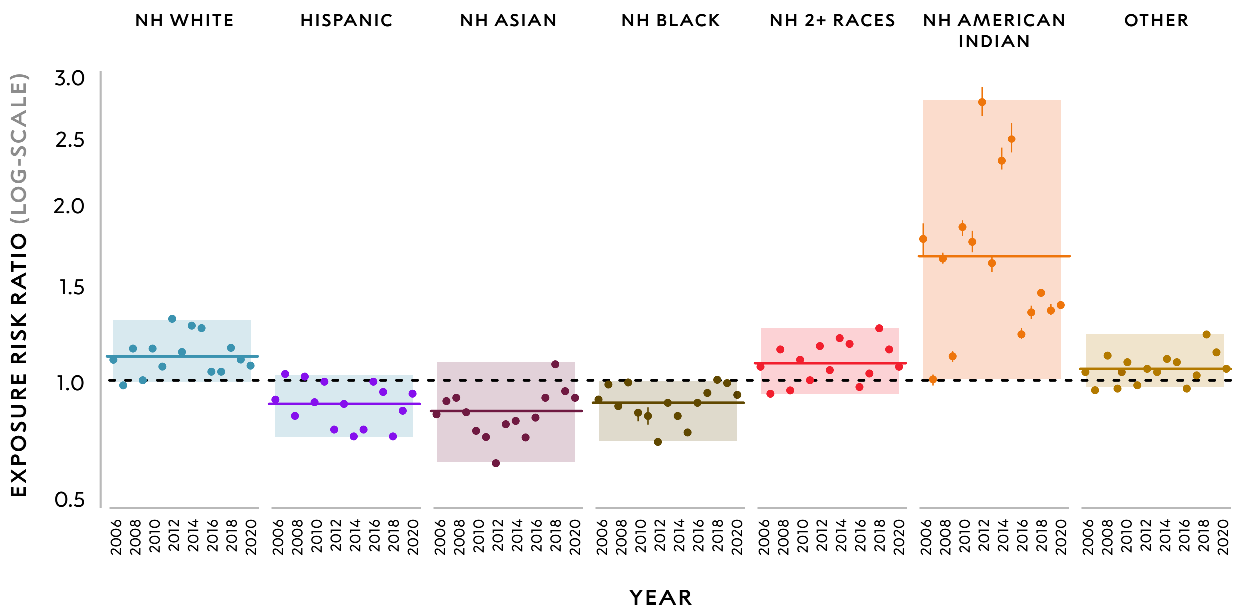
2. Figure 5.3 shows the average wildfire smoke exposure risk for different ethnic groups in California. The shaded areas of the graph represent a 95% confidence interval for each group's data set.

   a. Explain what a 95% confidence interval means in terms of this specific data.

   b. What conclusions can you make about the data?

   c. The graph for the Native American population (labeled on the graph as "Non-Hispanic American Indian") has the largest confidence interval. How does this affect the amount of uncertainty you have in the conclusions you make about the graph?

FIGURE 5.3
Wildfire Smoke Exposure, in California, 2006–2020



CONNECTIONS TO EVERYDAY LIFE

3. Johan wants to see if drinking more water will help clear up his acne. He decides to experiment by drinking 8 glasses of water every day for a month. He tracks his skin's appearance by taking pictures daily and noting the number of pimples. By the end of the month, Johan observes a lot less acne than at the beginning of the month.

   The following things happened during his experiment. Explain whether each is related to random error, a systematic error due to equipment or experiment design, or a systematic error due to a confound. Explain how each might have affected his results.

   - **a** Changes in weather, schedule, and diet resulted in several days when Johan drank more than or less than 8 glasses of water in a day.
   - **b** Johan didn't always use the same glass when he drank the 8 glasses of water. On some days, he used a smaller glass.
   - **c** After 2 weeks, Johan ran out of face soap, so he bought a different brand of face soap.

4. How would you redesign Johan's experiment from item 3 to reduce sources of scientific uncertainty in his data?

5. American physicist and Nobel Laureate Richard Feynman said, "When a scientist…has a hunch as to what the result is, he [or she] is uncertain. And when he [or she] is pretty…sure of what the result is going to be, he [or she] is still in some doubt. Scientific knowledge is a body of statements of varying degrees of certainty—some most unsure, some nearly sure, but none absolutely certain."

   Discuss with your class how uncertainty in science can sometimes mislead people into thinking that science cannot provide trustworthy information. How would you address this issue when talking to someone who says that science does not provide true information?

EXTENSION

Are you interested in ways to improve air quality and health in your community? In 2018, a medical professor named Aruni Bhatnagar studied whether planting trees could help. He focused on neighborhoods in Louisville, Kentucky, where there were fewer trees compared to other parts of the city. These neighborhoods also had a highway running through them. Nearly 8,000 trees and shrubs were planted, and health data was collected from almost 500 residents.

Before (left) and after (right) image of an area where trees were planted along Watterson Expressway/Interstate 264 in Louisville, Kentucky, for the Green Heart Louisville Project.

In 2024, the researchers found that people living in the greener areas had 13%–20% lower levels of inflammation biomarkers, which are linked to heart disease risk. You can learn more about this project by visiting the Green Heart Louisville Project's Heal Study website. Then, brainstorm ways you could help improve and study air quality in your own community. Consider ways to manage scientific uncertainty in your design.

KEY SCIENTIFIC TERMS

confidence interval
confidence level
confound

ACTIVITY 6

Quantifying Scientific Uncertainty

COMPUTER SIMULATION

Motorbike taxis, also called boda bodas, are a common source of transportation in Uganda. Researchers have found a way to attach air monitoring sensors to them in order to gather air quality data from around the city.

6: QUANTIFYING SCIENTIFIC UNCERTAINTY

GUIDING QUESTION

How can you reduce random errors and systematic errors in data?

INTRODUCTION

How can you reduce scientific uncertainty when investigating air quality? In 2022, a team of Ugandan researchers led by Engineer Bainomugisha developed an air monitoring technology that predicts pollution patterns. The technology was designed for local conditions in the country's capital, Kampala. Because it is solar powered, it can be installed in areas not connected to the energy grid. Air monitoring sensors are placed on top of buildings and on motorbike taxis. The motorbike taxis are driven around the city and provide air quality data from multiple areas. This reduces the likelihood of systematic error in the air quality readings. The researchers then use artificial intelligence (AI) software to review real-time data and predict pollution levels. The resulting information is used to develop clean-air policies and local health policies. In this activity, you will examine how random errors and systematic errors can affect data and ways in which these errors can be addressed.

Professor Engineer Bainomugisha and student Priscilla Adong install air monitoring technology in Kampala, Uganda in 2019.

> If you need to review the concepts of reliability, accuracy, precision, and validation, you will find a Science Review at the end of this activity.

CONCEPTUAL TOOLS

PROBABILISTIC REASONING

SYSTEMATIC & RANDOM ERROR

SIGNAL & NOISE

MATERIALS LIST

FOR EACH PAIR OF STUDENTS

- COMPUTER WITH INTERNET ACCESS

FOR EACH STUDENT

- STUDENT SHEET 6.1 "Testing Air Quality Sensors"

PROCEDURE

PART A: TESTING A PHOTOMETER

1 Read the following fictional scenario.

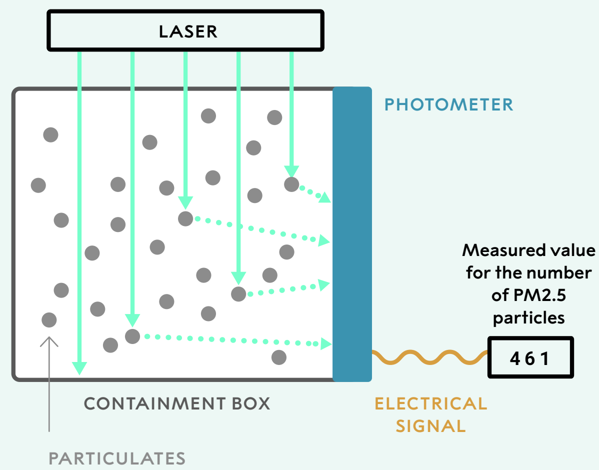
NEW STARTUP WANTS TO CLEAR THE AIR

A new startup company called Blase-Air has developed technology to improve air quality sensors. The technology is a new version of a light sensor called a photometer that is used inside air quality sensors, shown in Figure 6.1. Although it is still being tested, company engineers claim that the new photometer is more accurate than the old one.

Based on the test results, Blase-Air will decide whether they should use the new photometer in place of the older one in their air quality sensors.

The sensor works by flashing a laser into an air sample. The rays of the laser reflect off of the particulates and are collected by a photometer. The photometer converts the scattered laser signals into electrical pulses that are then used to calculate the number of PM2.5 particulates in the air sample. The device takes this measurement multiple times on a single air sample and provides an average of the measurements.

FIGURE 6.1
Air Quality Sensor



2 You and your partner will play the role of the team that is testing the new air quality sensor for the Blase-Air company. Follow your teacher's instructions for accessing the Scientific Uncertainty Simulation. This computer simulation contains the following three experiments:

EXPERIMENT 1 Investigate how well the original photometer (Photometer 1) measures PM2.5 levels.

EXPERIMENT 2 Investigate whether an appliance running next to Photometer 1 affects the data from Photometer 1 and, if so, how.

EXPERIMENT 3 Compare data from Photometer 1 to the new photometer (Photometer 2).

3 Conduct 5 trials of Experiment 1. Observe how the graphs change as you conduct each trial.

4 Select the SHOW TRUE VALUE and SHOW ERROR BARS buttons. In Figure 6.2, the true value refers to the actual number of PM2.5 particulates present in the air sample being tested. Error bars are a visual representation of the amount of uncertainty in a measurement, presented as a ± (plus or minus) range of numbers that fall above and below the average. In this simulation, the error bars were calculated using a 95% confidence interval. This means that if the experiment was repeated 100 times, the true value would be within the range of the error bars 95 times.

In Table 1 on Student Sheet 6.1, "Testing Air Quality Sensors," record the average, the error bar range, and the true value from your 5 trials.

FIGURE 6.2
Sample Graphs from Scientific Uncertainty Simulation

5 Graph 1, “Average of All Trials, Photometer 1,” plots all the individual data points vertically, while Graph 2, “Current Average, Photometer 1,” shows how the average of the data changes with more trials. Compare the data in your Graphs 1 and 2 from the simulation by discussing the following with your partner:

- your observations of each graph.
- what Graphs 1 and 2 might look like if there were a) no error, b) random error, and c) systematic error.

6 Conduct more trials until the error bar is ± 3. In the bottom row of Table 1 on the student sheet, record the number of trials it takes to get to an error bar range of ± 3, along with the average and the true value.

HINT: If you hover over the average line, you can see the average for any trial in the past. When you have SHOW ERROR BARS toggled on, you can see how the error bar range changes as you add more data points.

7 Use the results of Experiment 1 to discuss the following with your partner:

- how the number of trials affected the true value, the average, and the error bars.
- whether conducting more trials makes your average more accurate (closer to the true value) and why or why not.

8 Proceed to Experiment 2. Repeat Steps 3–6 for Experiment 2 and record your data in Table 2 on the student sheet.

9 Select the SHOW DATA FROM EXPERIMENT 1 button to see the data from Experiment 1 (when there was no appliance on). With your partner, determine whether the electric field of the appliance affected the measurements by:

a comparing the size of the error bars for Experiment 1 (no appliance on) and Experiment 2 (appliance on).

b comparing the averages for both Experiment 1 (no appliance on) and Experiment 2 (appliance on) to the true value.

10 Discuss with your partner what type(s) of error (systematic, random, or both) are likely to be occurring in Experiments 1 and 2. Be sure to explain your reasoning.

PART B: COMPARING TWO PHOTOMETERS

11 Proceed to Experiment 3 to investigate how well the air quality sensor functions with the old photometer (Photometer 1) compared with the new model (Photometer 2).

  a Repeat Steps 3, 4, and 6 for Experiment 3 and record your data in Table 3 on the student sheet. Make sure to record data for both Photometer 1 and Photometer 2.

  HINT: It may be helpful to have one partner watch and record data for Photometer 1 while the other watches and records data for Photometer 2.

  b Examine both graph types by selecting the CHANGE GRAPH TYPE button at the top of the screen to see how the average changed over time.

12 Use the data from Experiment 3 to determine which photometer:

  a has an average measurement that is most accurate (closest to the true value).

  b had less scientific uncertainty in the data.

  HINT: For each photometer, consider the size of the error bars as well as how spread out the data points are.

  c required fewer trials for the average to reach error bars within ± 3 of the true value.

13 As a class, evaluate the claim that the new Photometer 2 is more accurate than the old Photometer 1. Decide which photometer Blase-Air should use in their air quality sensor design.

BUILD UNDERSTANDING

1. Consider how the data changed during the experiments in the computer simulation. How does conducting more trials affect how close the average is to the true value when there is:

  a random error?
  b systematic error?

2. Error bars play an important role in communicating about scientific uncertainty in data. Describe how the size of the error bars can affect scientific uncertainty in a conclusion. Include an example from the computer simulation in your response.

3. Think back to the concepts of signal and noise from Activity 2: Signal and Noise. Explain how the presentation of data in the computer simulation helped differentiate the signal from the noise.

④ Figure 6.3 shows air quality data from five sensors in the same location. Each line represents a different sensor.

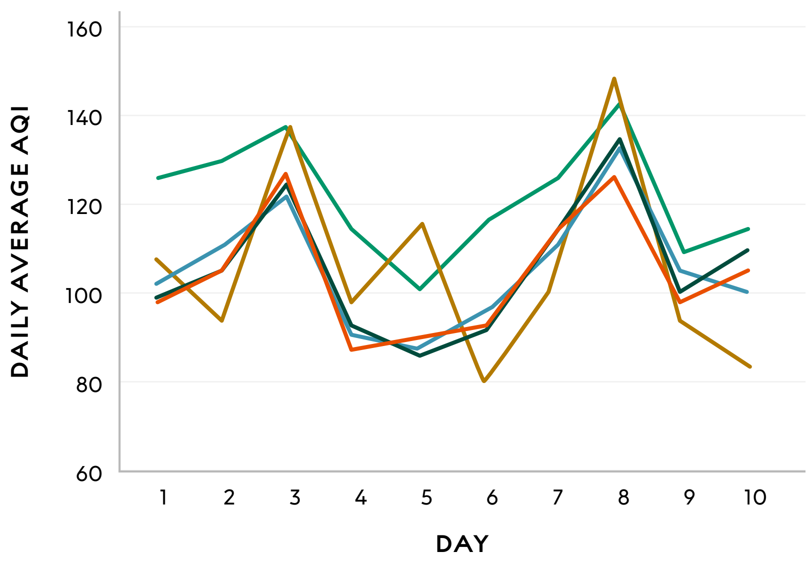


FIGURE 6.3
Data from Five Air Quality Index (AQI) Sensors

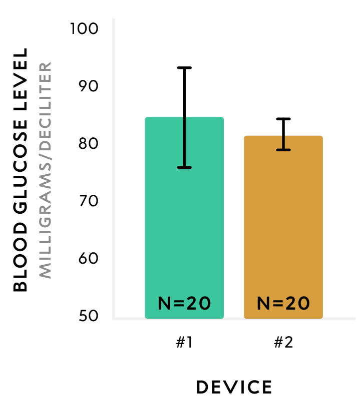
- **a** Which sensor is most likely to have a systematic error? Explain your reasoning.
- **b** Which sensor appears to have random error? Explain your reasoning.

CONNECTIONS TO EVERYDAY LIFE

⑤ Glucose is a type of sugar that your body uses for energy. Imagine that you read about a study comparing two different glucose meters for measuring glucose levels in the blood of patients with diabetes. Each glucose meter measured the same blood sample 20 times (N = 20). Figure 6.4 shows glucose level averages, shown with error bars, that were calculated for each glucose meter.

- **a** What can you determine about the uncertainty of the data collected by each device?
- **b** Can you be sure which device is measuring blood glucose levels closer to the true value? Explain why or why not.

FIGURE 6.4
Comparing Glucose Meters



KEY SCIENTIFIC TERM

error bars

SCIENCE REVIEW

Reliability, Accuracy, Precision, and Validation

Scientific explanations depend on relevant, accurate, and reliable data. Data is considered *reliable* if it is able to be reproduced consistently. For example, if an air sensor near you measures the AQI three different times in immediate succession and all three times it measures an AQI score of 32, then the sensor is very reliable. If a sensor measured an air sample three times and read 32, 59, and 67, then the sensor is not reliable. *Accuracy* is the closeness of a measured value to a standard or true value. For example, if your crowdsourced air sensor at home measures the same AQI as a nearby government air monitoring station that is higher quality and carefully calibrated, then you can say that your at-home air sensor is accurate. *Precision* is how close measurements of the same item are to each other. The farther the points are from each other, the lower the precision. For example, you placed two air quality sensors next to each other in your backyard and checked the AQI readings three times in an hour. Sensor 1 measured 32, 35, and 26. Sensor 2 measured 20, 45, and 35. Sensor 1 has a higher precision than Sensor 2.

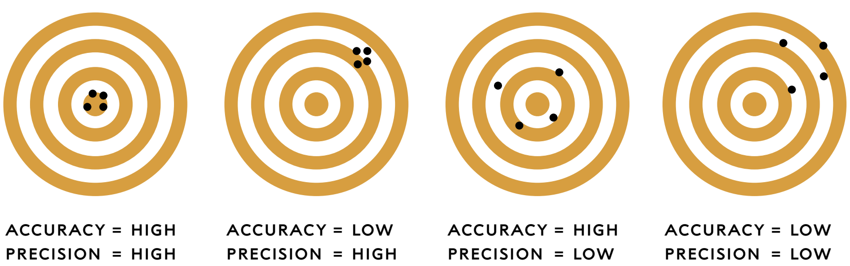


FIGURE 6.5
Accuracy vs. Precision

Validation is a process of determining the accuracy of a measurement. For example, when a scientist wants to determine the accuracy of her air sensor equipment, she can validate it in the lab by observing how close the measurement is to an air sample with a known concentration of air pollutants.

ACTIVITY 7

Reducing Scientific Uncertainty

INVESTIGATION

Sunshine, temperature, wind, and rain are some of the factors that you may consider when going outside to fly a kite. These are some of the same factors that affect outdoor air quality.

7: REDUCING SCIENTIFIC UNCERTAINTY

GUIDING QUESTION

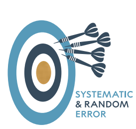
What are ways to collect and analyze data to reduce scientific uncertainty?

INTRODUCTION

Probabilistic reasoning helps assess the chances of different outcomes. People use this kind of thinking all the time in daily life, such as deciding which line in the grocery store will be the fastest or how likely it is to rain on a particular day. Important decisions often have to be made when there is scientific uncertainty. Data may be incomplete, unclear, or different from one source to another. In this activity, you will apply probabilistic reasoning to the analysis of air quality data in a fictional community and investigate the effect of systematic error on conclusions.

Changes that occur in a community, such as construction projects, may sometimes affect local air quality.

CONCEPTUAL TOOLS



MATERIALS LIST

FOR EACH PAIR OF STUDENTS

- SET OF 6 COLORED PENCILS: green, yellow, orange, red, purple, and maroon
- RULER

FOR EACH STUDENT

- 2 STUDENT SHEETS 7.1 “Map of Kairoba”

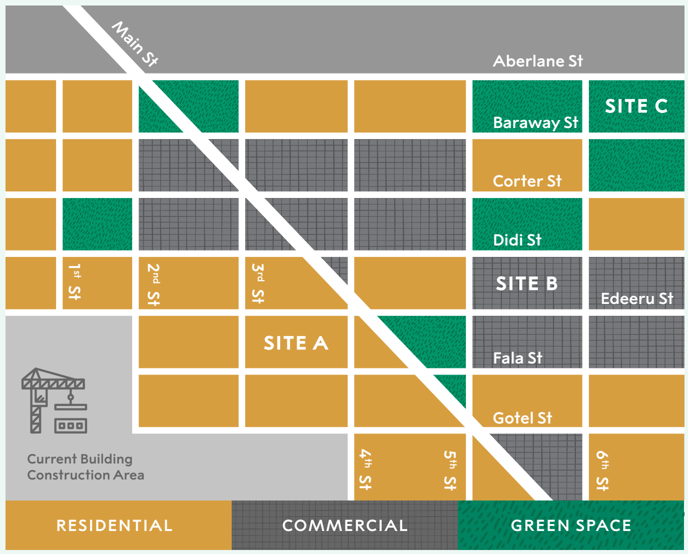
PROCEDURE

PART A: KAIROBA’S AIR QUALITY

1 Read the following fictional scenario.

THE KAIROBA JOURNAL

Local Residents Protest Against New Construction



The city of Kairoba is experiencing a population boom, and city planners are building more housing for the growing population. Many residents who live in neighborhoods near new construction have complained that the air quality has become noticeably worse since building started. They report an increase in asthma and coughs. Local officials point out that the construction work has created new jobs and will provide new housing in Kairoba. For the past 10 years, some residents who were concerned about local air quality bought and installed air quality sensors at intersections near their homes. They plan to use the data from these air quality sensors to influence where future housing construction takes place.

2 Your group of four will role-play air quality experts making a recommendation to the city about which of three locations to build another housing development: Site A, Site B, or Site C. Discuss with your group how a construction project might affect the air quality (both short-term and long-term) at each potential housing site and what type of data could support your ideas.

3 Residents have provided some initial air quality data from their sensors. Your first step is to map the sensor data found in Table 1 on a town map. Decide which half of your group will map the data for October 25 and which half will map the data for November 1 (one week after work at the construction site began).

TABLE 7.1
Kairoba's Air Quality Index (AQI) Data

SENSOR	SENSOR LOCATION	AQI ON OCT 25	AQI ON NOV 1
1	ABERLANE & 2ND	22	36
2	ABERLANE & 5TH	57	39
3	BARAWAY & MAIN	30	45
4	BARAWAY & 6TH	87	35
5	CORTER & 4TH	44	38
6	DIDI & 2ND	61	84
7	DIDI & 5TH	91	81
8	EDEERU & 6TH	91	91
9	FALA & 5TH	55	48
10	GOTEL & 6TH	101	141

4 Record your chosen date of data collection on the line provided at the top of Student Sheet 7.1, "Map of Kairoba." Work with your partner to map the data for your chosen date on your copy of Student Sheet 7.1 by writing the AQI value at the intersection of each pair of streets on the map. Note that the town of Kairoba is laid out on a grid, with numbered streets going from west to east and named streets going from north to south in alphabetical order. Main Street intersects diagonally across the town.

5 Use colored pencils to circle each AQI value on your map with the color of the AQI category it falls into, as shown in Figure 7.1.

FIGURE 7.1
Air Quality Index (AQI)

AQI CATEGORY (COLOR)	INDEX VALUE	DESCRIPTION OF AIR QUALITY
Good (green)	0–50	Air quality is satisfactory, and air pollution poses little or no risk.
Moderate (yellow)	51–100	Air quality is acceptable. However, there may be a risk for some people, particularly those who are unusually sensitive* to air pollution.
Unhealthy for Sensitive* Groups (orange)	101–150	Members of sensitive* groups may experience health effects. The general public is less likely to be affected.
Unhealthy (red)	151–200	Some members of the general public may experience health effects. Members of sensitive* groups may experience more serious health effects.
Very Unhealthy (purple)	201–300	Health alert: The risk of health effects is increased for everyone.
Hazardous (maroon)	301 AND HIGHER	Health warning of emergency conditions: Everyone is more likely to be affected.

* According to the American Lung Association, sensitive groups include children under 18, adults over 65, people with chronic heart or lung disease, people who are pregnant, and people with diabetes. Adults who are active outdoors, including outdoor workers and frequent outdoor exercisers, can be considered sensitive because of prolonged exposure to outside air.

6 With your partner, examine your maps and discuss the following questions.

- What is the range of AQI data on this date?
- What can you conclude about the air quality of Kairoba on this date?

7 Calculate the average AQI on your date and record it in the space provided at the top of Student Sheet 7.1. Then, discuss with your partner how the average AQI compares to the range of AQI data you both mapped.

8 Compare your mapped data and AQI average with the data from the other half of your group (who looked at data from the other date). Then discuss the following questions:

- How did the air quality compare on October 25 vs. November 1?
- Does the mapped data support the claim from the article that the air quality is not affected by construction? Why or why not?
- Explain what additional data would reduce uncertainty in this conclusion.

PART B: MAPPING ADDITIONAL DATA

9 Other local residents shared their data from three additional sensors installed close to the construction site. Record AQI data from these three sensors for the original date and times, as shown in Rows 11–13 of Table 7.2. (All previously collected data from Table 7.1 is shaded.) Plot the additional data from the 3 new sensors on your corresponding Oct. 25 and Nov. 1 maps.

10 Residents also collected data from all 13 sensors on Nov. 8 while construction on the foundation of the building was ongoing (as shown in the last column of Table 7.2). You will receive another copy of Student Sheet 7.1. On your new copy of the student sheet:

a plot all the data for Nov. 8.

b circle each AQI number with the color of the AQI category it falls into, as shown in Figure 7.1.

c calculate and record the average AQI for Nov. 8. Then, discuss with your partner how the average AQI compares to the range of AQI data that you mapped and how it affects your conclusions.

TABLE 7.2
Kairoba's Air Quality Index (AQI) Data with Additional Sensors

SENSOR	SENSOR LOCATION	AQI ON OCT 25	AQI ON NOV 1	AQI ON NOV 8
1	ABERLANE & 2ND	22	36	32
2	ABERLANE & 5TH	57	39	42
3	BARAWAY & MAIN	30	45	40
4	BARAWAY & 6TH	87	35	42
5	CORTER & 4TH	44	38	29
6	DIDI & 2ND	61	84	14
7	DIDI & 5TH	91	81	40
8	EDEERU & 6TH	91	91	49
9	FALA & 5TH	55	48	40
10	GOTEL & 6TH	101	141	51
11*	FALA & 2ND	81	200	221
12*	GOTEL & 3RD	103	199	212
13*	GOTEL & 4TH	76	259	301

* Newly discovered air quality data.

11 Discuss with your group what you can conclude about the air quality of Kairoba on November 8.

12 With your group, place the three maps in the order of their dates. Then discuss:

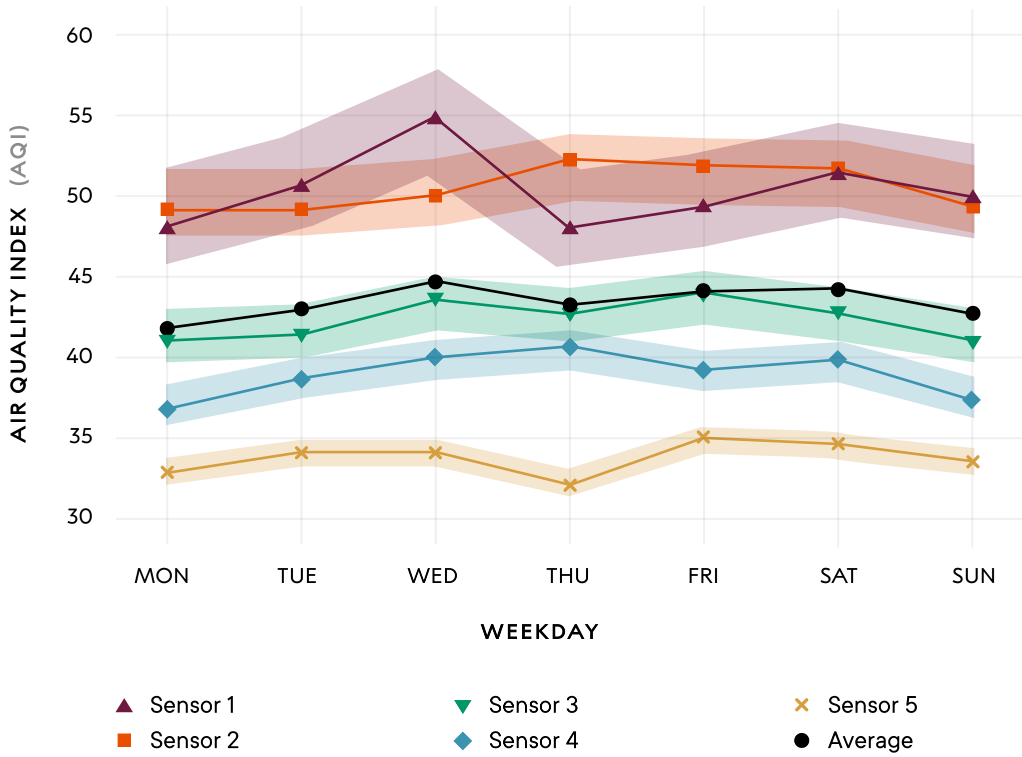
- What systematic error from the initial data set was corrected for when you found the three additional sensors?
- Does this larger data set support the article's claim that the construction has no impact on AQI? Use the data to explain your answer.
- What additional evidence would you want to collect to reduce uncertainty in your conclusion?

BUILD UNDERSTANDING

① Describe at least two possible sources of systematic error when collecting and analyzing air quality data. Then explain how you could address each of these errors to reduce your uncertainty in the data.

② The graph in Figure 7.2 shows daily average AQI at 5 sites in 1 city over the course of a week. The shaded areas indicate the 95% confidence interval of the daily averages at each site. The black line represents the average of the 5 sites for each day.

FIGURE 7.2
AQI Readings for 5 Sites in 1 City



a. Based on the confidence intervals of each sensor, identify which sensor has the least uncertainty and which has the most uncertainty. Describe how this affects your conclusions about the data.

b. A city official proudly claims that the city's AQI is always good (0–50). Do you agree? Support your answer with evidence.

c. Would you be more concerned about your local air quality if you lived near Sensor 1 or Sensor 5? Explain.

3. It is time for the city of Kairoba to make a decision about which site to build. As part of their decision-making, they have created a table of additional considerations. Based on your work in the activity and the information in Table 7.3, where would you recommend the city build? Support your answer with evidence and identify the trade-offs of your decision.

TABLE 7.3
Additional Housing Considerations

	SITE A	SITE B	SITE C
NUMBER OF PROPOSED APARTMENTS	100	100	75
RELATIVE COST TO BUILD	$$	$	$$$

CONNECTIONS TO EVERYDAY LIFE

4. The U.S. Environmental Protection Agency (EPA) provides air quality forecasts for cities and counties around the United States. Look at the air quality forecast in Figure 7.3, which shows PM2.5 levels over a period of six days.

   a What do you predict the PM2.5 levels are likely to be on Monday? Explain your answer, using probabilistic reasoning.

   b What additional information could reduce uncertainty in your prediction?

FIGURE 7.3
Air Quality Forecast

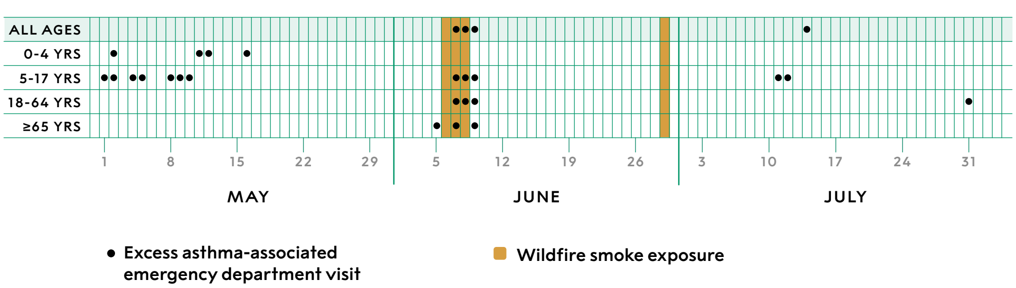
TODAY	TOMORROW	THURSDAY	FRIDAY	SATURDAY	SUNDAY
GOOD PM2.5	MODERATE PM2.5	MODERATE PM2.5	GOOD PM2.5	GOOD PM2.5	GOOD PM2.5

5. Examine Figure 7.4, which shows dots when emergency room (ER) visits due to asthma are higher than average.

   Suppose you want to know if wildfire smoke was a cause for this increase. Explain what conclusions you could make from the data in this graph. Be sure to include the following in your explanation:

   - Describe the patterns you observe between wildfires and ER visits due to asthma.
   - Describe the part of the graph that looks like a signal.
   - Considering the relationship between wildfires and ER visits due to asthma, describe which day(s) appear to be noise.
   - Explain what could have created this noise in the data.
   - Explain what conclusions you can make about the relationship between wildfires and ER visits due to asthma.

FIGURE 7.4
Emergency Room Visits for City Region 1



ACTIVITY 8

Collecting Experimental Data for Predictions

LABORATORY

The way in which a wildfire spreads is affected by the type and amount of vegetation.

8: COLLECTING EXPERIMENTAL DATA FOR PREDICTIONS

GUIDING QUESTION

How can you use experimental results to make predictions about the real world?

INTRODUCTION

Wildfires can blanket an area with smoke, causing air quality problems for people nearby. Since smoke can travel hundreds or even thousands of miles, it can also affect air quality far from the fire. Using probabilistic reasoning to predict air quality changes from wildfires is challenging because of varying factors such as wind, weather, and the landscape.

Fire scientists investigate the type of plants that burn in a wildfire to help predict how the fire will behave. In this activity, you will take on the role of a fire scientist and explore how different plants, or fuel sources, affect wildfires. Grasses, trees, and shrubs are all examples of fuel sources for wildfires.

Fire is a natural part of many ecosystems. Fire scientists study the causes and benefits of fires and help prevent destructive wildfires.

If you need to review the concepts of combustion, thermal energy, calories, and the law of conservation of energy, you will find a Science Review at the end of this activity.

CONCEPTUAL TOOLS

PROBABILISTIC REASONING

SIGNAL & NOISE

MATERIALS LIST

FOR EACH GROUP OF FOUR STUDENTS

- METAL SODA CAN
- 50 mL GRADUATED CYLINDER
- RULER WITH CENTIMETERS (cm)
- WATER 50 mL PER SAMPLE
- THERMOMETER
- ALUMINUM FOIL
- WIRE GAUZE
- METAL TONGS
- 2–3 PAPER CLIPS
- RING STAND, WITH CLAMP AND RING SUPPORT
- LONG-REACH LIGHTER OR MATCHES
- TIMER
- BALANCE THAT MEASURES TO 0.1 GRAMS
- SAMPLES OF OUTDOOR VEGETATION

FOR EACH STUDENT

- STUDENT SHEET 8.1 “Testing Fuel Sources”
- SAFETY GOGGLES

SAFETY NOTE

Make sure to carefully follow all safety instructions from your teacher.

- Wear safety eyewear during this investigation.
- Keep long hair tied back and loose sleeves rolled up.
- Clear all items near your test area and be especially careful not to get your hair or clothing near the flame. If anything besides the fuel sample begins to burn, inform your teacher immediately.
- Keep a cup of water nearby as a fire safety precaution.
- Note that the wire gauze and burning samples may become quite hot to the touch during the investigation. Use metal tongs to pick up the wire gauze and the burned sample.

PROCEDURE

PART A: SETTING UP YOUR EXPERIMENT

1 With your group of four, read the following fictional scenario.

REGIONAL PARK AT RISK OF EXTREME FIRE

Fire scientists in Meso City are worried that a small fire in Koheegee Park could quickly grow into a large, uncontrollable blaze. A hotter climate and drier conditions make this more likely. If the fire reaches the Koheegee Reservoir, it could affect the water quality for local residents since the reservoir provides about half of the city’s water. To understand how fires might spread and how to prevent a large fire, the scientists are testing how different types of fuel sources in the area catch fire (ignite) and burn (combust).

2 Your group will play the role of the Meso City fire scientists. Your goal is to make predictions about how a wildfire might affect different parts of Koheegee Park. You will investigate two characteristics of fuel sources:

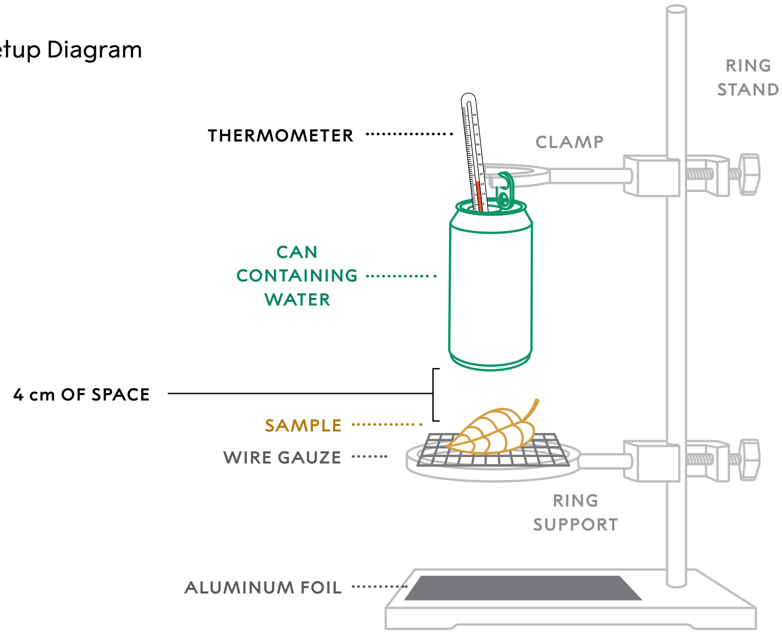
ignition time: how easily or quickly a material will catch on fire

heat of combustion: amount of energy released as heat when a substance is burned

3 As a group, choose two different fuel sources provided by your teacher to test. Record your fuel sources in your science notebook, along with observations about how dry or fresh they are, and a hypothesis as to which fuel source you think will ignite faster, burn hotter, or produce more smoke.

4 Follow your teacher's instructions for setting up the calorimeter. A *calorimeter* is a device used to measure the heat released or absorbed during a chemical reaction or physical change. Figure 8.1 shows a calorimeter constructed using a ring stand. Make sure there is space both above and underneath the wire gauze so air can flow through for your sample to burn.

FIGURE 8.1
Calorimeter Setup Diagram



5 Add 50 milliliters (mL) of water to the can. In Table 1 on Student Sheet 8.1, "Testing Fuel Sources," record the mass of the water in grams (g).

`HINT: For water, 1 mL equals 1 g.`

6 Hang the can from the clamp, using a paper clip if needed. Position the bottom of the can about 4 cm (1.6 inches) above the top of the wire gauze, as shown in Figure 8.1. The bottom of the can should be above the flame of the burning sample but not so low that it will extinguish the flame.

PART B: TESTING YOUR FUEL SOURCES

7 Record a brief description of each sample and its initial mass in Table 1 on Student Sheet 8.1. (You may want to mass your sample on the wire gauze so that you can easily calculate the change in mass after burning.)

8 Place the thermometer in the can. Measure the initial temperature of the water and record it in degrees Celsius (°C) in Table 1 on the student sheet.

9 Position your sample on top of the wire gauze. Light the fuel sample according to your teacher's instructions. Use your timer to measure the time it takes to ignite. Record the ignition time in Table 1.

HINT: Use the following tips to light the fuel sample:

- Hold the flame a couple of centimeters below the edge of the sample.
- Make sure there are no drafts blowing on the lighter or match flame.
- If necessary, use the metal tongs to adjust the sample position.
- If your sample fails to ignite after 15 seconds of applying the flame from the lighter, stop using this sample. Choose another sample to test and start over.

10 After the sample ignites, use tongs to promptly slide the sample under the can and start the timer to measure the length of time it burns. As the sample burns, observe:

- the speed at which the flame spreads through the sample.
- the size of the flame.
- how much smoke is produced.
- how long it takes for the sample to finish burning.

Record your observations in Table 1 on your student sheet.

11 Your sample may not burn completely. As soon as the sample stops burning, record the final temperature of the water in Table 1.

12 When the sample has cooled completely, mass the sample again. Record the final mass in Table 1.

13 Repeat Steps 8–12 for your other sample. You may need to replace the water in the can to reset the temperature. Make sure the can has cooled down before replacing the water.

14 Clean up and safely dispose of your samples before moving on to Procedure Part C.

PART C: ANALYZING YOUR DATA

15 Calculate the energy released as heat by the sample as it burned.

- **a** Record the mass of water for each sample in Table 2 on your student sheet.
- **b** Calculate the temperature change of the water by subtracting the initial water temperature from the final water temperature. Record this value for each sample in Table 2 on your student sheet.
- **c** A *calorie* is the amount of energy it takes to raise the temperature of 1 gram of water by 1 degree Celsius. Use the following formula to calculate the energy released by the sample in calories and then record this value for each sample in Table 2.

energy released by sample (calories) = **mass of water** (g) X **temperature change of water** (°C)

16 Use the following steps to calculate the heat of combustion. Since you began with different masses of samples, this calculation will help you compare them to one another.

- **a** Calculate the change in the mass of the sample by subtracting the final mass (after burning) from the initial mass (before burning). Record this value for each sample in Table 2.
- **b** Use the following formula to calculate the heat of combustion. Record this value for each sample in Table 2.

$$\text{heat of combustion (calories/gram)} = \frac{\text{energy released by the sample (cal)}}{\text{change in mass of sample (g)}}$$

17 Share your results with other groups that tested a similar type of sample and with other groups that tested a different type of sample. Use your data and observations to discuss the following:

- Which fuel sample(s) burned most quickly? Which sample(s) burned most slowly? How might this be a factor in a real wildfire with natural fuel sources such as grasses, dry vegetation, and trees?
- Which fuel samples had the greatest heat of combustion? How might this be a factor in a real wildfire?

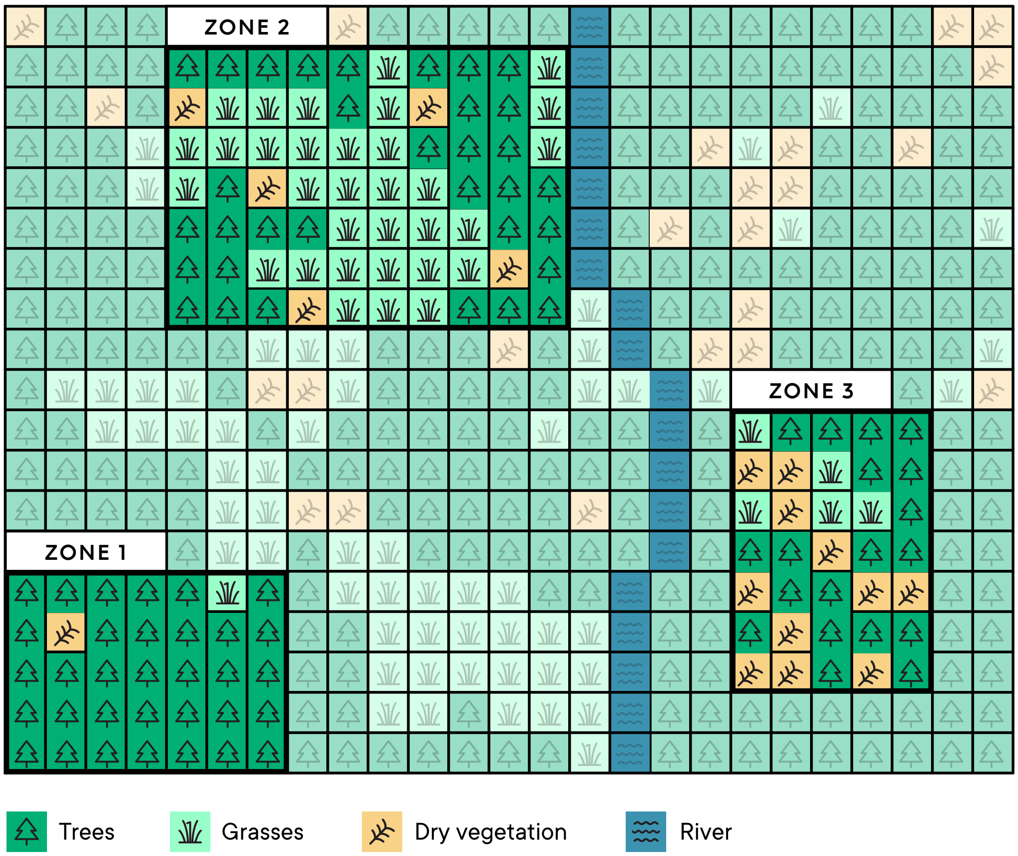
  HINT: Fires that release more energy can burn for longer and may be more difficult to contain.
- Which fuel samples produced the most smoke? How might this be a factor in air quality conditions from a real wildfire?

BUILD UNDERSTANDING

① You conducted an experiment to compare two different fuel sources. Which of the fuel sources do you think is more likely to be ignited in a wildfire? Use the results from your experiment as well as probabilistic reasoning to explain your prediction.

② Examine Figure 8.2, which shows a map of a fictional place called Koheegee Park.

FIGURE 8.2
Map of Koheegee Park with 3 Zones



- **a** Which of the fuel samples tested by your class would best represent each type of fuel source (trees, grasses, dry vegetation) found in Koheegee Park?
- **b** Which section of the park (Zone 1, 2, or 3) is most at risk from a rapidly spreading wildfire? Explain.

③ *Fuel load* is the amount of combustible material in a given area, measured as weight per unit of area. Examine Table 8.1, which shows the annual fuel load in four different regions by latitude.

TABLE 8.1

Annual Fuel Load for Four Regions, 2010–2019 (measured in petagrams)

REGION	LIVE FOLIAGE	LIVE WOOD	DEAD FOLIAGE	DEAD WOOD
A (50°N–90°N)	3	76	27	21
B (23.5°N–50°N)	4	61	19	10
C (23.5°S–23.5°N)	9	347	25	34
D (50°S–23.5°S)	1	18	4	5

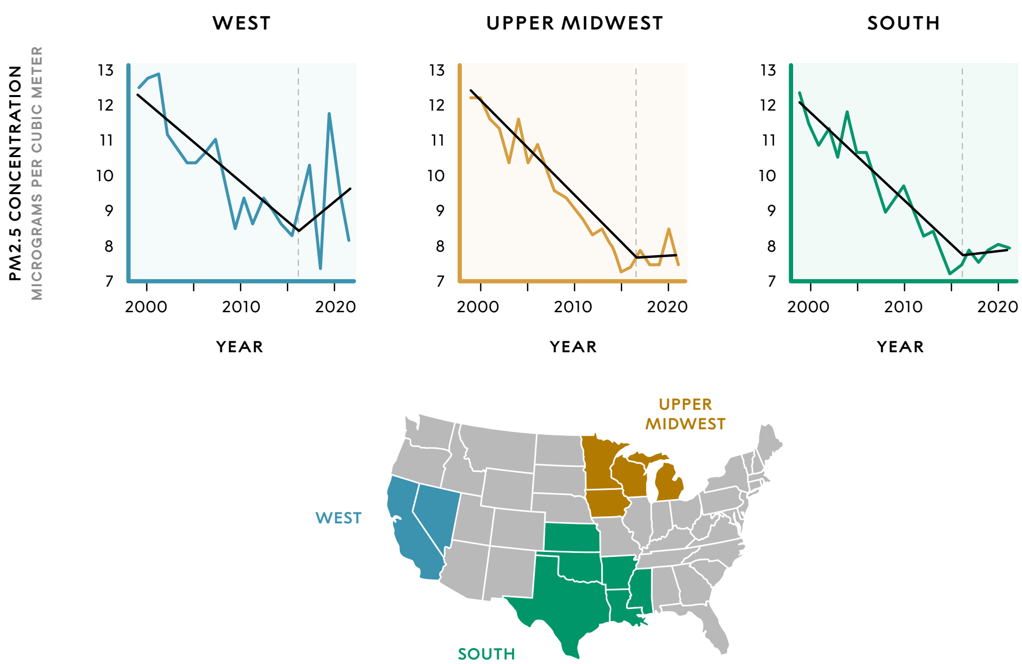
a Based on this data and the results from your experiment, which region do you think is most at risk for wildfires? Explain your reasoning.

b What other factors could influence the wildfire risk in each region?

(4) Researchers examined the contribution of wildfire smoke to average annual PM2.5 concentrations in different regions of the United States by using a combination of ground-based and satellite-based air pollution data from 2000 to 2022. Examine their results, which are shown in Figure 8.3.

FIGURE 8.3

Regional Trends in the United States of PM2.5 Concentrations, 2000–2022



a For each graph, describe the patterns you see in the data both before and after 2016, the year indicated by the dashed vertical line.

b Based on this data, which part of the United States is most likely to have PM2.5 levels that are influenced by wildfire? Explain your reasoning.

c Explain how the trend line on the graph helps you differentiate the signal from the noise.

CONNECTIONS TO EVERYDAY LIFE

5. Imagine you have been saving to go on a trip to an outdoor amusement park during a school break. A ticket costs $100. If you buy your ticket one month in advance, you will get a 20% discount. So far, the weather has been mild and warm with little rain and no snow. Table 8.2 shows average local weather data from the last 5 years for the month you plan to go.

TABLE 8.2
Five-Year Average of Local Weather Data during Annual School Break

YEAR	MINIMUM TEMPERATURE	MAXIMUM TEMPERATURE	TOTAL PRECIPITATION
2020	-1.1°C (30°F)	5.6°C (42°F)	0.1 cm (0.04 in)
2021	3.9°C (39°F)	10.0°C (50°F)	0.2 cm (0.08 in)
2022	0.6°C (33°F)	9.4°C (49°F)	1.0 cm (0.4 in)
2023	11.6°C (53°F)	19.4°C (67°F)	0.0 cm (0.0 in)
2024	5.0°C (41°F)	14.4°C (58°F)	3.0 cm (1.2 in)

a Use probabilistic reasoning to explain whether or not you would buy a ticket in advance.

b What information would reduce the uncertainty in your decision?

EXTENSION

Learn more about predicting and preventing wildfires by watching a video about fire scientists at the Missoula Fire Sciences Lab.

SCIENCE REVIEW

Combustion and Thermal Energy

Combustion is the burning of a substance (the fuel). Combustion converts the fuel into heat, light, sound, water vapor (gas), carbon dioxide, and tiny solids called particulate matter that form smoke. Ash is anything that is not burned completely and is leftover. Combustion requires oxygen to transform the chemical bonds of the fuel into other types of energy. *Thermal energy* is the energy of an object due to the motion of the particles (atoms or molecules) within the object. The amount of thermal energy an object has affects its temperature. The process of combustion is shown in Figure 8.4.

The *heat of combustion* can be measured by calculating the amount of energy released as heat from a fuel while it burns, using a calorimeter. One type of calorimeter is shown in Figure 8.5. It works by absorbing the energy from a combustion reaction contained within a water bath. The amount of energy that enters the water must be equal to the amount that was released by the reaction. This is due to the *law of conservation of energy* that says that energy cannot be created or destroyed, so the total amount of energy in a system is constant. This means that changes in temperature of the water can be used to measure the energy that came from a combustion reaction.

FIGURE 8.4
Combustion

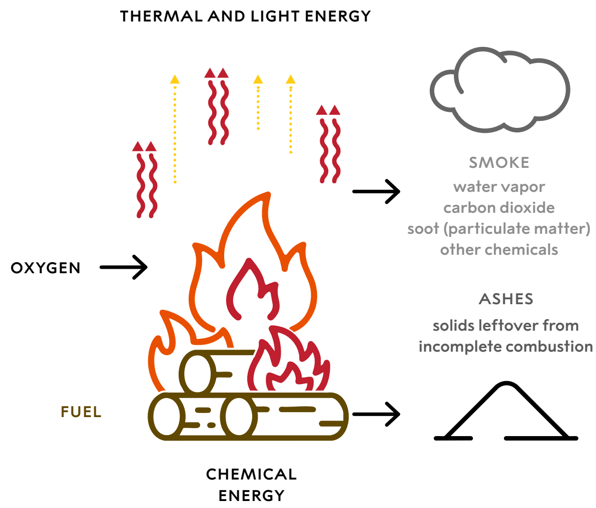
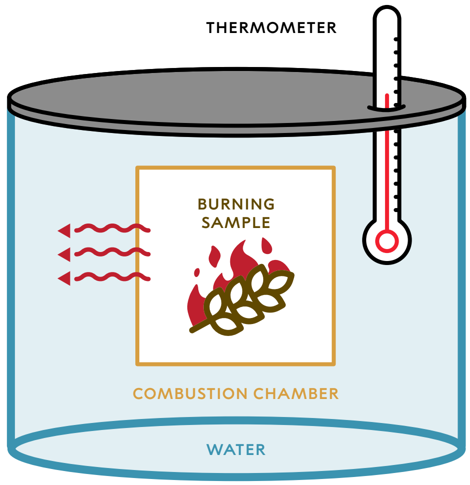


FIGURE 8.5
Calorimeter



ACTIVITY 9

Probabilistic Modeling

MODELING

A firefighting helicopter releases a water drop over a wildfire.

9: PROBABILISTIC MODELING

GUIDING QUESTION

How can probabilistic reasoning be used to predict an outcome?

INTRODUCTION

Many factors affect air quality, some of which are unpredictable. For example, in the summer of 2017, a wildfire started near Brian Head, Utah. The fire quickly grew from 20 to 202 hectares (50 to 500 acres) in just 3 hours. Fortunately, the town had used probabilistic reasoning to plan for wildfire risks and had partially cleared 4 hectares (10 acres) of nearby forest to reduce fire spread. This strategy saved many homes, even as the fire spread almost 28,000 hectares (70,000 acres) in the opposite direction. A year later, resident Lynn Mulder said, "Circumstances including wind changes and the immediate responses of firefighting agencies and the availability of aircraft were vital, but we believe that this small area of clearance saved the town."

National Oceanic and Atmospheric Administration (NOAA) researchers test a new AI fire monitoring and modeling technology at a research facility in Colorado.

Predicting outcomes can be difficult when information is incomplete or uncertain. To help, scientists use probabilistic models. A probabilistic model is a tool that uses patterns in data to predict the likelihood of different outcomes. These models are often computer programs that can process large amounts of data. The computer programs then use probabilistic reasoning to predict the most likely outcomes, sometimes using artificial intelligence (AI). In this activity, you will investigate how a probabilistic model can help predict the spread of a fire.

CONCEPTUAL TOOLS

MATERIALS LIST

FOR EACH PAIR OF STUDENTS

- 2 NUMBER CUBES
- SET OF 8 COLORED PENCILS

FOR EACH STUDENT

- 2 STUDENT SHEETS 9.1 "Map of Koheegee Park"
- 2 STUDENT SHEETS 9.2 "Modeling Wildfire Spread"
- 2 STUDENT SHEETS 9.3 "Data from the Wildfire Model"

PROCEDURE

PART A: SIMULATING A WILDFIRE

1 With your partner, read the following fictional scenario.

PROBABILISTIC MODELS HELP STOP WILDFIRE SPREAD

A firefighter monitors a controlled burn.

A team at the Department of Forest and Fire Protection in Meso City has worked with local fire scientists to develop a new probabilistic model they say will help fight wildfires in Koheegee Park. The model was developed to predict how and where a wildfire will spread, using factors such as vegetation, wind patterns, and local geography. The team hopes to use the model to direct firefighting resources and prevent small fires from burning out of control. Rangers also plan to use the model for forest management. By knowing which areas of the forest are most vulnerable to fires, they can perform controlled burns to remove dry vegetation. Controlled burning is a form of forest management that burns off underbrush, dead trees, and other dry vegetation in order to reduce wildfire risk and maintain forest health.

2 Imagine that a wildfire has just started in the middle of Koheegee Park. As a fire scientist at the Meso City lab, your job is to predict how the fire will spread over the next 8 hours, using the Wildfire Model.

a. Examine the map on Student Sheet 9.1, "Map of Koheegee Park," and the location of the three fuel sources within the park. Label the fire on this student sheet as Fire A.

b. On Student Sheet 9.2, "Modeling Wildfire Spread," draw your prediction of the general area you think the fire will spread the most over the next 8 hours and list the factors that you considered when making your prediction. Label the fire on this student sheet as Fire A.

3 With your partner, read the guidelines for conducting the Wildfire Model in Table 9.1.

4 With your partner, conduct the model by using the information in Table 9.2 to determine:

a. the probability of each wind direction.

b. the probability of each wind speed.

c. the fuel type that is least likely to spread a fire.

HINT: You may want to revisit the Science Review on probability found at the end of Activity 1.

5 With your partner, use the Wildfire Model to simulate fire spread for Hours 1–8. For each hour, you will:

a. complete each action in the order listed in Table 9.2.

b. record the wind direction, wind speed, and fuel sources on Student Sheet 9.3, "Data from the Wildfire Model." Label the fire on this student sheet as Fire A.

c. track the spread of the fire on an hourly basis by coloring squares on the map on Student Sheet 9.1. Use the colors indicated in the first column of Student Sheet 9.3 to represent the fire spread during each hour.

6 On Student Sheet 9.2, sketch the shape and direction that the fire spread after Hour 8 and describe the factors that had the greatest influence on the spread of the fire.

7 Share your results and your wildfire map with the class.

TABLE 9.1
Wildfire Model

- Every round represents 1 hour.
- The fire will start in the middle square on the map (with the fire icon) prior to Hour 1.
- Once a square catches on fire, it will continue to burn all 8 hours (throughout the model).
- The fire cannot spread onto squares that are already burning.
- The fire cannot spread across the river (except when strong winds cause a firestorm).

TABLE 9.2
Conducting the Wildfire Model

FACTOR	ACTION
Wind direction	The fire spreads 1 square in the direction of the wind **from every burning square.** Roll a number cube to determine the wind direction: ⚀ → **N** ⚁ OR ⚂ → **E** ⚃ OR ⚄ → **S** ⚅ → **W**
Wind speed	Roll a number cube for the wind speed: ⚀ OR ⚁ OR ⚂ → 5 mph No additional spread from wind. ⚃ OR ⚄ → 15 mph The fire spreads 1 more square in the direction of the wind **from every burning square.** ⚅ → 30 mph Firestorm! The fire spreads 3 more squares in the direction of the wind **from every burning square**. If the fire reaches the river, it crosses to the other side because embers are blown by the strong winds. The river does not count as a square.
Fuel source	Some fuel sources spread fire to surrounding areas. **For every burning square,** the fire spreads as follows: **Trees** → The fire does not spread to any additional squares. **Grasses** → The fire spreads 1 additional square in the direction of the wind. **Dry vegetation** → The fire spreads 1 additional square in every direction.

PART B: CONTAINING THE FIRE

8 Imagine you have a chance to fight a wildfire starting at the same location in Koheegee Park as you did in Procedure Part A. Your goal is to minimize the spread of the fire. Use your knowledge of the Wildfire Model and the map of Koheegee Park to decide where to send resources to stop the fire.

   a Label your second copies of your student sheets as Fire B.

   b On your second copy of Student Sheet 9.2, sketch your prediction of the general area you think the fire will spread the most in 8 hours and list the factors you considered when making your prediction.

   c Repeat Procedure Step 5 for Hour 1, recording the data on your second copy of Student Sheet 9.3 and tracking the spread of the fire by coloring squares on your second copy of Student Sheet 9.1.

   d **Before completing any actions for Hours 2–8**, you may send a firefighting helicopter to make a water drop onto one square of fire. Fire cannot spread into or out of this square in all future hours. Choose the square (on Student Sheet 9.1) where you would like to have the water drop and mark it with an X.

   e Repeat Step 5 for Hours 2–8 and then repeat Step 6 after Hour 8.

9 Compare your results with the results from another pair of students. Determine:

   - in which direction the fire moved the most.
   - what factor(s) most influenced the outcome.
   - how the different factors interacted in a way that caused the fire to spread.

BUILD UNDERSTANDING

1. Based on the Wildfire Model from this activity, make a recommendation to Koheegee Park Rangers about what areas would benefit from controlled burning. Support your reasoning with evidence from the model.

2. Kairoba City is north of Koheegee Park, while Meso City is east of Koheegee Park. According to the Wildfire Model, which city is more likely to have poor air quality as a result of the fire in Koheegee Park? Explain your reasoning.

3. Consider the factors that contribute to wildfire in the model.
   - **a** What are some ways that the model is realistic?
   - **b** What are some other factors that could affect wildfire spread that are not included in the model?
   - **c** How is the Wildfire Model an example of a probabilistic model?

4. Suppose you want to use the Wildfire Model to plan water drops to contain a fire in Koheegee Park. You have numerous helicopters and can drop water on as many locations as you want. However, each water drop costs $2,000 and can be dangerous to firefighting crews that will need to navigate difficult weather conditions and terrains.
   - **a** A probabilistic wildfire model predicts a fire to spread south of the park. People there are warned, and firefighters are sent to assess the situation and do a water drop if needed. When they arrive, there is no fire. What might be one consequence of this false positive?
   - **b** Describe how a probabilistic wildfire model could result in a false negative and what one consequence of this false negative could be.

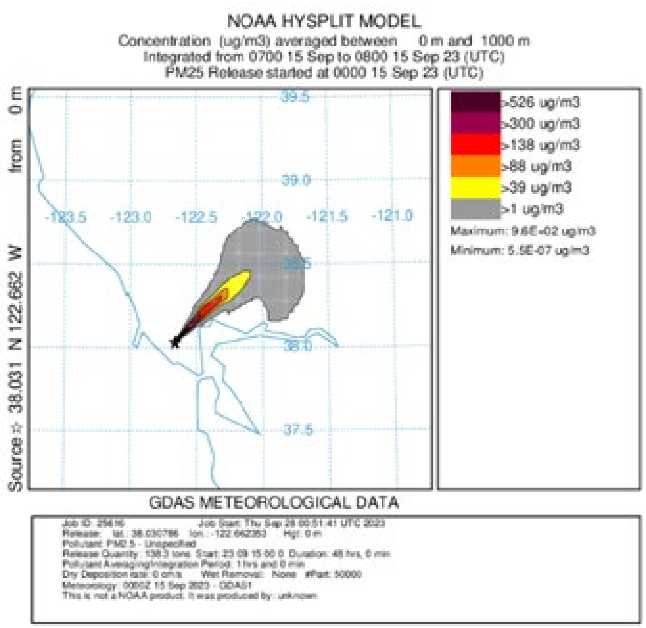
5. Imagine 2 different wildfire models are being tested to make predictions about local wildfires. Model A predicts that a fire at a particular location will burn between 1,000 and 2,000 hectares (2,471–4,942 acres) under certain wind and weather conditions and has a 95% confidence level. Model B predicts that a fire at the same location will burn between 1,000 and 2,000 hectares under certain wind and weather conditions and has a 90% confidence level. Explain which model you would prefer to use when making wildfire predictions for this area and why.

CONNECTIONS TO EVERYDAY LIFE

⑥ Probabilistic reasoning is a key component of several decision-making systems, including those used in AI tools. For example, an image-generating AI tool learns to make images by looking at millions of images in a database. It finds patterns in the data it is given and uses those patterns to project how things should look. A problem arises because the data the AI learns from can be unbalanced or biased. For example, if most of the pictures the AI learns from show only certain types of people (such as men as doctors and women as nurses), the AI might recreate those stereotypes in its images. Discuss with your class whether this represents random error, systematic error, or both in the AI image-generator system.

EXTENSION

The spread of smoke can also be modeled using probability. Evaluate where smoke from a controlled burn spreads over an 8-hour period in your own local area, using National Oceanic and Atmospheric Administration's (NOAA) online HYSPLIT computer model.



KEY SCIENTIFIC TERM

probabilistic model

ACTIVITY 10

Applying Probabilistic Reasoning

CARD-BASED INVESTIGATION

High winds can knock trees onto powerlines, sparking wildfires in dry, fire-prone areas.

10: APPLYING PROBABILISTIC REASONING

GUIDING QUESTION

How can you use probabilistic reasoning to reduce risk?

INTRODUCTION

The risk of wildfires starting from sparks from electric power lines is a growing concern. To reduce this risk, many electric companies turn off power lines during bad weather. These power shutoffs help prevent wildfires. In the past, power shutoffs were based on historical records of past wildfires. But in 2023, Professor Paolo Bocchini and student Xinyue Wang from Lehigh University in Pennsylvania created a new probabilistic model to predict the risk of fire from power lines during strong wind storms. They looked at factors such as how close power lines are to vegetation, how wind affects the cables, wind strength, and how long the wind lasts. According to Professor Bocchini, their team used a careful probabilistic approach to the problem. This new model can help people better predict the probability of a wildfire due to a spark from a power line and decide if a power shutoff is needed. In this activity, you will use probabilistic reasoning to help a fictional power company decide where and when to turn off power.

Signs can alert visitors to the level of fire risk in an area.

CONCEPTUAL TOOLS

MATERIALS LIST

FOR EACH GROUP OF FOUR STUDENTS

- CALCULATOR
- SET OF 4 FIRE RISK CARDS

FOR EACH STUDENT

- STUDENT SHEET 10.1 “Determining Fire Risk”

PROCEDURE

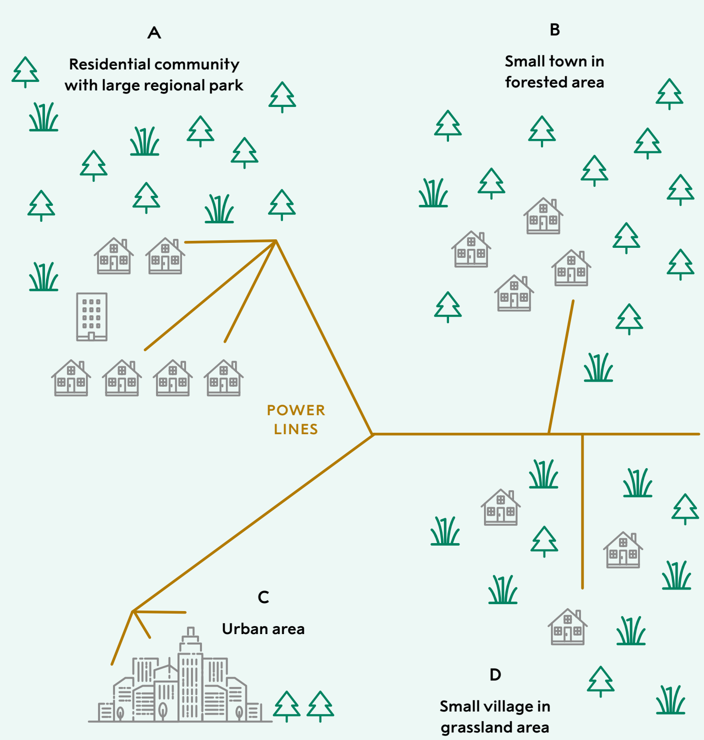
1 With your group of four, read the following fictional article.

POWER SHUTOFFS LIKELY DUE TO HIGH WINDS

Wind gusts of up to 72 kilometers per hour (45 miles per hour) are forecast over the next 3 days in Futura County. This, in addition to dry conditions throughout the region, has officials concerned. When winds are strong, like those predicted, it can cause branches to fall onto electrical power lines. This can start a fire, which then spreads quickly by the wind.

The local power company has shared the following map of areas of concern. Each of these four locations may experience a power shutoff during this period of particularly high winds.

FIGURE 10.1
Map of Major Power Lines to Each Area



2 Work with your partner to examine Figure 10.1.

a In the first row of the table on Student Sheet 10.1, "Determining Fire Risk," record the number of major power lines terminating in each location.

b In the second row of the table, describe the type and amount of vegetation at each location.

c In the third row of the table, describe the location of power lines relative to the vegetation at each location.

3 Based on the data you recorded on your student sheet so far, make an initial prediction about which location(s) has the highest risk for wildfire for the next 8 hours.

a Record your prediction in your science notebook and explain your reasoning.

b Determine how sure you are of your decision, using a scale of 0–100% where:

0% = there is no chance your prediction will be correct
50% = your prediction is just as likely to be wrong as it is to be correct
100% = you are absolutely sure that your prediction will be correct

c Share your prediction and reasoning with the rest of your group. Remember to listen to and consider the ideas of other group members. If you disagree with others, explain why you disagree.

4 Using your set of Fire Risk cards, examine Fire Risk 1: Vegetation Dryness by Location. Record notes about vegetation dryness for each location in the "Current Dryness Level" row on the student sheet.

Clearing dry vegetation is one way to create a defensible space and reduce wildfire risk.

5 The frequency of fires in the past has often been used to determine the risk of fires today. Examine Fire Risk 2: Past Fire Data by Location. Calculate the percentage likelihood of a significant fire in any given year in each location based on the past number of wildfires greater than 41 hectares (100 acres) in the last 50 years. Record this percentage in the "Likelihood of Fire Each Year (%)" row on the student sheet.

HINT: Divide the number of past wildfires for each location by the length of the time period (50 years) to get the average number of fires per year. Then multiply by 100 to get a percentage.

6 You will use Fire Risk cards 3 and 4 to determine the size of a possible fire at each location. Examine Fire Risk 3: 8-Hour Forecasted Maximum Wind Speed in Each Location, which shows the forecasted maximum wind speeds over the next 8 hours for each location. Record this data in the "Forecasted Maximum Wind Speed" row on the student sheet.

7 Examine the graph on Fire Risk 4: Wildfire Size Relative to Maximum Wind Speed and describe:

- any patterns you see in the graph.
- what conclusions you can make about the relationship between maximum wind speed and fire size.
- what role the size of error bars plays in your conclusions about the data.

8 Based on the data for each location on Fire Risk Card 3 and the graph on Fire Risk Card 4, determine the size of a possible fire at each location. Record this information in the "Size of Possible Fire Under Forecasted Maximum Wind Speed (hectares or acres)" row on the student sheet.

9 Based on all the data in Table 1 on your student sheet so far, revise your prediction about which location(s) has the highest risk for wildfire in the next 8 hours.

a Record your revised prediction in your science notebook and explain your reasoning.

b Determine how sure you are of your decision and compare how it changed as you gathered more data. Use a scale of 0–100% where:

0% = there is no chance your prediction will be correct
50% = your prediction is just as likely to be wrong as it is to be correct
100% = you are absolutely sure that your prediction will be correct

c Share your revised prediction and reasoning with the rest of your group. Remember to listen to and consider the ideas of other group members. If you disagree with others in your group, explain why you disagree.

d Work with your group to describe at least two sources of scientific uncertainty, including systematic errors or random errors, that could have affected the data or your predictions.

10 The power lines directed to the different communities have different energy loads based on local energy needs. Table 10.1 shows the energy required by each community and the vulnerability of each community to a loss of power, where a higher number indicates increased vulnerability. Vulnerability is a result of multiple factors such as income, health, and age. During a power outage, there are no traffic lights, no household refrigeration, and no power for medical devices. The ability of customers to cope with power outages varies, with some populations more vulnerable to its effects.

Examine the data in Table 10.1. Discuss with your group the trade-offs of shutting off the power to each location. Record notes for each location in the "Trade-Offs of Power Shutoff" row on your student sheet.

TABLE 10.1
Energy Load and Vulnerability Index

LOCATION	ENERGY LOAD CARRIED BY A POWER LINE (kW)	POPULATION	VULNERABILITY INDEX OF LOCAL POPULATION 0 = LOWEST VULNERABILITY 10 = HIGHEST VULNERABILITY
A	2,453	4,672	4
B	185	1,109	2
C	1,013	20,941	2
D	200	477	9

11 Use all the information on your student sheet to discuss which communities should have a power shutoff for the next 8 hours in order to reduce wildfire risk. Share your ideas and reasoning with the rest of your group. If you disagree with others in your group, explain why you disagree.

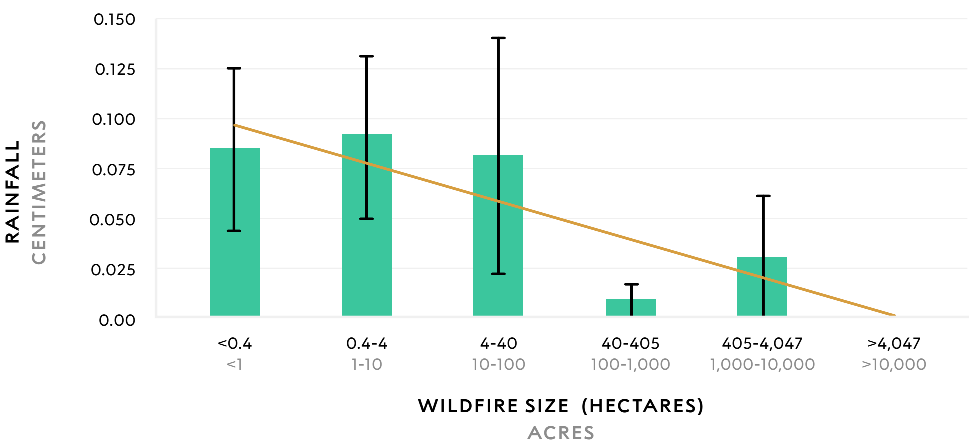
12 Follow your teacher's instructions to share your thinking with the class about which location(s) should experience a power shutoff to reduce wildfire risk.

BUILD UNDERSTANDING

1. With the available data and under current conditions, in which location(s) do you recommend shutting off the power during the next 8 hours to reduce wildfire risk? Support your answer with evidence and identify the trade-offs of your decision.

2. Figure 10.2 shows the average size of wildfires in this region when there are different amounts of rainfall during the wildfire events. Explain what conclusions you can make based on the data in the graph. Be sure to include the following in your explanation:
   - Describe what patterns you observe between wildfire size and rainfall.
   - Explain what conclusions you can make about the relationship between wildfire size and rainfall.
   - Explain at least two possible sources of scientific uncertainty, including scientific errors, that may have affected the data.

FIGURE 10.2
Regional Wildfire Size During Different Rainfall Amounts



3. The power company decides to install sensors along power lines that would result in an automatic shutoff if a tree branch touches a power line.
   - **a** What would be a false positive in this situation?
   - **b** What would be a false negative in this situation?
   - **c** Which would be a greater concern for community safety: a false positive or a false negative from the shutoff sensor? Explain your reasoning.

CONNECTIONS TO EVERYDAY LIFE

4. Your friend says that since there is uncertainty in science, no one can really know anything in science. Explain why you agree or disagree with your friend. Support your answer with an example from this unit.

STUDENT GLOSSARY

accuracy
the closeness of a measured value to a standard or true value

air quality index (AQI)
a scale that measures air quality that ranges from 0 to 500; lower numbers indicate better air quality

calibrate
compare the readings of a scientific tool with those of a standard in order to check the instrument's accuracy

calorie
the amount of energy it takes to raise the temperature of 1 gram of water by 1°C

calorimeter
a device used to measure the heat released or absorbed during a chemical reaction or physical change

combustion
burning of a substance (the fuel)

confidence interval
the range of data expected to contain the true value

confidence level
a statistical measure of the probability that the true value is within a specified range

confound
a factor that can distort or hide the relationship between two variables being investigated in a study

control
a basis of comparison for checking the effects of an experiment

crowdsourcing
when data or ideas are collected by many people in a community

error bars
a visual representation of the amount of uncertainty in a measurement, presented as a ± (plus or minus) range of numbers that fall above and below the average

false negative
a type of error when something is incorrectly identified as absent but is actually present

false positive
a type of error when something is incorrectly identified as present

fuel load
the amount of combustible material in a given area, measured as weight per unit of area

heat
the transfer of thermal energy from one substance to another

heat of combustion
amount of energy released as heat when a substance is burned

hectare
a metric unit for measuring land area, equal to 2.47 acres or 10,000 square meters

hypothesis
a possible explanation for observations, facts, or events that may be tested by further investigation

ignition time
how easily or quickly a material will catch on fire

law of conservation of energy
energy cannot be created or destroyed, so the total amount of energy in a system is constant

micrometer (μm)
one millionth of a meter, which is equal to 0.001 millimeters, or about 0.00004 inches; also known as micron

noise
information that hides, distracts from, or falsely resembles the meaningful information that is being investigated

particulate matter (PM)
microscopic particles suspended in the air that are so small that they can be inhaled

petagram
a unit of mass equal to 10^{15} grams (1 billion metric tons)

PM2.5
particulate matter in the air that has a diameter of 2.5 micrometers (μm) or less, also known as fine particles

PM10
particulate matter in the air that has a diameter of 10 micrometers (μm) or less, includes PM2.5 particles

precision
how close measurements of the same item are to each other

probabilistic model
a tool that uses patterns in data to predict the likelihood of different outcomes

probabilistic reasoning
a way of making predictions or drawing conclusions based on how likely something is to happen, especially when there is not enough clear data

probability
the measure of the likelihood that a specific event will occur

qualitative data
data that consists of categorizations or verbal descriptions, rather than numbers

quantitative data
data that consists of numbers; for example, measurements or quantities

random error
a difference between an observed and true value that has no consistent pattern and is caused by chance and/or unpredictable factors

range
the lowest and highest values in a data set

reliable
able to be reproduced consistently

scientific error
the difference between a measured or observed value and the true value of a quantity

scientific uncertainty
an understanding that there are limits to data and conclusions about the natural world, and additional data and/or investigations can lead to increased surety or new questions

signal
meaningful information about the phenomenon that is being investigated

systematic error
a difference between an observed and true value in a consistent direction, often caused by experimental equipment or design

thermal energy
the energy of an object due to the motion of the particles (atoms or molecules) within the object

trade-off
a desirable outcome given up to gain another desirable outcome

true value
the actual number that would be found if the measurement could be made without error

validation
a process of determining the accuracy of a measurement

variable
a feature, factor, or result that can change or vary

CREDITS

Cover Merijn Jansen;
v terawat/iStock, 2023;
2 Stylianos Stamatoglou/iStock, 2023;
5 Sean Locke/Stocksy, 2018;
6 United States Environmental Protection Agency, 2025;
7 hapabapa/iStock, 2020;
10 Common Ground Health, 2016;
15 numismarty/iStock, 2015;
16 Delphine Farmer/Indoor Chemistry: The CASA Experiment, 2020;
17 enginakyurt11/iStock, 2023;
19 National Interagency Coordination Center, 2024;
22 PurpleAir, 2022;
26 United States Environmental Protection Agency, 2024;
31 Wirestock/iStock, 2023;
32 Meg Buscema/Georgia State University, 2024;
33 Kristina Marie Duncan/ The Regents of the University of California, 2025;
36 monkeybusinessimages/iStock, 2018;
37 United States Environmental Protection Agency, 2023;
39 Dawar Hameed Butt/Nature, 2024;
41 Environmental Protection Agency, 2022;
42 SashaBuzko/iStock, 2008;
43 Harry Cabluck/AP Photo, 1972;
45 Johanna Lepeule et al./ Environmental Health Perspectives, 2012;
46 Francisco Blanco/iStock, 2024;
48 Fabián Ponce/iStock, 2021;
49 Joan Casey et al./Proceedings of the National Academy of Sciences of the United States of America, 2024;
51 Mike Wilkinson/The Nature Conservancy, 2024;
53 emretopdemir/iStock, 2018;
54 AirQo, 2019;
56 The Regents of the University of California, 2025;
62 Dave Jonasen/iStock, 2024;
63 Modfos/iStock, 2016;
68 Celeste McFarlane et al./ Aerosol and Air Quality Research, 2021;
70 Centers for Disease Control and Prevention, 2023;
72 Mike Lewelling/National Park Service, 2013;
73 Jacob Lund/Alamy, 2019;
74 Viktor Cvetkovic/iStock, 2020;
79 Marshall Burke et al./Nature, 2023;
83 MariusLtu/iStock, 2019;
84 Fire Weather Testbed/National Oceanic and Atmospheric Administration, 2024;
85 David Carrera, BLM/Flickr, 2020;
90 National Oceanic and Atmospheric Administration, 2025;
92 photovs/iStock, 2024;
93 burroblando/iStock, 2024;
95 Pamela Au/Alamy, 2015;
98 John R. Weir, Sam D. Fuhlendorf, Angela M. Reid/ Oklahoma State University, 2017;

Graphs and figures may have been modified from original for instructional purposes.