FOREST WATCH

STUDY GUIDE

Introductory
&
Math Activities

FOREST WATCH

STUDY GUIDE

INTRODUCTORY ACTIVITIES

Contents

WHAT IS A SCIENCE PROTOCOL?

A

ACTIVITY/ TEACHER RESOURCE

How can science protocols help you as a scientist share data?

Forest Watch is a collaborative scientific investigation consisting of students, classroom teachers and research scientists. As part of this investigation, students will be asked to follow a set of protocols. In these introductory activities students will develop an understanding of the term science protocol and its application to scientific investigation.

Materials

- age appropriate cartoons with 6 to 8 frames
- scissors

Procedure

Order out of chaos — cartoon 10-minute introduction

The objective of this activity is to demonstrate that a proper sequence or protocol is sometimes needed to reach desired results. To accomplish this objective students will put in proper sequence randomly ordered frames from a cartoon.

Procedure

- Find a cartoon with six to eight frames that you can cut into one-frame pieces. Number the frames on the back for future reference.
- Rearrange the frame pieces and tape together, out of order. Photocopy and make copies for student teams.
- Give the scrambled cartoons to student teams and ask them to notice if something is wrong.
- Students cut out and rearrange the frames in sensible order.
- Show students the original cartoon.

At the conclusion of this activity, emphasize that some things need to be in a specific sequence to give desired results. The frame containing the punch line will not have clear significance when viewed out of context. In a scientific investigation all elements need to be in the correct sequence to lead to a meaningful conclusion.

OBSERVING PLANT GROWTH AND DEVELOPMENT

ACTIVITY

What are science protocols and why are they necessary?

Many scientific investigations require protocols, a set of very precise directions, which enable scientists to collect, analyze and share reliable data. This activity will help you understand why protocols are important and how they are constructed.

In this activity, you will be part of a group of students who will design and carry out an experiment in which you will compare the growth and development of two different seed types. You will plant and observe the seeds as they germinate and grow into young seedlings.

Steps for developing sound scientific protocol:

1. Design an experiment that will help you answer a researchable question.
2. Conduct the experiment.
3. Collect, organize and analyze your data.
4. Share your data and conclusions.
5. Review and adjust your experimental design if necessary.

Procedure

During this activity your group will:

1. Complete a Research Proposal Form:
   a. Develop a researchable question: A question that can be answered through observation, experimentation and measurement.
   b. Provide a materials list: A list of all the supplies you will need during this experiment.
   c. Write a procedure: Detailed steps for performing the experiment. Your teacher will provide you with a time frame for completion of your work.
   d. Develop requirements for data collection and analysis: What type of data will be collected? How will the data be organized and analyzed?
2. Submit the research proposal form to your teacher for review, suggestions and approval.
3. Conduct your research experiment: Plant the seeds using the procedure you developed with your group. Collect and record your data for the amount of time specified by your teacher.
4. Share what you have learned: After the data have been collected, your group will share your results and observations with the class. Be prepared to answer several questions. What difficulties were experienced during the experiment? Which data proved most useful when comparing observations? Were you able to answer the researchable question? How could this experiment be improved? What are some other researchable questions that could be investigated regarding plant growth?
5. Complete a summary worksheet.

RESEARCH PROPOSAL FORM

* Please submit this form for approval by your teacher before you begin your experiment

GROUP MEMBERS ______________________ ______________________

______________________ ______________________

______________________ ______________________

CLASS ______________________ DATE ______________________

Your Researchable Question

Write a question that can be answered through direct observation and/or measurement. Don't make your question too complicated or involved. Even simple questions may be difficult to answer in a scientific experiment.

Materials List

List all the materials you will need to complete this experiment. Your teacher will let you know which supplies you may choose from.

Procedure

List the steps your group will follow to complete this experiment. Include labeled diagrams where appropriate.

Data Collection

Make a list of the types of information you will collect during this experiment. What will you observe or measure about the seeds and seedlings that will help you answer your researchable question?

How will the data be organized?

How often will the data be collected?

Make a data table to help organize your observations.

Comments

Add any additional information that is important for your study in the space below.

Teacher Signature ______________________ Date ___________________

PLANT GROWTH SUMMARY WORKSHEET

Name ____________________________

Date ____________________________

The Researchable Question

1. What was the researchable question that your group investigated?

2. Did your results provide an adequate answer to your researchable question? How could you change the question you asked to make it better?

3. List at least 3 more examples of researchable questions that could be used in a study of plant growth.

Standard Protocol

1. What observations or measurements did other groups make to study differences in plant growth?

2. Did all groups use the same criteria to compare the growth of the plants? Which observations or measurements were the most useful in answering their researchable questions? Why?

3. What problems or unexpected outcomes were encountered during the experiment?

4. In science experiments, it is often important to run multiple trials of an experiment in order to make sure you are observing as many possible outcomes as you can. Did any of the other groups have the same researchable questions as yours? Could you trade your data for theirs?

   What precautions would you have to take before you combined your data together?

5. What changes could you make in your procedures to standardize them and make it possible to use each other's information in the analysis of your researchable question?

Summary

What are science protocols and why are they necessary?

OBSERVING PLANT GROWTH AND DEVELOPMENT

TEACHER RESOURCE

Introduction

Science investigations require protocols. Protocols are necessary for the exact duplication, verification and validation of experimental results. Protocols may appear to be similar to the "cookbook approach" used for convenience and time management in school laboratory manuals. This activity will help your students understand why protocols are important and how scientists use them to collect, analyze and share data.

Guiding Question

What are science protocols and why are they necessary?

Materials

- 15 or more seeds each from two contrasting types
- potting soil
- containers for growing seeds (may be supplied by students)
- metric ruler or measuring tapes
- graduated cylinder
- water

Background

In this activity, student work groups design their own scientific protocol in order to answer a researchable question. The researchable question should focus on differences that may be observed and measured between the germination and growth of two plant seed varieties. Students begin the activity by completing a research proposal form. The purpose of the form is to give student groups an opportunity to work through the steps necessary to complete the experiment.

The research proposal form asks the student groups to:

1. Develop a researchable question: A question that can be answered through observation, experimentation and measurement.

   Before designing the experiment, students will need to formulate a researchable question. This question should help students to decide on the types of observations and measurements they will collect during the experiment. It will be necessary to guide students through the process of how to ask a researchable question.

Examples of feasible researchable questions:

- How long will it take for the first seedlings to appear?
- Which seed variety has greater % germination?
- After 10 days, which variety has the tallest seedlings?
- What are some differences that can be observed in the leaf formation of the 2 varieties of plants?
- Which seed type will germinate first?

Examples of nonresearchable questions:

- How do plants grow?
- How healthy are the seeds?

2. Provide a materials list: A list of all the supplies they will need during the experiment. You might want to provide a master list from which they can choose.
3. Write a procedure: Detailed steps for performing the experiment. You should provide your students with a reasonable time frame for completion of their work.
4. Requirements for data collection and analysis: What type of data will be collected? How will the data be organized and analyzed?

Carefully review your students research proposals making appropriate recommendations. In this activity, it's OK if some mistakes are made. One of the objectives is to demonstrate the importance of feasibility in a study.

After the data have been collected, lead the class in a discussion of their results.

Questions for discussion might include:

- What difficulties were experienced during the experiment?
- What conclusions could be drawn from your observations?
- Were you able to answer the researchable question?
- Did all groups get the same results?
- Is it possible to compare the data of various groups? Would it be useful to do so?
- What data proved most useful when comparing observations with different groups?
- How could this experiment be improved?
- Why is it necessary to have a large data set from which to draw conclusions?
- What is a science protocol?
- Why are science protocols necessary?

After the class discussion, students should complete the Summary Worksheet. Select examples of your student's work to submit for possible posting on the Forest Watch Web site.

A

Management Suggestions

Consider using this activity near the beginning of the school year. It could provide a foundation for many units, not just those in Forest Watch.

Seeds that germinate easily and rapidly, such as radish and clover, are best to use in this investigation. Seedlings should be allowed to grow for at least 2 weeks after the seeds have begun to germinate. Try Wisconsin Fast Plants at http://fastplants.cals.wisc.edu/main.html.

To set up the experiment, two sets, or varieties, of seeds should be used. Check science catalogs for availability, but some suggestions include:

- Ozone tolerant vs. intolerant varieties of the same plant, for example, tobacco.
- Monocotyledonous vs. dicotyledonous varieties such as corn and beans or radish and wheat.

Seeds that have been exposed to an environmental pollutant such as a 1% solution of chlorine bleach vs. a control, nontreated set.

BACKGROUND

Guiding Question

How do plant communities in the temperate deciduous forest change over time?

Succession is the regular, predictable change in plant communities in an area over a period of time. A community of plants becomes established in a particular area because it is well-suited or best adapted to the environmental conditions in that place. Over time, however, the organisms in the community begin to cause changes in their environment. For example, as plants grow tall they may shade out light to the ground and cause temperature and moisture levels to change. As conditions shift, other plants better adapted to the new environment may colonize the area and out-compete the original community. Eventually, a new community will be established. This process may continue until a climax or final community becomes established. The climax community is usually stable and self-reproducing. Biodiversity and amount of biomass changes as succession continues. Communities in the earliest and latest stages may have much less diversity than the intermediate phases.

Communities move through successional phases in response to changes in both their biotic (living) and abiotic (non-living) environment. Since we do not have the time to observe succession in one area over many years, we can instead observe various communities near your school occurring along a successional continuum from an old field to a forest community. We will measure several biotic and abiotic attributes of these communities.

Do you have an agricultural field or animal pasture that has been abandoned near your school? If you could observe that field for many years, you would notice many changes taking place. In the first year, low-growing annual and perennial plants such as ragweed and grasses will usually colonize the field. In the following year, perennial herbs such as goldenrod, milkweed and asters will become more important. By the third or fourth year, young woody shrubs and tree saplings will begin to grow among the herbs. These first woody plants are usually shade-intolerant species, meaning they only do well in sunlight conditions. Over the years, as the trees grow larger, the shade they create begins to change the light, temperature and moisture conditions in the soil below. As plants die and decay, the makeup of the soil itself changes as well. Many of the seeds from the perennial herbs and first generation trees are unable to survive in the shade. Seeds from new, shade-tolerant species, however, are able to germinate and survive in the darker conditions. Many of these shade tolerant herbs, shrubs and trees will be species of the eventual climax forest. The entire process may take a hundred years or more to reach maturity. Climax communities vary in different regions of the country and world depending on the local and regional conditions of the area.

In this exercise, biotic measurements will include some plant identification and sampling and measuring of plant populations and height of the dominant vegetation. Abiotic measurements will include light intensity, relative humidity, air temperature, soil temperature and wind speed. We can then compare how the conditions are different in each area and speculate how and why they might change over time, and how communities will be affected.

TERRESTRIAL SUCCESSION

A

ACTIVITY

Can you observe secondary terrestrial succession and differentiate between local plant communities in different stages of succession?

Materials

- 100 meter string with 10 meter increments
- light meter
- sling psychrometer
- air thermometer
- soil thermometer
- piece of string 1 meter in length, hula hoop or sampling frame
- plant field guide
- hand-held anemometer
- meter stick
- clinometer

Procedure

1. Your teacher has laid out a 100-meter transect which runs through several plant communities near your school. The transect is marked at 10-meter intervals. You will need to take measurements at each of the 11 stations on the transect.

2. Your teacher will assign each group of students certain types of measurements to take and instruct you in the use of each instrument.

3. Record the data in the table provided. Share all class data for each station.

Data Table. Abiotic and Biotic Conditions in Communities near School.

Station #	Light Intensity at 1 meter	Air T at surface and 1 meter	Soil T	Wind Speed	relative humidity	common vegetation (species)	Height of dominant vegetation
1							
2							
3							
4							
5							
6							
7							
8							
9							
10							
11							

Comments

Interpreting the Data

1. How many community types did you observe along the transect? Community types may include old field, shrub/sapling, and forest communities.

2. How does the height of the dominant vegetation change from the old field to the forest community? Use a piece of graph paper and draw a scale profile of the vegetation along the transect. Include both the horizontal and vertical changes you observed. Label the community types.

3. Biomass is the measure of the amount of matter stored in the vegetation. Which community type stores the most biomass in the vegetation over long periods of time? Which has the greatest rate of turnover of biomass from year to year?

4. Is there a difference between air temperature near the ground and at one meter above the ground at the various stations? Are these stations associated with any particular community type?

5. Is there a difference in relative humidity between the old field and forest communities? What factor(s) might account for this difference? Does any of the other data you collected support your hypothesis?

6. Compare light intensity at the various stations. How does light intensity affect the types of organisms that can live in an area?

7. Describe how the dominant vegetation types change from community to community.

Communicating Your Results

Characterize the various communities that you have observed along the transect in terms of biotic and abiotic factors. What are the characteristics of communities in the early phases of succession? How does the diversity and amount of biomass held in the vegetation change as the area approaches the climax stages?

TERRESTRIAL SUCCESSION ACTIVITY

A

TEACHER RESOURCE

Introduction

In this activity, students will become familiar with the basics of secondary succession typical of an old field community in mid-latitude forests. The focus of this activity is to introduce the scientific concepts of change, systems, scale and sampling. Students are required to observe and measure, using various scientific instruments, characteristics of successional communities near their school. One of these communities, the forest community, will ultimately become the site of the Forest Watch study. Eventually, students may also begin to study some of the old field species, for example, milkweed and large-toothed aster, for the effects of ozone pollution. Students will become familiar with some of the tools they will be using in the Forest Watch protocols and will be introduced to the idea of precision and accuracy in measurement and data recording.

Guiding Question

Can you observe secondary terrestrial succession and differentiate between local plant communities in different stages of succession?

Materials

- 100-meter nylon rope with 10-meter increments
- sling psychrometer
- soil thermometer
- light meter
- air thermometer
- plant field guide
- piece of string 1 meter in length, hula hoop or sampling frame
- hand-held anemometer
- meter stick
- clinometer

Background

Much of New England is in various stages of secondary succession. One hundred years ago as much as 90 percent of New England was cleared of forests for grazing and agriculture. Today, New Hampshire, for example, is now covered with forests over 85% of the state.

The process of reforestation from field to forest is called secondary terrestrial succession. Succession is all about change, which makes it a great vehicle to study that scientific concept. This activity can be completed in conjunction with a unit on ecology in a biology, earth science, or environmental science class. The instructor will need to add support materials on communities and ecosystems and should review the correct technique for the use of the instruments and data collection.

Management Suggestions

Depending on your depth of interest and time available, plan on spending at least a couple of days before the lab introducing the idea of terrestrial succession in forest ecosystems. This can be accomplished through the use of science videos, overhead diagrams, notes, and slides or photos showing local examples.

Most of the instruments (soil thermometer, simple anemometer, light meter, air temperature thermometer) are fairly easy to use following the manufacturers' instructions. A psychrometer can easily be constructed by fitting a gauze sleeve over the bulb of a lab thermometer. Tape this thermometer to a gauzeless one. After wetting the gauze, the student can carefully wave the instrument until the "wet" bulb temperature has dropped to its lowest point. Record the dry bulb and wet bulb temperatures and use this information to calculate relative humidity. You will need a relative humidity table available in some biology and meteorology texts.

The height of the vegetation in the old field and young shrub communities can be measured using a meter stick. If you are unfamiliar with the use of the clinometer, consult the Measuring Tree Height Protocol or have the students estimate height by using their thumb or a pencil. For example, have a student whose height is known stand next to the tree. A second student should walk away from the tree stopping every once in a while raising a thumb or pencil to eye level and then comparing the height of the student to the thumb or pencil. When his/her thumb is the same "height" as the student next to the tree, measure how many thumbs from the base of the tree to the top and, using the actual height of the student, estimate the tree height.

Use the 1-meter string, hula hoop or sampling frame to sample the area directly below the flag or knot at each station for numbers and kinds of plants growing there. If you are using the string, tie one end of the string to the transect line and sweep a circle as the area to be sampled. Have the students identify the actual species, or just the type of plants (grasses, herbs, shrub, etc.) in the sample area. They can estimate the dominant species or type by observation or actual counting and comparing. Students will soon find that counting all the grass plants in the plot is time consuming and difficult to do! Have them brainstorm ideas about how to estimate the number of individuals in the plot by counting only a small section of the area. Additionally, students can calculate the area of the sampling plots and the density of the various plants within the plot.

The data collection can be done by groups in a variety of ways. One way would be to have one group do all the height measurements, another do soil and air temperature, another do wind speed and common vegetation, etc., sharing the class data among them. Or, ask each group to do all the measurements. This can only be done if there are enough instruments available for each group to have one of each. In addition, have groups begin at different stations so as to avoid crowding.

Additional Suggestions and Variations

The succession activity can be modified in several ways to emphasize particular science themes and lengthen the scope of study. It can be also be adapted for use with younger science students.

Activity—Investigating a Plot Over Time

Introduction

This activity is designed to continue throughout the school year and, in addition to observing differences and change between communities at one particular point in time, gives the students a chance to observe change through the seasons. The same plots could even be maintained and observed for several years by different groups of students. Students will measure and delineate sample plots at stations along the 100-meter transect (see Succession Activity). They will then make a series of observations and measurements on a regular basis throughout the school year. The types of measurements to be taken can vary according to the availability of instruments and time.

Materials (Per team of students)

In addition to the instruments suggested in the Terrestrial Succession Activity:

- 4 stakes
- 1 compass
- 4 meter sticks
- 1 84.9 cm piece of string
- 1 260 cm piece of string
- soil pH meter
- soil moisture indicator

Procedure

In the classroom

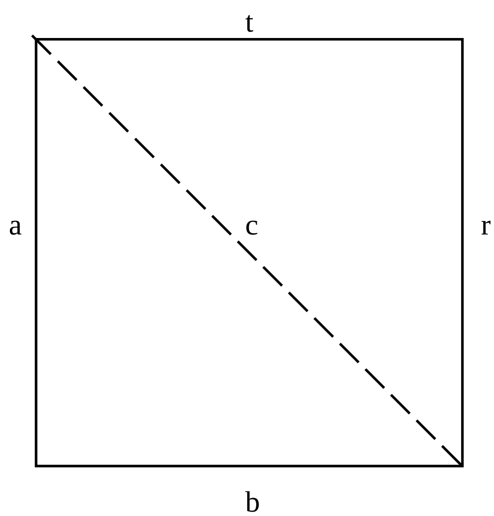
Introduce the concept of change and the topic of terrestrial succession. Generate a discussion about how the students might go about documenting change in plant communities near their school. Ask students what kinds of information they might collect and what tools would be necessary to study and compare the different plant communities. How large a study area would they need? How often should data be recorded?

At this point, you could introduce an optional approach to setting up and collecting data that may allow the students to experiment and determine on their own the importance of controlled, standardized collection methods. Instead of presenting the students with a detailed set of procedures as to what size and shape plot they should use and which types of data they could collect, provide an array of measuring tools and instruments and allow each group to choose their own. After the first collection period, assuming students have used different sized plots and are focusing on different types of data, you can ask them how and if it is possible to compare the plots to one another and why that is important. How did they decide how large to make the plot? How did they determine which parameters were important to study? Brainstorm, as a group, possibilities for better coordination of the data collection procedure that will allow students to make valid comparisons between their plots. Have the students agree, with your guidance, upon the size of the plot and what data to collect. They can then return to the field and set up new, more precise sampling areas.

To help students set up a very accurate square plot review the Pythagorean Theorem. Explain to the students that there is a method that can be used to standardize setting up the field study plots. This will help us set up a more accurate PSSP (Pixel-Sized Sampling Plot) for our Forest Watch Project.

Introduce the Pythagorean Theorem: $a^2 + b^2 = c^2$. Refer to the Establishing the Pixel-Sized Sampling Plot (PSSP) Protocol for more information.

Draw a square box on the board and label the sides **a**, **b**, **r**, and **t**. Draw a dotted line from the vertex of – **at** to the vertex of **br** and label this **c** as illustrated below:



If we know that each side of the study plot needs to be exactly 60 cm, then what formula could we use to help figure out the length of **c**?

($a^2 + b^2 = c^2$)

Given the length of **a** and the length of **b** = 60 cm

Calculate: 60 cm x 60 cm = 3600 cm^2

3600 cm^2 + 3600 cm^2 = 7200 cm^2

The square root of 7200 cm^2 = 84.9 cm

Setting up the plot

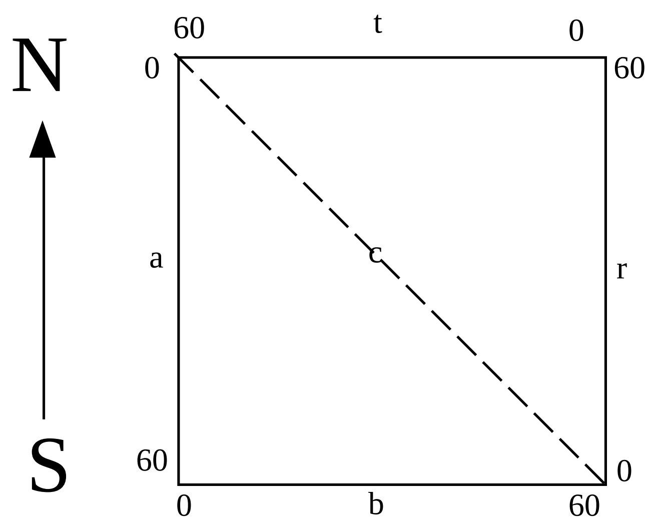
You may wish to review the procedure for setting up the PSSP in the Site Selection protocol. This process is similar.

In the field

Using the compass, determine N, S, E & W with your students. One side of the box should run N-S, as this is common in forestry studies. Lay a meter stick along this N-S line. Insert a stake at both the 0 cm and the 60 cm marks. Explain that this will represent **side a** of our square study plot.

Lay a second meter stick perpendicular to the first at the 60 cm mark of **side a**. This will represent **side b**. Hold the 84.9 cm piece of string at the 0 cm end of **side a**. This piece of string will represent the **hypotenuse c**. Stretch the end of the hypotenuse string to the 60 cm end of the **side b** meter stick. Insert a stake where they meet.

Lay a third meter stick perpendicular to **side b** and parallel to **side a**. This represents **side r**. Be sure the 0 cm mark on **side r** touches the 60 cm mark on **side b**. Take the 84.9 cm string and lay it from **–**ab to the 60 cm mark on **side r**. Where the two meet, put in a fourth stake. This creates **side t**. Check the accuracy of this stake by measuring the length of **side t**. The distance should be exactly 60 cm.



Working in teams, students are asked to identify, count and record each species of plant within his/her Field Study Plot (FSP). In addition, soil temperature, air temperature, soil pH, light intensity and soil moisture readings may be recorded. Students should revisit and record at the FSP at least three times throughout the year, for example, early fall, late fall and spring. More visits would be ideal.

Students could graph the data using any illustration from simple bar graphs to stem and leaf plots, to box and whisker graphs. (See Math Activities.)

Analysis

1. Create a bar graph to illustrate the number of each species of plant from each sampling period. Were there any changes in the number of plants of each species? How about in the number of overall plants within your Field Sampling Plot (FSP)?
2. Did any abiotic factors change within your FSP?
3. Compare your abiotic factors with other teams within the class. Were there any differences between the FSPs for the same sampling period? Were there any differences between the FSP's over the course of the year?
4. If so, what might account for those changes?
5. What further investigations would you suggest for next year? Why?

FIELD SCIENTISTS: ______________________________

____

____

____

DATE:________________ DATE:________________

Data Table. Investigating a Plot Over Time.

Plant Species	# of plants	soil temp.	soil pH	light intensity	soil moisture	# of plants	soil temp.	soil pH	light intensity	soil moisture

Observations and evidence of animals and other organisms:

Date Observation

Activity—Adopt a Tree

In cases where the succession lab and the year-long sample plots are not feasible to complete, another activity involving the observation of change in the natural world is to observe a tree as it changes over the course of the school year. Have students choose a tree and make periodic observations of it. The observations might be entered into a journal and could include pictures and photographs, collected leaves, observations of animals and insects, human influences, major changes such a leaf drop and bud production, poems, etc.

Activity—Chlorophyll Extraction in the Fall and Spring in Leaves such as the Sugar Maple.

Consult the Spectrophotometry Optional Protocol for the white pine adapting it to maple or other leaves that change color in the late fall. Compare the results to the photosynthetic green leaves of early fall or late spring.

INTRODUCTION TO REPRESENTATIVE SAMPLING

ACTIVITY

Can you determine a realistic, representative sample of the forest ecosystem to help assess forest health in your area?

Is it possible to assess the entire forest in the satellite image of your community?

Why do we study only one 30 m x 30 m sampling plot (PSSP)?

How many trees are being monitored by Forest Watch Students? Is this enough?

Procedure

With the guidance of your teacher, work as a team to answer and discuss the following questions.

1. Observe the Landsat scene of your state. Suppose you wanted to study the entire state to determine how many wetlands there are. What are some ways that you could accomplish this goal?

2. Observe the 512 pixel x 512 pixel Landsat Image of your local area. What information do you need in order to calculate the area of the image in km^2?

3. Calculate the size (area covered) of this image in km^2?

4. Suppose that you've decided to study the forested areas in your satellite image. Estimate the percentage of your image that is covered by forest.

   a. How could you determine the area, in km^2, of your estimated forest cover? Show your calculation below.

   Estimated area of your image covered by forest: _________________ km^2

5. Let's say that to begin your study you decide to set up study areas, or plots, within the forested region on the image. What shapes would you choose for your study plots?

   a. What plot shape might be the easiest to set up and why?

6. Suppose that you decide to use square plots that are 100 m^2 for your study. What are the dimensions of these plots?

7. How many 100 m^2 plots would it take to study the forested area of your satellite image? Begin by writing down the area in km^2 covered by forest. (You determined this question #4.)

Sample Calculations: Let's say that 75% of your satellite is covered by forest. Therefore, the area covered by forest is 177 km^2. Because our plots are 100 m^2 and the area has been determined in km^2, we must convert them to the same units. Therefore, convert 177 km^2 into m^2:

$$177\ km^2 \times \frac{1{,}000{,}000\ m^2}{1\ km^2} = 177{,}000{,}000\ m^2$$

Now find the number of 100m^2 plots that can be established in 177,000,000 m^2 area.

$$\frac{177{,}000{,}000\ m^2}{100\ m^2} = 1{,}770{,}000 \text{ plots}$$

Wow, you've got work to do!!

   a. Now calculate the number of plots that would be found in your satellite image using your own estimation of the area covered by forest. (See question 4) Show your work below.

8. Make a list of the kinds of information (data) you will collect in your study area?

9. How long do think it will take you to complete the study of one 100 m^2 plot? ___________

   a. Suppose you decide that it will take one hour to study one 100 m^2 plot. How many plots could you study in a normal work week?

   b. How many weeks are you willing to work in a year?

   c. Given the number of plots you have to study, how long would it take you to complete your study? Show your work.

10. To make the work go faster, you decide to hire some field researchers to help you with the study. How many researchers do you need and how much money will they want to make for their work in this study? Who's going to provide the funding for your study?

11. Let's be more realistic about the amount of our town represented in the satellite image that we are going to study. Suppose that we decide to study only 1/1000 of our town or .001.

    a. How much area is covered by .001 of your satellite image?

    b. How many plots could you set up in this area?

    c. How long would it take you to study just this .001 of your image? (Use the same 1 hr/plot and 40 plots / week schedule)?

12. Observe your image again. Estimate how much of the image is covered by .001.

    a. Now, how will you decide which .001 of the image to actually study?

    b. How will you distribute your study plots across the image?

    c. Will this study site be a representative sample of the entire area represented by the satellite image?

    d. Will your data truly reflect what someone would expect to find in a typical forested area of your town? Why or why not?

13. In the Forest Watch study, each school sets up a sample plot called a pixel-sized sampling plot or PSSP. Needle samples are collected from 5 white pine trees in each plot.

    a. How does the size of the sample plot that you will use in Forest Watch compare to the 100 m^2 plots described above?

    b. How might this influence the study?

    c. Ask your teacher for the number of schools involved in this study and map their locations.

    d. Calculate the total number of trees involved in the study.

    e. How are the sample plots distributed in the New England region?

    f. Do you think that we are sampling from enough trees?

    g. Do you think that the number of schools setting up plots and sampling from trees in their plots is enough to give a representative sample of New England forests? Why or why not?

14. Why or how is representative sampling important in scientific investigation?

15. As a research scientist, what are some of the limitations that you must consider in your efforts to obtain a representative sample?

INTRODUCTION TO REPRESENTATIVE SAMPLING

A

TEACHER RESOURCE

Introduction

This activity uses a local satellite image to focus students on the science concepts of sampling, scale and modeling. Through a teacher-guided question and discussion sequence, students discover the importance of representative sampling and how to choose a realistic plot size in a forest ecosystem. Students begin to understand that research goals often have limitations and that investigators must try to minimize those limitations through careful choice of sampling criteria and sampling plot size. This exercise provides an opportunity for students to become more familiar with the satellite image of their local area and to practice math skills.

Guiding Question

Can you determine a realistic, representative sample of the forest ecosystem to help determine forest health in your area?

Materials

- computer
- calculators
- worksheet
- local 512 x 512 pixel Satellite
- MultiSpec Program
- Landsat scene poster

Note: You may complete this activity without the computer and MultiSpec Program

Management Suggestions

Consider working through this activity as a class, discussing the various ideas as you proceed. Students may work with partners or in teams. This activity may be completed in a class period depending on your students' math skills.

Answer Key

1. Observe the Landsat scene of your state. Suppose you wanted to study the entire state to determine how many wetlands there are. What are some ways that could you accomplish this goal?

   Answers will vary. Possibilities for studying large areas include using a helicopter, satellite images, aerial photography, driving across the state in a grid, among others.

2. Observe the 512 pixel x 512 pixel Landsat Image of your local area. What information do you need in order to calculate the area of the image in km^2?

   You need to know how large each pixel is. For a TM image the pixel is 30 m x 30 m.

3. Calculate the size (area covered) of this image in km^2?

   For one side:

   There are 512 pixels across

   1 pixel is 30 meters on a side

   512 pixels x 30 m/pixel = 15,360 m

   15,360 m = 15.36 km

   Therefore, each side of the image is 15.36 km.

   The area covered by the image is then

   15.36 km x 15.36km = 236 km2

4. Suppose that you've decided to study the forested areas in your satellite image. Estimate the percentage of your image that is covered by forest.

   We will use 75% for this example.

   a. How could you determine the area, in $km^{2,}$ of your estimated forest cover? Show your calculation below.

   75% of 236 km^2 = 177 km^2

   50% = 118 km^2

   25% = 59 km^2

   10% = 24 km^2

   (Other percentages have been included if your students choose another percentage.)

   Estimated area of your image covered by forest: 177 km2

5. Let's say that to begin your research you decide to set up study areas, or plots, within the forested region on the image. What shapes could you choose for your study plots?

   Answers will vary. They may be rectangular, triangular, circular, square.

   a. What plot shape might be the easiest to set up and why?

   Answers will vary. For purposes of this example, use a square plot. Students may discuss how they could most easily set up a square plot.

6. Suppose that you decide to use square plots that are 100 m^2 for your study. What are the dimensions of these plots?

   10 m x 10 m

   (10 m on a side)

A

7. How many 100 m^2 plots would it take to study the forested area of your satellite image? Begin by writing down the area in km^2 covered by forest. (You determined this question #4.)

   Sample Calculations: *Let's say that 75% of your satellite is covered by forest. Therefore, the area covered by forest is 177 km2. Because our plots are 100m^2 and the area has been determined in km^2, we must convert them to the same units. Therefore convert 177 km^2 into m^2:*

$$177\ km^2 \times \frac{1{,}000{,}000\ m^2}{1 km^2} = 177{,}000{,}000\ m^2$$

   Now find the number of 100 m^2 plots that can be established in 177,000,000 m^2 area.

$$\frac{177{,}000{,}000\ m^2}{100\ m^2} = 1{,}770{,}000 \text{ plots}$$

Wow, you've got work to do![!]

   a. Now calculate the number of plots that would be found in your satellite image using your own estimation of the area covered by forest. (See question 4) Show your work below.

8. Make a list of the kinds of information (data) you will collect in your study area?

   Answers may include soil type, # of trees, different species of trees, wildlife signs, amount of rain-fall, light intensity, water quality, vegetation characteristics such as tree height, DBH, live crown height, etc.

9. How much time do you think it will take to complete a study of <u>one</u> 100 m^2 plot?

   Have students share some of their estimates of how long it might take. Remind them that it will depend on the type and amount of data that they choose to collect.

   Note: for the following sets of questions, refer to the introductory math activities for examples of converting units of measure.

   a. Suppose you decide that it will take one hour to study one 100 m^2 plot. How many plots could you study in a normal work week?

   Typical work week = 40 hrs, at 1 hour/plot means 40 plots could be studied in one week.

   b. How many weeks are you willing to work in year?

   Answers will vary. For this example let's say that they will work 48 weeks in a year, allowing for 4 weeks vacation. Generous for a beginning research position!

   c. Given the number of plots you have to study, how long would it take you to complete your study?

   Show your work below:

   With 4 weeks vacation per year :

$$1{,}770{,}00 \text{ plots} \left(\frac{1 \text{ week}}{40 \text{ plots}}\right)\left(\frac{1 \text{ year}}{48 \text{ weeks}}\right) = 922 \text{ yr.}$$

10. To make the work go faster, you decide to hire some field researchers to help you with the study. How many researchers do you need and how much money will they want to make for their work in this study? Who's going to provide the funding for your study?

    Answers will vary. Students should discuss this as a class. This might be an interesting opportunity for a discussion of bias in research (either accidental or on purpose) in relationship to where the funding originates. The kinds of research funded and supported by particular industries, for example, might only be those projects that would ultimately support the industry in some beneficial way.

11. Let's be more realistic about the amount of forest in the satellite image that we are going to study. Suppose that we decide to study only 1/1000 (or .001) of the forested region.

    a. How much area does .001 of your entire satellite image cover?

    .001 x 236 km2 = .236 km2

    b. How many plots could you set up in this area?

    $$.236\ km^2 \left(\frac{1{,}000{,}000\ m^2}{1\ km^2}\right)\left(\frac{1\ plot}{100\ m^2}\right) = 2{,}360\ plots$$

    c. How long would it take you to study just this .001 of your image? (Use the same 1 hr/plot, 40 plots /week, 48 weeks of work / year schedule)?

    $$2{,}360\ plots \left(\frac{1\ hour}{1\ plot}\right)\left(\frac{1\ week}{40\ hr}\right)\left(\frac{1\ year}{48\ weeks}\right) = 1.2\ years$$

12. Observe your image again. Estimate how much of the image is covered by .001.

    Students will need to decide on 1/1000 of the image.

    a. Now, how will you decide which .001 of the image to actually study?

    Students should try to decide what is most representative of the forest in the image knowing that they can only cover 1/1000 of the image for their study.

    b. How will you distribute your study plots across the image?

    Again, they should try to locate the 2,370 plots in areas that they think is most representative of the forested cover in the image.

    c. Will this study site be a representative sample of the entire area represented by the satellite image?

    Students may decide that they need to do some "ground truthing" before they decide where to locate their plots.

    d. Will your data truly reflect what someone would expect to find in a typical forested area of your town? Why or why not?

    Answers will vary.

13. In the Forest Watch study, each school sets up a sample plot called a pixel sized sampling plot or PSSP. Needle samples are collected from 5 white pine trees in each plot.

   a. How does the size of the sample plot that you will use in the Forest Watch study compare to the 100 m^2 plots described above?

   In Forest Watch, the plot size that we use is a 30 m x 30 m plot

   b. How might this influence the study?

   Smaller and fewer plots mean less area sampled.

   c. Ask your teacher for the number of schools involved in this study and map their locations.

   At this time there are about 92 schools that participate in the study.

   d. Calculate the total number of trees involved in the study.

   460 trees

   e. How are the sample plots distributed in the New England region?

   Students should refer to their maps that show the location of all the Forest Watch Schools

   f. Do you think that we are sampling from enough trees?

   Answers will vary.

   g. Do you think that the number of schools setting up plots and sampling from trees in their plots is enough to give a representative sample of New England forests? Why or why not?

14. Why or how is representative sampling important in scientific investigation?

   We need to have enough samples collected from a variety of areas that truly represent the object of our investigation. Obtaining a representative sample is a very complex thing to do. At best, researchers are able to get close to obtaining representative samples. Many limitations occur that prevent a "perfect" sample. It is important to be as unbiased in our sampling efforts as possible. For example, if we are assessing forest health, we want to sample a variety of trees on a continuum of healthy to unhealthy.

15. As a research scientist, what are some of the limitations that you must consider in your efforts to obtain a representative sample?

   Time, funding, access to the study sites, trained field researchers, etc.

AN INTRODUCTION TO TAXONOMY AND DICHOTOMOUS KEYS

A

BACKGROUND

Guiding Question

How can you identify organisms using a dichotomous key?

In these activities, you will learn how scientists use keys to help them identify organisms that have been classified. You will learn how to make a key and use a key to identify some of the trees that are found in the group called conifers. This group of trees includes the white pine. What kingdom does the white pine belong to?

Scientists refer to organisms by using either their common name, such as white pine or by using their scientific name, such as *Pinus strobus*. Each organism has only one scientific name, but could have many local common names. By using scientific names in their communication, scientists can be sure that they are talking about the same organism.

The scientific names of organisms are based on the classification system of living things. Binomial nomenclature is a system of naming species of organisms so that each species has a unique name. As the term implies, each species has two names. Together these two words are what we call the scientific name. The first word is the genus name. The genus name is capitalized and represents a large group. The second word is the species name. The species name is not capitalized and represents a small specific group. For example, the scientific name for humans is *Homo sapiens* (genus *Homo* and species *sapiens*). *Homo sapiens*, translated from Latin, means "Man, wise." All humans are found within this genus and species group. Humans and their closest relatives are found in the larger taxonomic family called the Hominidae. The Hominidae, in turn, are found in the taxonomic order called Primates. Grouping continues until the largest classification level called a kingdom is reached. What class, phylum and kingdom do humans belong too?

Classification is a way of separating a large group of closely related organisms into smaller subgroups. With a classification system, identification of an organism is easier. Classification teaches you how to organize things, store things and find things fast. Taxonomy is the science of classifying organisms. Scientists look at many characteristics of living things when they organize them into groups. These may include the structure, function, biochemistry, behavior, genetic systems, embryonic development, ecological interactions, nutritional patterns and evolutionary histories of organisms. Taxonomy is a branch of science that is constantly changing. As new information is discovered, taxonomists must review the categories into which they may have originally placed a specific organism.

A dichotomous key is a tool used by scientists to help determine or identify the scientific name of an organism. Scientists can develop very specific keys for the organisms they are interested in identifying in the field or, after collection, in the lab. Using a dichotomous key is based on making careful observations about organism traits and then comparing the observations to two choices in the key. Only one choice can be correct. By following several sets of choices moving from very general characteristics to very specific ones, an identification of the organism can be made. The identification can only be as good as the key itself. And, as stated above, scientists must be willing to change or modify keys as information about organisms changes with new discoveries.

CLASSIFYING SHOES AND MAKING DICHOTOMOUS KEYS

ACTIVITY

Materials

- shoes
- needle or leaf samples (or drawings)
- pencil and paper

Procedure

1. Each student in class should take off one shoe and put it into a pile in the center of the room.
2. Discuss with your classmates some of the similarities and differences you observe amongst the shoes.
3. Divide the shoes into two piles. Each pile is separated from the other by one feature. With your classmates and teacher, begin recording the shoe features in a class chart on the blackboard or overhead projector. Repeat the division and recording of differences until each pile is reduced to one shoe. Each time you end a line with a shoe, write the name of the shoe's owner. In the end, you will have constructed a key to the owners of the shoes in your original pile.
4. To test your key, take all the shoes and pile them together again. Ask someone from another room to come and randomly choose a shoe. Have them use your key to identify the owner!
5. Now you are ready to practice making a key on your own. The needles in the diagram on the accompanying page are all from different pine trees and drawn to actual size. Note that each bundle contains different numbers, types and lengths of needles. Design a key that will classify each tree. It's really a lot of fun!
6. Trade your key with a classmate to check its accuracy.

DESIGNING A DICHOTOMOUS KEY FOR CONIFER TREES

Name(s)

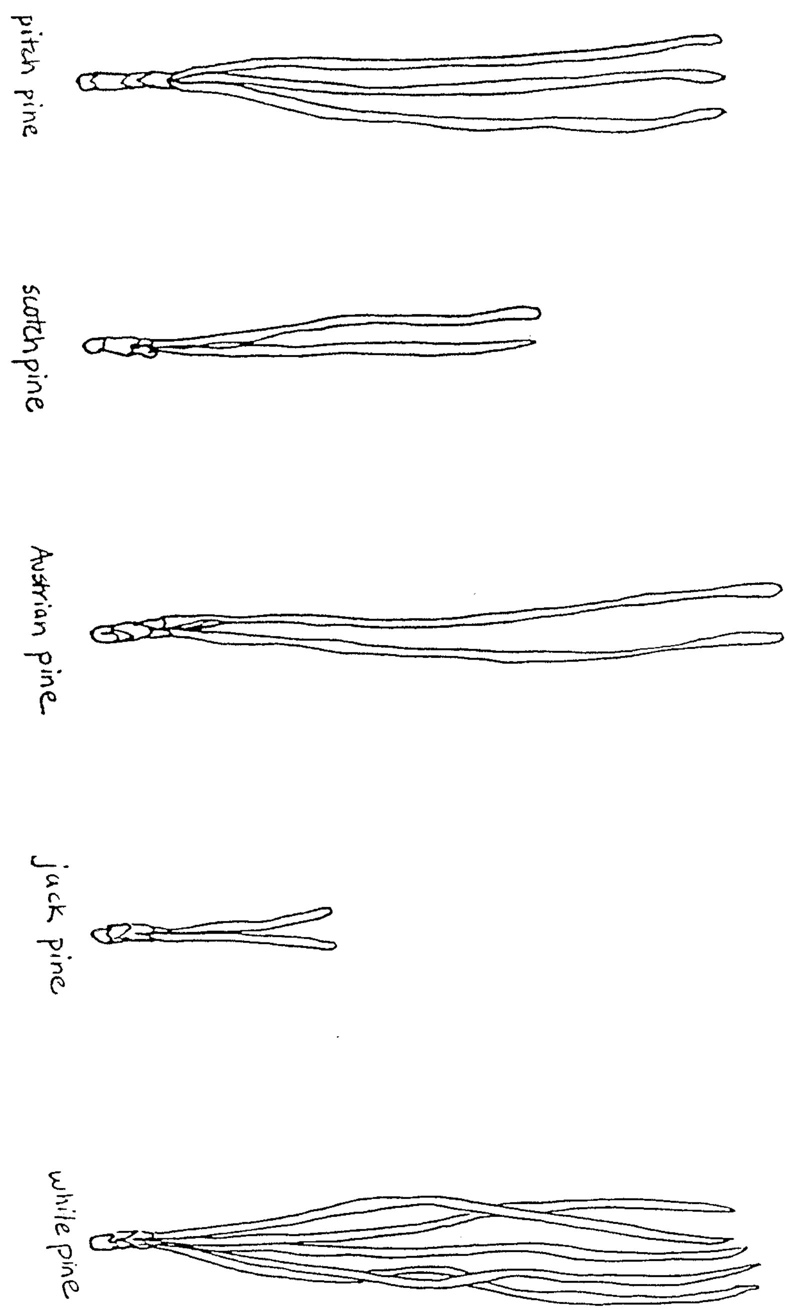
Date:

Activity 1:Using needle structure to make a key to some common conifer species.

1. Use the empty table below to start your dichotomous key for the five types of pine needles. Continue with as many choices as you need to identify every needle example.

   1. A. ....
      B. ....
   2. A. ....
      B. ....

2. Look at some of your classmates' dichotomous keys. Are they exactly like yours?
   Why or why not? Can different keys be developed to correctly identify the same organisms?


pitch pine
scotch pine
Austrian pine
jack pine
white pine

ACTIVITY

Materials

- variety of woody tree samples
- Key to Selected Woody Tree Species of New England

Procedure

The keys that you have constructed in Activity 1 are known as dichotomous keys. At each point in the key you must choose between two characteristics to describe some trait of an organism. Only one choice can be correct. In this activity, you will use a dichotomous key of selected woody tree species to help you identify the white pine trees that are the subject of our study.

1. Obtain a copy of the tree key from your teacher.
2. Your teacher will arrange several leaf or needle branch samples around the tables in your classroom. Use the key to identify at least three of the leaf and needle samples. In the spaces on the student worksheet, write down each of the dichotomous key choices you and your partner select on your way to finally identifying the scientific name of your tree. Place the name of the tree in the last blank space on the worksheet. When you are done, tape one of the leaves or needles onto the sheet next to its dichotomous key choices. Remember, each time you come to a new sample you should start at number 1 and move from there as the directions instruct.
3. Once you feel comfortable using the key, you are ready to identify some of the white pine trees on your PSSP and choose which 5 trees you will sample. You may be interested in identifying some of the other trees that occur on your PSSP in addition to white pine.

Using a Key to Identify Trees

Look at the example dichotomous key called the Key to Selected Woody Tree Species of New England.

You will be identifying at least three different tree/leaf samples. Everybody must classify tree/leaf sample #1. Then select two other tree samples to classify. Use the empty line grids below to record your three sets of classification selection choices

1...

Tape sample here

....

....

....

....

....

1...

Tape sample here

....

....

....

....

....

1...

Tape sample here

....

....

....

....

....

A Key to Selected Woody Tree Species of New England

The following dichotomous key to woody tree species commonly found in New Hampshire is meant to serve as an introduction for students. The key is not intended to be inclusive of all species in an area. It is designed to be a non-technical tool for showing students how to use such a key. References to several keys are given at the end of this simple key.

1. A. Leaves needle-like, scale-like, or narrow and flattened, less than 4 mm wide, Gymnosperms 2
   B. Leaves broad, flattened, more than 4 mm wide, deciduous, Angiosperms. 8
2. A. Leaves needle-like or narrow and flattened 3
   B. Leaves scale-like .... White cedar (*Thuja* sp.)
3. A. Leaves needle-like 4
   B. Leave narrow and flattened. 7
4. A. Needles in groups of 2-5 5
   B. Needles attached singly. .... Spruce (*Picea* spp.)
5. A. Needles in groups of 2-3 6
   B. Needles in groups of 5, 60-100 mm, soft and thin .... White pine (*Pinus strobus*)
6. A. Needles in groups of 2, 90-160 mm, soft. .... Red pine (*Pinus resinosa*)
   B. Needles in groups of 3, 70-120 mm, stiff .... Pitch pine (*Pinus rigida*)
7. A. Needles less than 10 mm long, on short stalks .... Hemlock (*Tsuga* spp.)
   B. Needles over 10 mm, not on stalks. .... Fir (*Abies* spp.)
8. A. Leaves simple 9
   B. Leaves compound 15
9. A. Leaves opposite on branch 10
   B. Leaves alternate on branch 11
10. A. Leaves lobed .... Maple (*Acer* spp.)
    B. Leaves with small teeth. .... Arrow wood (*Viburnum* spp.)
11. A. Leaves lobed 12
    B. Leaves with teeth 13
12. A. Lobes pointed .... Red Oak (*Quercus rubra*)
    B. Lobes rounded .... White Oak (*Quercus alba*)
13. A. Leaf stalks round 14
    B. Leaf stalks flattened .... Poplar (*Populus* spp.)
14. A. Buds long, sharp .... Beech (*Fagus* spp.)
    B. Buds short, not sharp .... Birch (*Betula* spp.)
15. A. Leaves opposite on branch .... Ash (*Fraxinus* spp.)
    B. Leaves alternate on branch 16
16. A. Leaflets few (5-7) .... Hickory (*Carya* spp.)
    B. Leaflets many (9 or more) .... Walnut (*Juglans* spp.)

Technical Keys to Common Species

Batson, W.T. 1975. Genera of the Eastern Plants. Published by the author, Printed by the State Printing Co., Columbia, SC.

Grimm, W.C. 1983. The Illustrated Book of Trees: the comprehensive guide to more than 700 trees of Eastern North America. Stackpole Books, Harrisburg, PA.

Sargent, C.S. 1965. Manual of the Trees of North America. 2 vols. Dover Paperback Publications, Inc., NY.

Watts, M.T. 1963. Tree Finder. Nature Study Guild, USA.

AN INTRODUCTION TO TAXONOMY AND DICHOTOMOUS KEYS

A

TEACHER RESOURCE

Introduction

In these activities, students will learn how scientists use keys to help them identify organisms that have been classified. Students will learn how to construct and use a dichotomous key. They will use a key to identify some of the trees that are found in the group called conifers. This group of trees includes the white pine.

Materials

- paper/pencil
- shoes
- Key to Selected Woody Tree Species of New England

Background

Scientists refer to organisms by using either their common name, such as white pine or by using their scientific name, such as *Pinus strobus*. Each organism has only one scientific name, but could have many local common names. By using scientific names in their communication, scientists can be sure that they are talking about the same organism.

The scientific names of organisms are based on the classification systems of living things. Binomial nomenclature is a system of naming species of organisms so that each species has a unique name. As the term implies, each species has two names. Together these two words are what we call the scientific name. The first word is the genus name. The genus name is capitalized and represents a large group. The second word is the species name. The species name is not capitalized and represents a small specific group. For example, the scientific name for humans is *Homo sapiens* (genus *Homo* and species *sapiens*). *Homo sapiens*, translated from Latin, means "Man, wise." All humans are found within this genus and species group. Humans and their closest relatives are found in the larger taxonomic family called the Hominidae. The Hominidae, in turn, are found in the taxonomic order called Primates. Grouping continues until the largest classification level called a kingdom is reached. What class, phylum and kingdom do humans belong to?

Classification is a way of separating a large group of closely related organisms into smaller subgroups. With a classification system, identification of an organism is easier. Classification teaches you how to organize things, store things and find things fast. Taxonomy is the science of classifying organisms. Scientists look at many characteristics of living things when they organize them into groups. These may include the structure, function, biochemistry, behavior, genetic systems, embryonic development, ecological interactions, nutritional patterns and evolutionary histories of organisms. Taxonomy is a branch of science that is constantly changing. As new information is discovered, taxonomists must rethink the categories into which they may have originally placed a specific organism.

A dichotomous key is a tool used by scientists to help determine or identify the scientific name of an organism. Scientists can develop very specific keys for the organisms they are interested in identifying in the field or, after collection, in the lab. Using a dichotomous key is based on making careful observations about organism traits and then comparing the observations to two choices in the key. Only one choice can be correct. By following several sets of choices moving from very general characteristics to very specific ones, an identification of the organism can be made. The identification can only be as good as the key itself. And, as stated above, scientists must be willing to change or modify keys as information about organisms changes with new discoveries.

Management Suggestions

One way to do the shoe activity is with the entire class being guided by the teacher in the development of the dichotomous key. Use the board or overhead to create one master key based on student input.

Student-generated needle keys may vary somewhat from each other, but they must use the dichotomous approach as illustrated in the student example for shoe identification. Students may consult the Key to Selected Woody Tree Species of New England to help them set up the wording for their keys. Have metric rulers available so students may incorporate measurement as one of their traits if they so desire.

Time needed to complete activities:

Classifying Shoes and Making a Key: 25-30 minutes.

Using a Key to Identify Common Tree Species in New England: 30-35 minutes.

Answer Key

Classifying Shoes and Making a Key

1. An Example of how a key might be set up is found below.
    1. A. Shoes with laces .... .go to 2
       B. Shoes with no laces.. go to 4
    2. A. Brown shoes. .... go to 3
       B. Not brown .... go to 6
    3. A. Lt. Brown shoes... Brian
       B. Suede shoes.. Sandy
    4. A. Boots .... Trina
       B. Not boots... go to 5
    5. A. Flat .... Gail
       B. High .... Lori
    6. A. Blue .... Rob
       B. White .... Peggy

A

2. Student-made dichotomous keys of the five conifer needle samples may all be different. Differences do not make one key right or wrong. What you should look for is the pattern of two distinct choices provided for each observed trait and that every sample can be identified easily and correctly by using only one pathway on the key. See the example below.

1. A. Needles in groups of 2 .... go to 2
   B. Needles in Groups of 3 or more .... go to 3
2. A. Needles 40 mm or less .... Jack Pine
   B. Needles greater than 40 mm .... go to 4
3. A. Needles in groups of 3 .... Pitch Pine
   B. Needles in groups of 5 .... White Pine
4. A. Needles 60 mm long .... Scotch Pine
   B. Needles 80 mm long .... .Austrian Pine

Using a Key to Identify Tree

All students need to be able to identify the white pine for the Forest Watch study. Therefore, one of the tree samples (e.g. sample #1) that your students identify in this activity should be from a white pine. The remainder of the answers will vary based on the types of tree samples that you have displayed in the lab. Use the Key to Selected Tree Species of New England to help you visualize the pattern. Below is an example of how the students should record their answers.

1.A. Leaves needle-like, scale-like, or narrow and flattened, less than 4 mm wide, ...2

Gymnospermae

2. A. Leaves needle-like or narrow and flattened...3 Tape sample here
3. A. Leaves needle-like...4
4. A. Needles in groups of 2-5....5
5. A. Needles in groups of 5, 80-130 mm, soft...White Pine (*Pinus strobus*)

Note: You may want to add descriptive diagrams to accompany this key. For instance, you can include leaf form examples e.g., simple vs. compound, alternate vs. opposite and others.

For Further Study

Have students set up dichotomous keys for other groups of objects, e.g., their video or audio tapes or CD's, cell types in the body, rocks or minerals.

Have students use dichotomous keys to identify and name other organisms that you may have collections of in the lab, e.g., insects, wildflowers, and weeds.

The activity Creating Imaginary Organisms and Making a Key to Identify Them could be used to extend the idea of using keys to identify organisms.

CREATING IMAGINARY ORGANISMS AND MAKING A KEY TO IDENTIFY THEM

ACTIVITY

Imagine that your class is part of a research team of scientists on an expedition to the rainforest. During the expedition organisms that have never been identified are discovered. Before the organisms can be named, their characteristics must be described. Each student group will create, describe and draw four imaginary organisms.

Procedure

For each organism your group will:

1. Complete a “New Species Report”.

   Use the New Species Report pages provided by your teacher to identify the characteristics or traits of your newly discovered organism. Be as creative as you can. You should create four new and different species.

2. Make a colored picture or three-dimensional model of each organism.

   Use the characteristics that you listed in your New Species Report to draw and label a diagram picture of each new organism. If you have the materials, you might make a three-dimensional creature.

3. Determine a unique common name and scientific name for each organism.

   Both common and scientific names are often designed to be descriptive of the organism. Use an interesting or unusual characteristic of your new organism to help you name it. Look at examples of Latin names to help you find one that fits. Remember, a scientific name has two parts: genus and species. You could place some of your creatures in the same genus if they have some, but not all, characteristics in common.

4. Make a dichotomous key to help other scientists identify the new organisms.

New Species Report

SUBMITTED BY _________________

SPECIMEN #__________________

You may wish to name your new organism after you describe it using the guidelines in numbers 1-7 below.

COMMON NAME OF ORGANISM ________________

SCIENTIFIC NAME OF ORGANISM ________________

1. To which kingdom (plant, animal, protist, fungi, monera) does the organism belong?

2. What characteristics does the organism have that make it a member of this kingdom?

3. What characteristics does the organism have for survival in the rainforest? (Ex. A caterpillar may be green to blend in with the vegetation)

4. How does the organism get its food? What characteristics does it have for food-getting?

5. How does the organism protect itself?

6. What characteristics would be most helpful in identifying the organism?

7. Morphology of the organism:
   a. size and shape

   b. body regions (ex. head, thorax, abdomen)

   c. appendages for movement

   d. type of body covering

   e. Sensory organs (organs for smell, sight, touch, taste, pain)

Labeled Diagram Showing Important Characteristics

SPECIMEN # _______

CREATING IMAGINARY ORGANISMS AND MAKING A KEY TO IDENTIFY THEM

TEACHER RESOURCE

Introduction

This activity will help students identify characteristics that can be used to classify and identify organisms, introduce the student to binomial nomenclature, and give the student an opportunity to practice making and using dichotomous keys.

Guiding Questions

Can you create and draw an Imaginary Organism?
Can you use binomial nomenclature?
Can you make a dichotomous key?
Can you use a dichotomous key to identify organisms?

Materials

- Preliminary Report form for each student team
- Drawing paper
- Pencils, crayons, markers

Management Suggestions

A discussion of binomial nomenclature and instructions for how to construct a dichotomous key can be found in the background section for the activity called An Introduction to Taxonomy and Dichotomous Keys. Students will have fun making up Latin-sounding names. Or use real Latin names from a dictionary or biology text that provides definitions for terms.

Students will be asked to imagine that they are part of a science research team that has just discovered four organisms in the rainforest that have never been identified and classified. The students will be asked to describe and draw the imaginary organisms that they have created. For each organism each group will:

- Complete a “New Species Report”.
- Make a colored picture or three-dimensional model of the organism.
- Determine a unique common name and scientific name for each organism.
- Make a dichotomous key to help other scientists identify the new organisms.

Divide the class into groups of convenient size. Each group will create four new organisms. Determine the total number of organisms that will be created so that you can prenumber all the organism sheets before you hand them out to the students. Remind the students that they should not put the real name of the organism on the labeled diagram sheet otherwise you won’t be able to use these sheets as unknowns from which to develop dichotomous keys. When the students have completed their creatures, go around the room and make a master list of specimen number and name. You can use this list later to check and verify keys. Until the end, students will only know the names and numbers of their own creations.

Tack the completed diagrams as a large group on a bulletin board so that all groups can clearly see them. Have the students create, either individually or in their groups, a dichotomous key to the new organisms. Trade the keys around and randomly choose one or more of the organisms for the students to identify with someone else’s key. Continue trading and evaluating keys as long as you want!

NEW SPECIES IDENTIFICATION REPORT

Teacher Master Key to New Organisms

SPECIMEN # SCIENTIFIC NAME/COMMON NAME

FOREST WATCH

STUDY GUIDE

MATH ACTIVITIES

Contents

ACCURACY OF MEASUREMENT

ACTIVITY

Background

Accuracy is important when you are measuring things. Machines, such as cars and wristwatches, work better when their parts are constructed using very accurate measurements. Measurements taken during scientific investigations need to be as precise and accurate as possible. Accuracy in measurement helps insure that products and results will be reliable. In this activity, you will learn some tips and tools that will help you be accurate in reporting your measurements and calculations.

Rounding of Numbers and Accuracy

Example:

In order to find the diameter of a tree, you must use a tape measure to find the circumference of the trunk. Suppose you record the circumference as 63.2 cm. This number, 63.2, has 3-digit accuracy. The following measurements all have 3-digit accuracy:

0.632 meters
632 mm
632,000,000 nanometers

The zeroes in 632,000,000 are called "place holders" and do not mean the measurement is any more accurate than to 3-digits.

Use the formula $C = \pi d$, or $d = C/\pi$, to find the diameter of a tree with a circumference equal to 63.2 cm. You can use the π key on your calculator. Carry all the decimal places on your calculator until the end of any problem. Once you reach the end of the problem, remember that you still have 3-digit accuracy. So even if you use 3.141592654 for the value of π, in the end, you have to round your answer to 3-digit accuracy; the accuracy of your actual measurement.

Here's how it should look:

$$d = 63.2/\pi = 20.11718481 \approx 20.1 \text{ cm}$$

20.1 cm is the reported diameter.

How do you decide what level of accuracy you should use in reporting of measurements?

First you need to determine your least accurate measurement. Suppose you want to find the average of the following five tree circumferences that were measured in centimeters.

88.6, 75.7, 88, 65.3, 79.6

(88.6 + 75.7 + 88 + 65.3 + 79.6) / 5 = 79.44 cm

Now you have to think about the accuracy of your answer. If the 88 cm measurement was measured to the nearest centimeter rather than the nearest millimeter, it is your least accurate measurement in the data set. This is true even though some of the other measures are accurate to millimeter measurements. Since your least accurate measurement determines the overall accuracy of your data set, your answer must be reported as 79 cm.

Suppose you had measured the tree and found the circumference to be 88 cm and no millimeters. If you had recorded your measurement as 88.0 cm, you are indicating that your measurement is accurate to the nearest millimeter and that you checked to see if there were any mm beyond the 88 cm mark. The “0” in this case is not a place holder. It indicates the accuracy of the measurement to millimeters.

If students take many measurements of needle fascicle length with a metric ruler, to the nearest millimeter, you can average those measurements and report your results to the nearest tenth of a centimeter. This is true even if some measurements have 4 digits and others have 3. Look at the example below.

99.3, 99.8, 102.4, 103.2, 101.1

(99.3 + 99.8 + 102.4 + 103.2 + 101.1)/5 = 101.16 ≈ 101.2 cm

Summary

- Round numbers, such as averages, to the least accurate measurement you used when the measurements were collected for your data set.
- When computing numbers to a designated degree of accuracy (2 digit, 3 digit, etc.), round your answer to degree of accuracy <u>after you complete</u> your computation.

Exercises

1. Why is it important to measure accurately? Provide an example of a case where accuracy in measurement is critical, for example, can you think of some application in the field of medicine, school grading or clothing manufacturing.

2. How would you indicate that 24 cm is accurate to the nearest tenth of a centimeter?

3. Average the following measurements of tree heights and appropriately round your answer: 29.5 m, 54.2 m, 73 m, and 68 m.

4. Why do you think it is important to round your final answer to the accuracy of you least accurate measurement?

ACCURACY OF MEASUREMENT

B

TEACHER RESOURCE

Materials

- pencil
- calculator

Management Suggestions

[illegible]

Answer Key

1. Why is it important to measure accurately? Provide an example of a case where accuracy in measurement is critical, for example, can you think of some application in the field of medicine, school grading or clothing manufacturing.

   The more accurate a measurement, the more reliable the results and dependable the end product will be. The more accurate the measurements when constructing a doorframe, the better the door will fit in it.

2. How would you indicate that 24 cm is accurate to the nearest tenth of a centimeter?

   Put a zero as a placeholder to show that the tenths place has a value of zero: 24.0 cm.

3. Average the following measurements of tree heights and appropriately round your

   answer: 29.5 m, 54.2 m, 73 m, and 68 m.

   29.5 m + 54.2 m + 73 m + 68 m = 224.7/4 m ≈ 56 m ...round to the least accurate measure.

4. Why do you think it is important to round your final answer to the accuracy of your least accurate measurement?

   If you do not round to the least accurate measure, you are assuming, or guessing, that the next place value in that least accurate measure is zero.

SCIENTIFIC NOTATION

ACTIVITY

Background

Scientific notation is a way of writing large or small numbers in a shorthand, easier form. A number written in scientific notation consists of a number between one and ten (not including ten) multiplied by some power of 10.

Scientific notation for large numbers

The large number 565,000 is written in standard form.

The same large number can be written in scientific notation as 5.65×10^5.

To write a number in scientific notation, count how many places you must move the decimal point in the standard form in order to form a number that is greater than or equal to one but less than ten. In this case, you would get the number 5.65. Because the decimal point was moved 5 places to the left, the number is written in scientific notation as 5.65×10^5.

Exercises:

1. The average distance between the moon and earth is 240,000 miles.Write this standard number in scientific notation.

2. The sun is approximately 9.3×10^7 miles away from the earth. Write this scientific notation in standard form.

3. Light travels at the speed of about 186,282 miles per second. Find the distance in miles that light travels in one day. Write your answer in both standard form and scientific notation.

Scientific notation for small numbers

You have seen how large standard numbers can be written in scientific notation. Small numbers (numbers less than one) can also be written in scientific notation.

The small number 0.00927 is written in standard form.

The same small number can be written in scientific notation as 9.27×10^{-3}.

To write a number in scientific notation, count how many places you must move the decimal point in the standard form in order to form a number that is greater than or equal to one but less than ten. In this case, you would get the number 9.27.

To write 0.00927 in scientific notation, you need to move the decimal three places to the right. The number in scientific notation is 9.27×10^{-3}. Notice that the exponent is negative.

When the standard number is less than one, the power of 10 is a negative number. When the standard number is greater than one, the power of 10 is a positive number. Study the table and notice the pattern.

Number in standard form	2370	237	23.7	2.37	0.237	0.0237	0.00237
Scientific Notation	2.37×10^{3}	2.37×10^{2}	2.37×10^{1}	2.37×10^{0}	2.37×10^{-1}	2.37×10^{-2}	2.37×10^{-3}

Exercises:

4. Use scientific notation to write the following terms.

   a) 0.00523

   b) 0.00000185

   c) A spectrometer is an instrument that measures visible and non-visible wavelengths of light in nanometers. A nanometer is one billionth of a meter. The red band of visible light is centered at 650 nanometers on the light spectrum. Represent 650 nm in meters.

5. A common influenza virus is approximately 100 nanometers in diameter. Write this number in standard form and in scientific notation in meters.

6. The Visible InfraRed Intelligent Spectrometer (VIRIS) is an instrument that can measure infrared light that is only 0.70 micrometers in length. A micrometer is a millionth of a meter. Write this number in standard form and in scientific notation in meters.

Summary

- Standard number form refers to a number written in a customary way, with each digit in its proper place value (565,000).
- Scientific notation refers to a number written in a non-standard form that allows very large numbers or very small numbers to be written in a more concise way (5.65×10^{5}).
- Scientific notation for large numbers consists of a number that is greater than or equal to one but less than 10 and is multiplied by a positive power of ten.
- Scientific notation for small numbers consists of a number that is greater than or equal to one but less than 10 and is multiplied by a negative power of ten.

SCIENTIFIC NOTATION

B

TEACHER RESOURCE

Materials

- calculator

Answer Key

Large Number Exercises:

1. The average distance between the moon and earth is 240,000 miles. Write this standard number in scientific notation.

   2.4×10^5

2. The sun is approximately 9.3×10^7 miles away from the earth. Write this scientific notation in standard form.

   93,000,000 miles

3. Light travels at the speed of about 186,282 miles per second. Find the distance in miles that light travels in one day. Write your answer in both standard form and scientific notation.

   186,282 x 60 x 60 x 24 =

   186,282 x 86,400 =

   16,094,764,800 miles

   (or)

   1.609×10^{10}

Small numbers Exercises:

4. Use scientific notation to write the following terms.

   a) 0.00523 *5.23×10^{-3}.*

   b) 0.00000185 *1.85×10^{-6}.*

   c) A spectrometer is an instrument that measures visible and non-visible wavelengths of light in nanometers. A nanometer is one billionth of a meter. The red band of visible light is centered at 650 nanometers on the light spectrum. Represent 650 nm in meters.

   standard: 0.00000065 m, scientific notation: 6.50×10^{-7} m

5. A common influenza virus is approximately 100 nanometers in diameter. Write this number in standard form and in scientific notation in meters.

   standard: 0.0000001 m, scientific notation: 1.0×10^{-7} m

6. The Visible InfraRed Intelligent Spectrometer (VIRIS) is an instrument that can measure infrared light that is only 0.70 micrometers in length. A micrometer is a millionth of a meter. Write this number in standard form and in scientific notation in meters.

   standard: 0.0000007 m, scientific notation 7.0×10^{-7} m

AN INTRODUCTION TO STEM-AND-LEAF PLOTS

B

ACTIVITY

Background

Forest Watch is a long-term scientific study being conducted by students, teachers and researchers at the University of New Hampshire. Much of the project involves observing, collecting, analyzing and communicating data about the white pine tree. The results can be used to help both students and scientists look for change in their trees possibly signaling a change in the health of the forest. In order to look for a change in the data, you need to be able to arrange the data and calculate certain descriptive statistics. Being able to interpret and understand what your data is telling you will enable you to better communicate your findings to others.

Procedure

In this activity, you will be using data on needle fascicle lengths that were collected by Forest Watch students at RHAM High School in Hebron, Connecticut. The data they collected appears on the data sheet handout. As you look at the data as it appears in table form, think about the following questions:

What is the average fascicle length of this data set?,

What are the longest and shortest fascicle lengths?,

What is the difference between the longest and shortest needle lengths?,

Are there any fascicle lengths that are a lot different in length from the others?

Can you think of any other questions you might ask? How easy and straightforward is it to answer any of those questions? It's not very easy because of the way the data is presented in the table. If you take the data, however, and rearrange it, you can make it a lot easier to answer not only the questions above, but many others as well.

1. Look again at the fascicle length data from the north side of the trees at RHAM High School. Take the data and rearrange it in a table or plot that will help you more easily answer the following questions: What is the range of the data set (difference between the longest and shortest fascicle length)? Are there any data points that seem much bigger or smaller than most of the others? Use the space below.

   Share your table or plot with other students and see if they can answer the questions easily from your diagram.

One way to rearrange your data is to make a long list of data points starting from the shortest and going to the longest fascicle length. Perhaps you did this. This would make it easy to determine the range of the data set and to easily see if any values are much smaller or much larger than most of the others. If you haven't done so already, place the data from the RHAM data set in order from smallest to largest in the space below.

Now you can find out even more from your data by taking your ordered set and turning it into a diagram called a stem-and-leaf plot. The example that follows shows you how to construct a stem-and-leaf plot. Stem-and-leaf plots are used for relatively small sets of data such as the data set you are using.

Example data set:	2.4	9.7	6.8	9.3	7.1	10.2	8.7	10.5	6.5
	6.7	6.6	7.0	7.7	8.5	8.8	10.0	10.1	9.5
	7.0	7.4	6.7	6.4	10.1	7.5	7.2	7.5	6.5
	8.0	8.7	5.0						

Notice that the smallest data point is 2.4 and the largest is 10.5. Knowing the spread of the data set helps you determine how to start constructing the stem-and-leaf plot.

In this stem-and-leaf plot, the "stems" are made up of the units digits 0 through 10. In Step 1 of building a plot, the stems are written vertically in a column and a short line is drawn to the right of each number.

In Step 2, the data values for tenths are entered to the right of the stems. These values represent the leaves of the plot. So, the "stems" represent units and the "leaves" represent tenths. For example, the data point 9.3 is represented as 9|3. The stem is 9 and the leaf is 3.

```
units|tenths
    0|
    1|
    2|4
    3|
    4|
    5|0
    6|8775564
    7|01525740
    8|50778
    9|573
   10|51012
```

In Step 3, if you are working from an unordered set like the example above, the “leaves” (tenths digits) have to be ordered from least to greatest. So,

units|tenths		units|tenths
0|		0|
1|		1|
2|4		2|4
3|		3|
4|	becomes…	4|
5|0		5|0
6|8775564		6|4556778
7|01525740		7|00124557
8|50778		8|05778
9|573		9|357
10|51012		10|01125

2. You can use the stem-and-leaf plot to make observations about the data. Compare the stem-and-leaf plot representation of the data and your representation of the data. In what ways is your representation better for making observations about the data? In what ways is the stem-and-leaf plot a better representation for making observations about the data?

Look back at the ordered stem-and-leaf plot in the example above. The following questions ask you to look more closely at the data values. The questions also include definitions of words that you will need to know as you learn to make interpretations about situations from the data.

3. The largest and smallest data values are called the <u>extremes</u>. What are the extremes in the data set above? What is the <u>range</u> (difference between the extremes) of data values for this data set?

4. Between which numbers do most of the data values seem to fall (be clustered close to)?

Besides the variation or range of data, it is also important to get a feeling for what statisticians call the center of the data set, or the measure of central tendency. This includes three measures, called the mean, mode and median. In this activity, you will learn about the mean and the mode. The median is discussed in another lesson.

5. Use the data set or stem-and-leaf plot to find the mean of the data values. The mean value is the average of the data values.

6. Use the data set or stem-and-leaf plot to find the mode of the data values. The mode is the most common data value. It is possible to have more than one "common" data value.

Do any of the data points seem to be considerably different from all the others?

A data value that is much greater or less than most of the others in the data set is called an outlier. When outliers are found in the data, questions may arise as to why those data points exist. The outliers could represent chance natural variation, measurement error, a broken needle fascicle, or other causes. If possible, always check your measurements for outliers to see if you can find a reason for the unusual value. As you conduct the Forest Watch protocols and activities, if you find an error in your measurement before you throw away your needles, try to correct your data.

7. The data value of 2.4 cm looks like it could be an outlier. What are some possible ways that this outlier might have been obtained?

8. Use the RHAM High School data from the south side of trees to make a stem-and-leaf plot.

9. Use the stem-and-leaf plot to answer the following questions.
   a. Between which two values do most of the data values fall?
   b. What are the extreme values?
   c. What is the range of the data? Are there any outliers?
   d. What is the mean value? What is the mode?

10. Use the stem-and-leaf plots from both the north and south sides of the trees to compare needle fascicle lengths between the two sides. What are some similarities and differences that you can easily find out from looking at the plots? What do you think may cause the differences you observe?

11. Obtain another data set(s) from your teacher. Create a stem-and-leaf plot(s) using this new information. Find the mean and the mode of the data set. Look for and record any possible outliers and determine the range. If you have more than one set of data and they are complimentary (north vs. south sides of the same type of measure, for example) compare the information you have about both the data sets. Make a statement about how similar or different the data are from each other.

AN INTRODUCTION TO STEM-AND-LEAF PLOTS

B

TEACHER RESOURCE

Introduction

Forest Watch is a long-term scientific study being conducted by students, teachers and researchers at the University of New Hampshire. Much of the project involves observing, collecting, analyzing and communicating data about the white pine tree. The results can be used to help both students and scientists look for change in their trees possibly signaling a change in the health of the forest. In order to look for a change in the data, you need to be able to arrange the data and calculate certain descriptive statistics. Being able to interpret and understand what your data is telling you will enable you to better communicate your findings to others.

This introductory activity will introduce students to the method involved in constructing a stem-and-leaf plot from a relatively small data set. Students will be able to see how organizing their data into a graph like a stem-and-leaf plot can help them better interpret the information. They will compare means and modes as measures of central tendency, try to determine if there are extremes in their data set and, if so, speculate about their importance.

Mathematics Topics

Data Representation
Central Tendencies
Mean, Range, Mode, Extremes, Outliers
Stem-and-leaf Plots
Data Analysis

Curriculum Standards in Mathematics

Standards are listed by number.

1. Mathematics as Problem Solving
2. Mathematics as Communication
3. Mathematics as Reasoning
4. Mathematical Connections
5. Algebra
10. Statistics

Materials

- graph paper
- ruler
- data sheet with data from needle fascicle lengths (found at end of Math Activities)

Management Suggestions

You may want to work through the first part of this activity on the overhead, creating the stem-and-leaf plot as you describe the steps. Students can then complete the activity with very little additional direction. You may want to have students work in cooperative groups of 3 – 5 students so they can readily compare their tables from problem #1.

Students will have the chance to use the definitions in this activity throughout the Forest Watch activities, especially those associated with measurement of needles. This skill, however, will be valuable in other science and math activities beyond the Forest Watch program.

You may want to show students that histograms are stem-and-leaf plots turned sideways. Stem-and-leaf plots are generally used for small data sets. Histograms are used for data sets that are too large to easily list each data value.

Answer Key

1. Look again at the needle fascicle length data from the north sides of trees at RHAM HS. Take the data and rearrange it in a table or plot that will help you more easily answer the questions: What is the range (difference between the longest and shortest fascicle length) of the data set? Are any of the data points significantly different from the others?

   Answers Will Vary. Look for some logical pattern.

2. You can use the stem-and-leaf plot to make observations about the data. Compare the stem-and-leaf plot representation of the data and your representation of the data. In what ways is your representation better for making observations about the data? In what ways is the stem-and-leaf plot a better representation for making observations about the data?

   AWV.

3. The largest and smallest data values are called the <u>extremes</u>. What are the extremes? What is the <u>range</u> (difference between the extremes) of data values for this data set?

   The extremes are smallest: 2.4 cm and largest: 10.5 cm.

   The range is 10.5 -2.4 = 8.1 cm

4. Between what numbers do most of the data values seem to fall (be clustered close to)? Is it the extremes or is it two other numbers in the data set?

   Answers may vary.

   Most of the data seem to fall between 6.4 cm and 10.5 cm. No, it is not the extremes.

5. Use the data set to find the <u>mean</u> of the data values. The <u>mean</u> value is the average of the data values.

   The mean of the data set is 7.8 cm.

6. Use the data set to find the mode of the data values. The <u>mode</u> is the most common data value. It is possible to have more than one "common" data value.

   The data set is actually multimodal: 10.1, 8.7, 7.5, 7.0, 6.7 and 6.5 are all listed twice, so there is no clear mode for this data.

7. The data value 2.4 cm looks like it could be an outlier. What are some possible ways that this outlier might have been obtained?

   The needle may have been broken off or chewed by an insect. The students may have made a mistake in measuring or recording the length of the needle.

8. Use the RHAM High School data from the south side of trees to make a stem-and-leaf plot.

```
 0|
 1|
 2|
 3|
 4|
 5|
 6|38899
 7|23445579
 8|0234456667899
 9|007
10|5
```

9. Use the stem-and-leaf plot to answer the following questions.

   a. Between what two values do most of the data values fall?

   Most of the data falls between 6.3 and 9.0.

   b. What are the <u>extreme</u> values?

   The extreme values are 6.3 and 10.5.

   c. What is the <u>range</u> of the data? Are there any <u>outliers</u>?

   The range of the data is 10.5 - 6.3 = 4.2. There are no outliers.

   d. What is the <u>mean</u> value? What is the <u>mode</u>?

   The mean of this data set is 8.09 cm. There are several data points that appear twice, so again, there is no clear mode for this data.

10. Use the stem-and-leaf plots from both the north and south sides of the trees to compare needle fascicle lengths between the two sides. What are some similarities and differences that you can easily find out from looking at the plots? What do you think may cause the differences you observe?

    The mean needle fascicle length for the north sides of the trees is shorter at 7.8 cm than the mean fascicle length of the needles on the south side at 8.09 cm. Neither data set shows a clear mode. The number of needles being measured in this data set, about 30 each, is relatively small compared to the total number of needles found on a tree. It is possible that the difference between the two sets has occurred by chance sampling variation. On the other hand, a more rigorous statistical test would be needed to tell if the difference between the two sets of data really are different statistically. One step in that direction would be to complete and interpret a box plot of the data covered in the next exercise.

11. Obtain another data set(s) from your teacher. Create a stem-and-leaf plot(s) using this new information. Find the mean and the mode of the data set. Look for and record any possible outliers and determine the range. If you have more than one set of data and they are complimentary (north vs. south sides of the same type of measure, for example) compare the information you have about both the data sets. Make a statement about how similar or different the data are from each other.

    AWV. Have students work with several data sets from different schools. Ask them to compare their results. Are the needle lengths different between schools? Why might they be different?

USING A STEM-AND-LEAF PLOT TO LEARN MORE ABOUT YOUR DATA

ACTIVITY

How can you determine the distribution of a data set?

Procedure

Several students in Forest Watch measured the heights of all the white pine trees in their study plot. They found the trees were 1.6 m, 12.2 m, 14.7 m, 16.8 m, 13.1 m, 3.1 m, and 1.5 m in height. They calculated the mean (or arithmetic average) of the tree heights and found it was 9.0 m.

The students then debated about whether the average height of the trees, 9.0 m, was the best description of the center of the tree height data. In other words, did 9.0 m provide a good, close description of the height of any given tree that was measured? One of the students, Joe said, "I think that the average height, 9.0 m, best represents the center height of the trees in our study plot." Do you agree or disagree with Joe's statement? Why? In this activity, you will explore other ways, besides the mean, to represent the center of data sets.

First, study the shapes of the example stem-and-leaf plots below. The shape of a data set is also called its distribution.

Mound-Shaped ("Normal")

3|4
4|699
5|244559
6|177
7|8

U-Shaped

2|257779999
3|2355678
4|588
5|2
6|11377
7|34459
8|235566789

Rectangular-Shaped

3|2578
4|137
5|0449
6|1456
7|3688

J-Shaped

1|3
2|2
3|12
4|78
5|115
6|222345578
7|12226677999

J-Shaped

1|33344467889
2|222345578
3|123456
4|78
5|15
6|2
7|12

A mound-shaped plot is often found when you graph the distribution of large data sets. It is sometimes called a bell-shaped or normal distribution. In a mound-shaped distribution, the data values are fairly symmetrical, with low data values balancing high data values. Most of the values are clustered in the middle of the graph. When the data are distributed in a mound shape, the mean (arithmetic average) provides a good representation of the middle or center of the distribution.

Sometimes, however, the data does not fall into a mound shaped distribution. Its shape may be more like one of the others in the examples above. When this occurs, how can we describe the "center" of the data set? Follow the example below.

Students at RHAM High School collected data values for needle fascicle lengths, measured in cm, found on the north side of study trees. A stem-and-leaf plot of the data appears below.

```
units|tenths
     0|
     1|
     2|4
     3|
     4|
     5|0
     6|4556778
     7|00124557
     8|05778
     9|357
    10|01125
```

1. Is the data distribution mound shaped? If not, how would you describe the shape of the distribution?

   When data is not mound shaped, a measure of central tendency called the median is often used to represent the center value.

2. The median is the middle data point in an ordered set (that is a set where the numbers are arranged from smallest to largest). In this case, because there is an even number of data points, the median lies between the 15th and 16th data value.

   a. Find the median value. It is the number found half way between the 15th and 16th data values.

   We can also look at this data more closely to see where most of the data points are found in relation to the median. The lower quartile (LQ), median, and upper quartile (UQ) values divide the data into four parts with approximately the same number of observations in each part. For this data set, locate the upper quartile value by counting up 8 digits from the median and locate the lower quartile value by counting down 8 digits from the median.

   b. What are the upper quartile and the lower quartile values?

   c. What percent of the data lies between the upper and lower quartile values?

   d. You can now calculate the interquartile range (IQR) by subtracting the lower quartile from the upper quartile. Calculate the interquartile range.

Some data sets have one or two numbers that are found either well above or well below not only the median of the data set, but the middle 50% of the data set as well. These numbers are called outliers. You can check to see if a data set has any outliers by using the equation below.

UQ = upper quartile

IQR = interquartile range

LQ = lower quartile

Upper outliers:

Data that are more than the UQ + 1.5(IQR) = 9.3 + 1.5(2.6) = 13.2 cm are upper outliers.

Lower outliers:

Data that are less than the LQ - 1.5(IQR) = 6.7 - 1.5(2.6) = 2.8 cm are lower outliers.

e. Check and use the results of these formulas to find any upper or lower outliers in the RHAM High School data.

3. Use the data values for needle fascicle lengths found on the south side of the study trees at RHAM High School to make a stem-and-leaf plot.

4. Use your stem-and-leaf plot to find the following.
   a. median
   b. upper and lower quartiles
   c. interquartile range
   d. upper and lower outliers
   e. Describe the shape of the data set.

5. How would you compare the needle fascicle lengths between the north and south sides of trees in the RHAM data set? Use the data distribution information in the stem-and-leaf plots and the calculations of median, quartiles, range and outliers to help you in your comparison. Write a short paragraph of your findings.

6. Recall that at the beginning of this activity you read that several students measured the heights of seven white pine trees in their study site. They found the trees were 1.6 m, 12.2 m, 14.7 m, 16.8 m, 13.1 m, 3.1 m, and 1.5m in height. The mean (or arithmetic average) of the tree heights was calculated to be 9.0 m. One student, Joe, thought that the best representation of the middle height of the trees was the average, 9.0 m.

   a. Find the median of the heights of the seven trees in this data set.

   b. Do you think the mean or the median best represents the middle height and the center of the data. Explain your answer.

7. Return to the five examples of stem-and-leaf plots at the beginning of this activity. Find and put a circle around the median data value in each of the examples. Remember that the median is the data point found in the exact middle of the ordered data set.

   Now, turn the page with the various stem-and-leaf plots sideways so that the stems of the data sets are toward you and the leaves are pointing away from you. Observe again the shapes or distributions of the data sets.

   A distribution is called <u>symmetric</u> if the upper and lower halves of the stem-and-leaf plot are almost mirror images of each other.

   A distribution is referred to as <u>skewed</u> if the mound of the data is concentrated either to the right or to the left of the center of the plot. It is <u>skewed to the right</u> if there is a <u>concentration of relatively small values</u>, with some scatter over a range of larger values. It is <u>skewed to the left</u> if there is a <u>concentration of relatively large values</u>, with some scatter over a range of smaller values. The scattered data on either end of these kinds of plots are sometimes referred to as "tails". The tails act to "pull" the data either to the right or to the left of the center of what would be a symmetrical distribution.

8. For each of the shapes in the examples found at the beginning of this activity, find the median and state whether the data set is symmetrical, skewed left or skewed to the right or neither.

   a. The mound-shaped data set.

   b. The U-shaped data set.

   c. The J-shaped data sets.

   d. The Rectangular-shaped data set.

USING A STEM-AND-LEAF PLOT TO LEARN MORE ABOUT YOUR DATA

TEACHER RESOURCE

Introduction

This activity serves as a follow-up and extension to the introduction to stem-and-leaf plots activity. In that activity, students learned how to take a data set and present it in the form of a stem-and-leaf plot. The concepts of mean and mode were reviewed. Students used the stem-and-leaf plots to identify the extremes, possible outliers and the range of the data set.

In this activity, students will look more closely at the distribution of the data set as pictured in the stem-and-leaf plot. They will be asked to think about which measure of central tendency, the median or the mean, is the most appropriate to explain the center of a particular data set. The mode is not discussed in this activity, but you could review it if you wish. Students will distinguish between several distribution shapes and will learn the difference between symmetric and skewed data distributions. They will calculate the median, 25% and 75% quartiles, the interquartile range and outliers of a data set.

Curriculum Standards in Mathematics

Standards are listed by number

1. Mathematics as Problem Solving
2. Mathematics as Communication
3. Mathematics as Reasoning
4. Mathematical Connections
5. Algebra

10. Statistics

Materials

- graph paper
- ruler
- data sheet with data from needle fascicle lengths (found at end of Math Activities)

Management Suggestions

This activity is a prerequisite for several other mathematics activities. Students can do this activity with very little direction. You may want to have students work in cooperative groups of 3 – 5 students so they can readily compare their work. At the end of the activity, ask them to make up some data sets that show various types of data distribution patterns. Teachers have found that students enjoy writing their own mathematics problems and having other students solve them. Students also deepen their understanding of mathematical concepts when they write problems for others to solve.

If you feel your students know the terms, or can learn them easily with a few minutes discussion, you may want to skip this activity and go directly to the Box Plots Activity. For younger students, you may want to skip the formal definition of outliers. You could define outliers as data values that are very different from most others in the data set.

Regardless of what teaching path you choose, you should stress why outliers may be of importance. While there are always chance occurrences and possible anomalies in any type of data, many times unusual data points signal an error in data gathering, copying, etc. The importance of errors in measurement of samples should be discussed and corrected if possible. This should help reinforce the importance of accuracy in every phase of data collection and manipulation.

You should continue to stress the difference between the mean (arithmetic average) and median. Students need many experiences finding the mean and median of data sets and thinking about which value best represents the "center". The mean represents the center best when the data set is very large or normally distributed. Many data sets that Forest Watch students use are small. In these situations, the median may better represent the center.

Be sure students learn the meaning of the words in this activity. These terms are essential for students' interpretation and description of data sets in the Forest Watch activities.

For younger students, you may want to develop games to help students learn the vocabulary words. One example of using a game to learn definitions is "Go Fish." Students write each word on a 3 X 5 card and the definition on another. Students then shuffle the "deck" of words and definitions. Each student draws seven cards. The first student reads aloud the word or definition on one of their cards. The student with the corresponding definition or word gives the first student the card, reading the definition or word aloud. Other students verify the match. The first student then places the "pair" on the table in front of her. If no student has the matching card, the first student must "go fishing" in the stack. Play continues as in regular "Go Fish." The student with the largest number of pairs wins.

Some teachers regularly put quiz and test scores on the board in stem-and-leaf plots. They then ask their students to find the mean, median, and upper and lower quartiles and note the range in which 50% of the scores fall. By asking the students to find the median and mean, teachers can encourage students to reflect upon and discuss whether the mean or median is a fairer "average" score.

B

Answer Key

1. Is the data distribution mound shaped? If not, how would you describe the shape of the distribution?

   The data is not mound shaped. The data is somewhat rectangular but does not have any particular shape. There is an outlier.

2 a. Find the median value. It is the number found half way between the 15th and 16th data values.

   The median for this data set is 7.5 cm. It just so happens that both the 15th and the 16th value are 7.5 cm. In other cases, you would need to calculate the average between the two values.

   b. What are the upper quartile and the lower quartile values?

   The upper quartile is 9.3 cm and the lower quartile is 6.7 cm.

   c. What percent of the data lies between the upper and lower quartile values?

   50% of the data lies between the upper and lower quartiles.

   d. You can now calculate the <u>interquartile range</u> by subtracting the lower quartile from the upper quartile. Calculate the interquartile range.

   The interquartile range is 9.3 - 6.7 = 2.6 cm. A small range indicates that most of the data is similar to or clustered near the median. A broad range indicates a broad spread of the data.

   e. Check and use the formulas to find any upper or lower outliers in the RHAM HS data .

   There are no data values greater than 13.5 cm, so there are no upper outliers. 2.4 is the lower outlier.

3. Use the data values for needle fascicle lengths found on the south side of the study trees at RHAM High School to make a stem-and-leaf plot.

   0|
   1|
   2|
   3|
   4|
   5|
   6|38899
   7|23445579
   8|0234456667899
   9|007
   10|5

4. Use your stem-and-leaf plot to find the following.

   a. median *8.25 cm*

   b. upper and lower quartiles *8.7 cm, 7.4 cm*

   c. interquartile range *8.7 -7.4 = 1.3 cm*

   d. upper and lower outliers *none*

   e. Describe the shape of the data set. *The data is mound-shaped.*

5. How would you compare the needle fascicle lengths between the north and south sides of trees in the RHAM data set? Use the data distribution information in the stem-and-leaf plots and the calculations of median, quartiles, range and outliers to help you in your comparison. Write a short paragraph of you findings.

   The needles on the south side of the tree demonstrate a more normal (mound-shaped) distribution than the needles on the north side of the tree. The median in the north is 7.5 cm while the median in the south is almost a centimeter longer at 8.25 cm. The interquartile range of the north at 2.6 cm is about two times as wide as the interquartile range in the south at 1.3 cm. The data from the north has one outlier, a low data value, while the data from the south show no unusual data points in relation to the others. Overall, the data from the north side of the tree seems to have more variation than the data from the south. It shows no definite pattern of needle fascicle length.

6 a. Find the median of the heights of the seven trees in this data set.

   The median is 12.2 m.

   b. Do you think the mean or the median best represents the middle height and the center of the data. Explain your answer.

   The median may be a better representation of the data set. The two very low data points pull the mean to a lower value than the median, thus masking the fact that there are four trees that are fairly tall in comparison to the short ones in the data set.

8. For each of the shapes in the examples found at the beginning of this activity, find the median and state whether the data set is symmetrical, skewed left or skewed to the right or neither.

   a.The mound-shaped data set. *median 5.45 cm, symmetrical (mound)*

   b.The U-shaped data set. *median 5.2 cm, symmetrical (U- shaped)*

   c.The J-shaped data sets. *a. median 6.5 cm, skewed left,* *b. median 2.4 cm, skewed right*

   d.The Rectangular-shaped data set. *median 5.4 cm, symmetric*

USING BOX PLOTS TO COMPARE DATA SETS

B

ACTIVITY

How can you compare two sets of needle data to see if they are similar to each other or different?

Background

This activity will help you draw diagrams that will let you make comparisons between two or more similar data sets. Why would you want to compare two or more sets of data? Suppose that you and your classmates have been collecting needle length data for five white pine trees on your school grounds for several years. One reason for collecting the data is to find out if the needle length is changing over time or if the average needle length remains about the same over the years. You can check for change in your needle length by constructing, comparing and analyzing diagrams of the data called box plots.

Procedure

1. A group of students involved in Forest Watch have speculated that the needle fascicles on different sides of a tree may be different from one another in length. They asked the question: Is the average length of needle fascicles longer on the north or on the south side of a tree, or are they about the same? What do you think? Give reasons for your answer.

2. Suppose you wanted to use data to find out if there is a difference in the average fascicle length between the north and south sides of a tree. How could you use the data that you and your classmates have gathered over the years to compare the two sides? Brainstorm some ideas. How might you use a picture or diagram to compare the two data sets (north and south)? Discuss with others some ideas they have.

The example below shows a sample data set of fascicle lengths from the north and south sides of a tree. The fascicles were measured to the nearest tenth of a centimeter. The data are presented as two stem-and-leaf plots, one for each set of data. Needle fascicle length (cm) from the north side of a tree.

Needle fascicle length (cm) from the south side of a tree.

units|tenths	units|tenths
0|	0|
1|	1|
2|4	2|
3|	3|
4|	4|3667789
5|9	5|1111233788999
6|4	6|134447
7|1122345555666799	7|0023
8|112222223	8|
9|01	9|
10|	10|

3. What do the stem-and-leaf plots suggest about the needle length of north and south side of that tree?

The information from the stem-and-leaf plots can be further used to construct box plots of the data. Box plots will give you another way to look at and compare the two data sets.

A box plot is defined by five numbers: a lower "whisker", the first quartile, the median, the third quartile, and an upper "whisker" value for a data set. The first and third quartiles define the extremes of the box. The median is indicated within the box. The values at about the 10th and 90th percentiles form the "whiskers" of the diagram.

4. Use the Sample High School needle data from the north and south sides of study trees to compute the following: (Use data from the stem-and-leaf plots above.)

	north side needles (cm)	south side needles (cm)
a. median		
b. lower (first) quartile		
c. upper (third) quartile		
d. interquartile range		
e. lower extreme (lowest value)		
f. upper extreme (highest value)		
g. lower outliers		
h. upper outliers		

The information in the table can now be used to construct box plots representing the data from the north and south sides of trees.

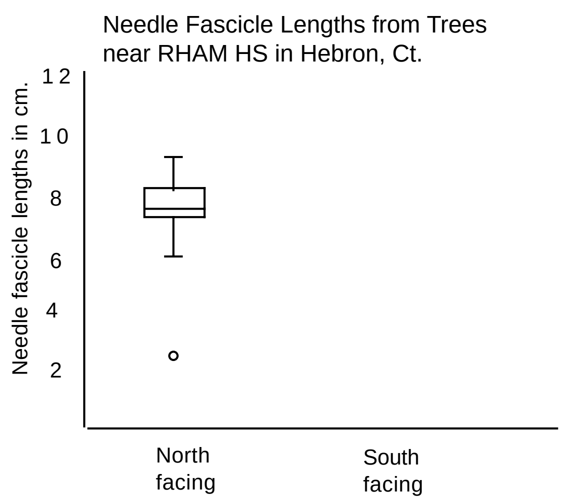
a. The first step in the process involves creating a grid onto which the boxes will be drawn. Observe the diagram below. In this case, the grid consists of a vertical line and a horizontal line that meet like the outlines of a graph. The vertical line is labeled “Needle Fascicle Lengths in Centimeters” and is marked with equally spaced values 0 through 12. The horizontal line is divided in half and labeled “ North Facing” and “South Facing”, one on each half.

b. The title of the diagram is <u>Needle Fascicle Lengths from Trees near Sample High School in Littletown, CT</u>.

c. Look at the box over the North Facing column on the diagram. Notice that the box has a horizontal line running through it on the inside. This line represents the median data point for the sample data set. The median for this data set is equal to 7.6 cm. The horizontal line that forms the top of the box is the upper quartile (8.2 cm) while the line forming the bottom is the lower quartile (7.3 cm) for the data set. Connected together, these lines form the “box”.

d. The vertical lines extending above and below the box that have smaller horizontal lines at the top or bottom are called the “whiskers”. The lower “whisker” represents the first data value, 5.9 cm, that is greater than the length computed for the lower outlier. The upper “whisker” represents the first data value, 9.1 cm, that is less than the length computed for the upper outlier. (These whisker values are <u>not</u> listed in your table above, you must find them in the original stem-and-leaf plot.)

e. The outliers, if there are any, are marked with small “o”s.



5. On the box plot diagram above, use the data that you calculated to draw a box plot of the data from the needle fascicles on the south side of trees.

Analysis and Interpretation

6. Use the box plot diagram to make a statement comparing the data sets from the north and south sides of trees. Can you determine from the box plot graphs whether or not the needle fascicle lengths are different from one another or not? Explain using all the various values that you calculated and plotted on the graph. For example, how do the medians and the lower and upper quartiles compare between the two sample sets? Are they similar to one another (is there overlap) or are they very different?

7. What information do box plots convey more readily than stem-and-leaf plots?

8. Given the following situations, decide whether you would use a stem-and-leaf plot, box plot, histogram, or bar graph to analyze a particular feature of the data. Explain your answers.

    a. You know the lengths of 20 needles and you want to know the range in which most of the lengths fall.

    b. You have data about the needle lengths from 50 schools in New England. You are interested in knowing the shape of the distribution of the needle lengths before you do further analysis.

    c. You have data about the needle lengths from 50 schools in New England. You are interested in comparing the lengths of the needles from the north and south sides of all the trees. You are particularly interested in the middle 50% of the needle lengths.

    d. You have data about the needle lengths from 50 schools in New England. You are interested in comparing the needle lengths in each state.

9. Use the data collected by students at your school (or from other schools) to make box plots of the data from the north and south sides of trees. You may use data from one year combining all from the north side of five trees and all from the south side of five trees. Or, if you have enough data, you may compare needle length from the same tree over two or more years. Think of other comparisons you could make.

10. Why would you be concerned if some of your data shows that needle length varied significantly from one year to the next? Speculate as to what might cause needle lengths to vary.

Communicating your Knowledge

Stem-and-leaf and box plots are statistical tools that can help us to more easily visualize the data found in a table. What are some of the advantages and disadvantages to using one or the other type of plot? How can these kinds of data representations help you to better describe your data than just using a data table and how can they help you tell if the data are similar or different from one another?

USING BOX PLOTS TO COMPARE DATA SETS

TEACHER RESOURCE

Introduction

In this activity, students will get a chance to explore ways to not only describe their data set using statistical measures, but to visually represent their information and readily make comparisons between two or more sets of data.

One way to use the data in a visual form is to build and analyze a diagram called a box plot. One type of box plot is defined by five numbers: a minimum (lower whisker), the first quartile, the median, the third quartile (together forming the box), and a maximum (upper whisker) value for a data set. In this exercise, students will incorporate more information into their plot as they will define outliers beyond the whiskers. Just as in a traditional box plot, the first and third quartiles define the extremes of the box and the median is indicated within the box. The whiskers, however, are the data values that are found at about the 10th and 90th percentiles. And the extremes, if they are calculated as outliers, are added beyond or below the whiskers.

After completing this activity, students should be able to use the descriptive statistics of a data set (median, quartiles, etc.) to construct and analyze a box plot. They should also be able to discuss the strengths and weaknesses of using the various ways to present their data and which way is most appropriate in a given circumstance: stem-and-leaf, bar graph, histogram and/or box plot.

Curriculum Standards in Mathematics

Standards are listed by number

1. Mathematics as Problem Solving
2. Mathematics as Communication
3. Mathematics as Reasoning
4. Mathematical Connections
5. Algebra

10. Statistics

Materials

- graph paper • ruler
- data values for needle fascicles on the north and south side of trees

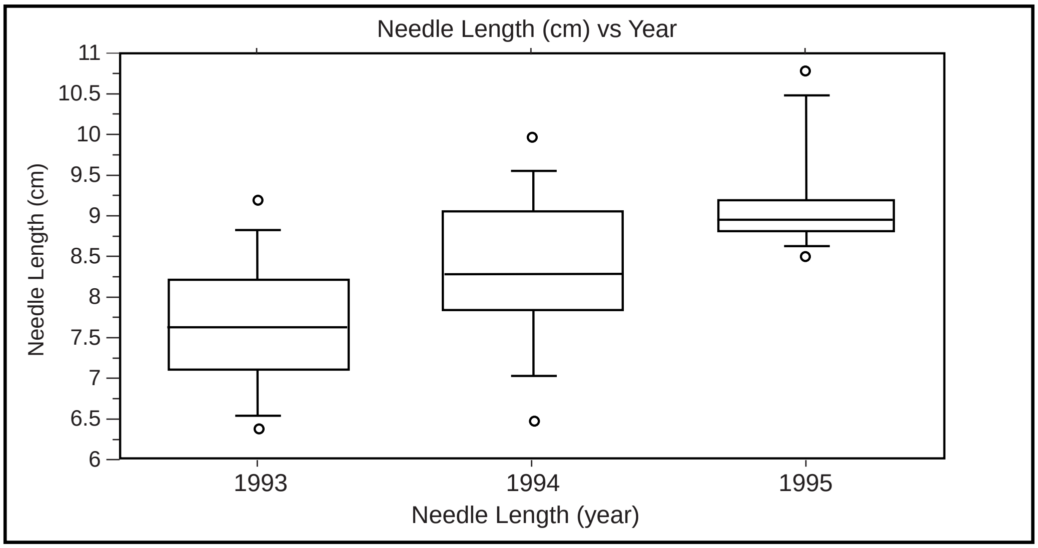
Management Suggestions

Before trying this activity, you should cover or review with your students the topics found in the two stem-and-leaf plot activities. Students must be able to find quartiles and use the equations presented in the stem-and-leaf activities to find outliers.

This activity helps students learn a powerful and widely used technique for comparing large and small data sets. Older students should be able to do this activity with very little direction. You may, however, want to provide more guidance for younger students.

Try arranging cooperative student work groups of three to five students each. This will allow advanced students to help others in their group and will provide small discussion groups for brainstorming while answering some of the questions.

Many statistical computer programs will draw box plots for you after you enter the data. Box plots have become widely used in the second half of the twentieth century. An example of a computer drawn box plot follows. The example box plot below was drawn using the statistical program: StatView.



Answer Key

1. A group of students involved in Forest Watch have speculated that the needle fascicles on different sides of a tree may be different in average length from one another. They asked the question: Is the average length of needle fascicles longer on the north or on the south side of a tree? What do you think? Give reasons for your answer.

   Some scientists think needle fascicles from the north facing sides of trees may grow longer because they receive less light. The reduced light causes the needles on the north side to elongate to produce more photosynthetic material.

2. Suppose you wanted to use data to find out if there is a difference in the average fascicle length between the north and south sides of a tree. How could you use the data that you and your class-mates have gathered over the years to compare the two sides? Brainstorm some ideas. How might you use a picture or diagram to compare the two data sets (north and south)? Discuss with others some ideas they have.

   Answers Will Vary

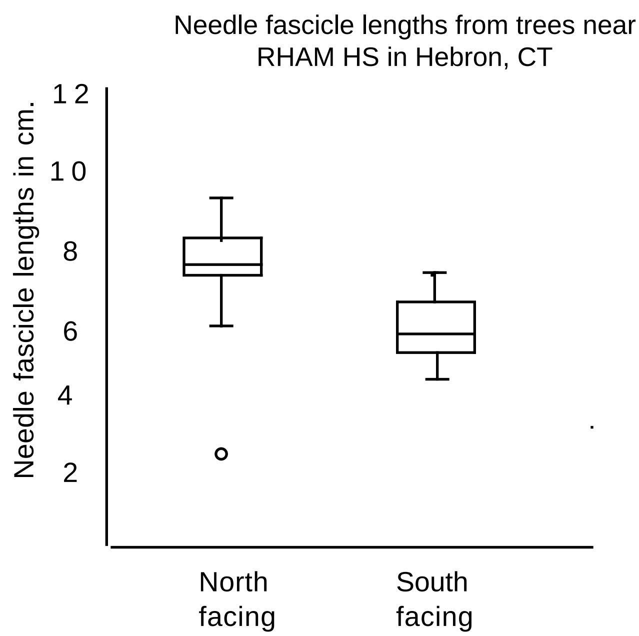
3. What do the stem-and-leaf plots suggest about the needle length of north and south side of that tree?

   The needles from the north side of the trees tend to be longer. Encourage students to find the median and the mode of the two data sets as well as the range.

4.

	north side needles (cm)	south side needles (cm)
a. median	7.6	5.75
b. lower (first) quartile	7.3	5.1
c. upper (third) quartile	8.2	6.4
d. interquartile range	0.9	0.7
e. lower extreme(lowest value)	2.4	4.3
f. upper extreme(highest value)	9.1	7.3
g. lower outliers	2.4	none
h. upper outliers	none	none

5.



6. Use the box plot diagram to make a statement comparing the data sets from the north and south sides of the tree. Can you determine from the box plot graphs whether or not the needle fascicle lengths are different from one another or not? Explain using all the various values that you calculated and plotted on the graph. For example, how do the medians and the lower and upper quartiles compare between the two sample sets? Are they similar to one another (is there overlap) or are they very different?

The box plots give a visual indication that the data sets are in fact different from one another and that the needles from the north side are generally longer than the needles from the south. Notice that the boxes themselves do not overlap one another. There is a little bit of overlap in the data *between the upper whisker of the south side data and the lower whisker of the north side data. The data from the north shows a slightly greater range or variation of data values than the south. This is illustrated by the slightly longer whiskers on the north box plot and the low outlier.*

7. What information do box plots convey more readily than stem-and-leaf plots?

   Box plots make it easier to see where 50% of each data set lies. They show the median and upper and lower quartiles of each data set clearly. This makes it easier to make comparisons. They also show any outliers.

8. Given the following situations, decide whether you would use a stem-and-leaf plot, box plot, histogram, or bar graph to analyze a particular feature of the data. Explain your answers.

   a. You know the lengths of 20 needles and you want to know the range in which most of the lengths fall.

   stem-and-leaf

   b. You have data about the needle lengths from 50 schools in New England. You are interested in knowing the shape of the distribution of the needle lengths before you do further analysis.

   histogram

   c. You have data about the needle lengths from 50 schools in New England. You are interested in comparing the lengths of the needles from the north and south sides of all the trees. You are particularly interested in the middle 50% of the needle lengths.

   box plot

   d. You have data about the needle lengths from 50 schools in New England. You are interested in comparing the needle lengths in each state.

   bar graph

9. Use the data collected by students at your school (or from other schools) to make box plots of the data from the north and south sides of trees. You may use data from one year combining all from the north side of five trees and all from the south side of five trees. Or, if you have enough data, you may compare needle length from the same tree over two or more years. Think of other comparisons you could make.

   AWV

10. Why would you be concerned if some of your data shows that needle length varied significantly from one year to the next? Speculate as to what might cause needle lengths to vary.

    Answers should include natural causes such as disease of the tree due to insects or fungi, abiotic factors such as poor air quality, drought or lack of nutrients as well as errors in measurement or biased sampling of needles.

Communicating your Knowledge

Stem-and-leaf and box plots are statistical tools that can help us to more easily visualize the data found in a table. What are some of the advantages and disadvantages to using one or the other type of plot? How can these kinds of data representations help you to better describe your data than just using a data table and how can they help you tell if the data are similar or different from one another?

Data tables consisting of unordered data values are not very useful in helping one describe and define a sample population. It is much more useful to determine some statistical descriptors of the sample such as the mean, median, mode and variation of the data and then to arrange that data in a visual manner such as a stem-and-leaf or box plot.

Stem-and-leaf plots are practical to use as long as there is not too much data to plot (no more than about 30 values). They can quickly allow one to assess the median, mode, shape or distribution as well as the range of the data set. They can be used to more easily determine the upper and lower quartile data values.

Box plots use the descriptive statistics of a data set to more easily allow for the comparison between two or more sets of information. This provides a quick way to determine whether or not the data sets are different from one another. This information can then be used to advance the question-why?!

ACTIVITY

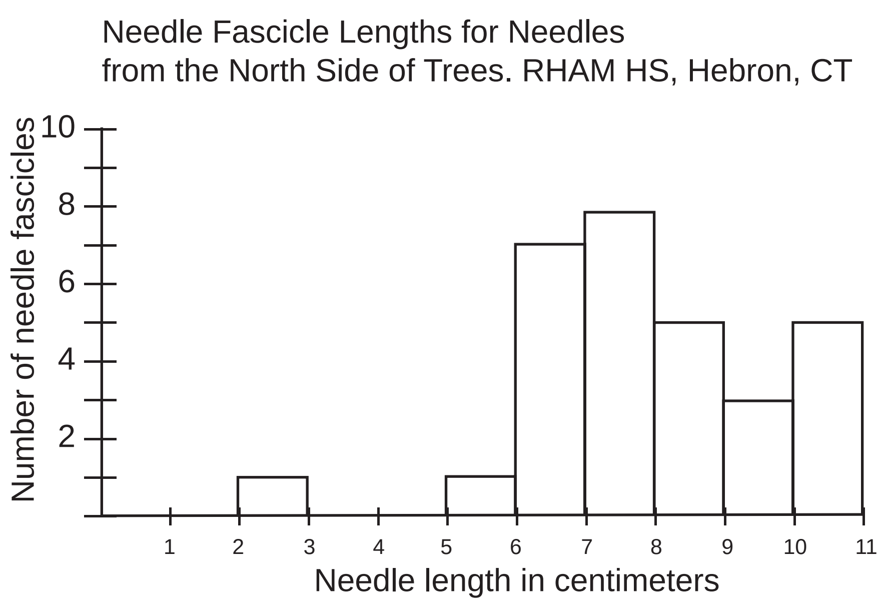
Background

Students in Forest Watch will be collecting data about their white pine trees for many years. Sometimes the data can be analyzed using stem-and-leaf plots, but as the amount of data grows or the way in which you want to look at your data changes, you might want to use other forms of graphs such as histograms or bar graphs to display your data. In this activity, you will learn how to make a histogram, sometimes called a frequency graph, and a bar graph.

Procedure

Part A- Histograms

Data sets often include a great many values. How can you represent the many values in such a way as to be able to more easily interpret them? One way to do this is to create a graph called a histogram. A histogram allows you to group together values that are close to one another and plot them as bars on a graph. The histogram below shows the needle fascicle length data for the north side of study trees at RHAM High School in Hebron, CT. Look at a copy of the needle fascicle length data and try to see how they were grouped as you interpret the graph below.



The bar between 6.0 and 7.0 includes data values that are equal to or greater than 6.0 cm and less than 7.0 cm. The needle measurements in cm are listed on the horizontal axis of the histogram and the actual number (or frequency) of values that fall into each measurement category are listed on the vertical axis.

Histograms Summary

The horizontal axis of a histogram is a measurement scale in which the horizontal axis is broken into intervals of equal size. To construct a histogram, break the range of values of a variable into convenient intervals and write appropriate values on the horizontal axis. The vertical axis lists the count of the data values that fall in each interval.

Histograms resemble stem-and-leaf plots turned on their side. Use stem-and-leaf plots for small data sets and histograms for larger data sets. It is generally not appropriate to place spaces between the bars because the horizontal axis has a measurement scale. Spaces are used, however, if the count for a particular interval is zero.

Histograms are sometimes called frequency graphs. Histograms can also be used to display relative frequencies. For relative frequencies, the vertical axis is marked in percentages. The percent of observations that fall into each interval is graphed. The percents must add to 100%.

Exercises:

1. Describe what the difference is between a number that is greater than or equal to twelve but less than thirteen and a number that is greater than twelve but less than thirteen? How would you interpret the temperature situation when someone tells you that the temperature today is going to be in the twenties? Can you think of other examples of the use of greater than or equal to?

   Use the histogram of needles fasciles length from the north sides of trees from RHAM, HS to answer the following questions.

2. How many needle fascicles in the histogram above were between 7.0 cm and 8.0 cm? (This means equal to or greater than 7.0 cm and less than 8.0 cm.)

3. What kind of information can you get about needle fascicle lengths from this histogram? Is there a clear mode of needle lengths on this graph? What is the mode?

4. What differences are there between the information you can obtain from a stem-and-leaf plot of the needle length data and from a histogram?

5. When do you think histograms would be more useful than stem-and-leaf plots?

6. Draw a histogram for the data from the south side of trees at RHAM High School. Describe the shape of the graph.

7. Compare the histogram of the north side needle lengths with the histogram of the south side needle lengths. How are they similar and different?

8. What information does a histogram allow you to read more or less easily than a box plot?

Part B– Bar Graphs

A bar graph is another tool that you can use to visually present groups of data. Bar graphs are used to compare data values taken at a given point in time. A line graph, on the other hand, is used to show trends in a measure over time.

Table 1 contains the 1996 average needle length information from four schools in the Forest Watch study. Use the data in the table to create a bar graph that will allow you to easily compare the average needle length between the schools. Be sure to title your graph. The vertical axis (y-axis) should represent Needle Length (cm) and should begin with the increment of 7.0 and increase by increments of 0.5. School name and location should be recorded on the horizontal axis (x-axis). You may wish to use the initials of towns and states to label this axis.

Table 1. 1996 Needle length data from four Forest Watch schools.

School	Average Needle Length (cm)
Kennebunk School Kennebunk, ME	7.5
Kenneth A. Brett School Tamworth, NH	7.3
Hanson School West Buxton, ME	7.6
Hollis School Hollis, NH	8.9

1. When you have completed your graph, write a short paragraph that describes the differences in the average needle length among the four locations.

2. What would happen to the bars if the increments on the vertical axis were increased from 0.5 to 2.0 (if you kept the size of the increments the same height)?

3. What would happen to the bars if the increments on the vertical axis represented 0.1 cm steps?

4. Summarize how a change of scale on your vertical axis could affect your interpretation of the graph?

5. Find the location of each of the schools on a map. Do you think location could be an influence on the length of needles? Explain why or why not.

A clustered bar graph allows you to compare two data sets in a similar manner as you did in the bar graph in the exercise above. Examine the following Data Table 2.

Table 2. 1996 Needle length data from north and south side of trees from four Forest Watch schools

School	Ave. Needle Length (cm) North	Ave. Needle Length (cm) South
Kennebunk School Kennebunk, ME	7.8	7.3
Kenneth A. Brett School Tamworth, NH	7.7	7.0
Hanson School West Buxton, ME	7.9	7.2
Hollis School Hollis, NH	9.0	8.8

6. Use the data in Table 2 to create a clustered bar graph. Use the appropriate choice of scale for your vertical axis. This clustered bar graph should be set up the same way as a bar graph, however, you should make sure that you place the North and South needle length bars for the same school next to each other on the graph.

7. Write a short paragraph that describes the difference in average needle length that you can observe from your graphs.

8. Compare North and South side needle averages. Is there a noticeable difference in needle average from one quadrant to another?

9. If so, brainstorm some factors that might account for this.

10. List three questions that you could ask as you observe this graph.

11. How would you research and answer the question: “Is the pattern of average needle length in the north and south sides of study trees indicated by this clustered bar graph the same pattern that is observed in other Forest Watch study trees?” Would you expect to see this same pattern for other study sites in the Forest Watch Project? Why or why not?

Bar Graphs Summary

Bar graphs are used to compare the amounts of different types of items or categories at one point in time as opposed to comparing the frequencies of one type of item over a continuum. The horizontal axis of a bar graph need not have any measurement scale but may simply identify the items being compared. Bar graphs usually have spaces between the bars. The vertical axis records the counts. Clustered bar graphs allow you to compare more than one type of data for each category.

12. For each of the following situations, choose which representation: a stem-and-leaf plot, histogram, bar graph, or box plot would provide the most information in the easiest manner. Provide reasons for your answers.
    a. You have a small set of data (n = 20) and you are interested in knowing the median and upper and lower quartiles.
    b. You have a large set of tree diameter data (n = 500) and you would like to know if the data is mound-shaped and whether the mean or the mode is an appropriate representation of the center of the data.
    c. You have data from a forest that includes the number of trees that show several different injuries: recent evidence of deer chewing on the bark, insects nibbling on the needles and branches broken in a storm. You would like to compare the number of trees with each type of damage.
    d. You have data representing the heights of 100 trees in a particular area. You would like to know the range of heights in which 50% of the trees fall.

13. Make histograms of all of the needle fascicle length data your class has collected from your study trees. Compare the lengths of needles from the north and south sides of study trees at your site.

HISTOGRAMS AND BAR GRAPHS

TEACHER RESOURCE

Introduction

This activity introduces one more way to explore data analysis. Histograms and bar graphs may allow students to work more efficiently and quickly with larger data sets than stem-and-leaf plots. After completing this activity, students should be comfortable drawing both histograms and bar graphs and know the strengths and weaknesses of each. They will end the exercise by matching example data sets with the appropriate visual representation: stem-and-leaf, histogram, bar graph or box plot.

Curriculum Standards in Mathematics

Standards are listed by number

1. Mathematics as Problem Solving
2. Mathematics as Communication
3. Mathematics as Reasoning
4. Mathematical Connections
5. Algebra
10. Statistics

Materials

- graph paper
- ruler

Management Suggestions

Use this activity after you have covered the stem-and-leaf plot lesson. There is also a reference to box plots, so it is necessary to at least show an example of a box plot and explain how they are used and interpreted.

Histograms are very similar to stem-and-leaf plots. However, histograms do not show each individual data value as you would find in a stem-and-leaf. A stem-and-leaf plot turned on its side resembles a histogram and both can be used to interpret the distribution of data: normal, skewed, bimodal, etc.

The following pages discuss the differences between bar graphs and histograms. The pages can be used as an overhead transparency or a written handout for students, if you wish.

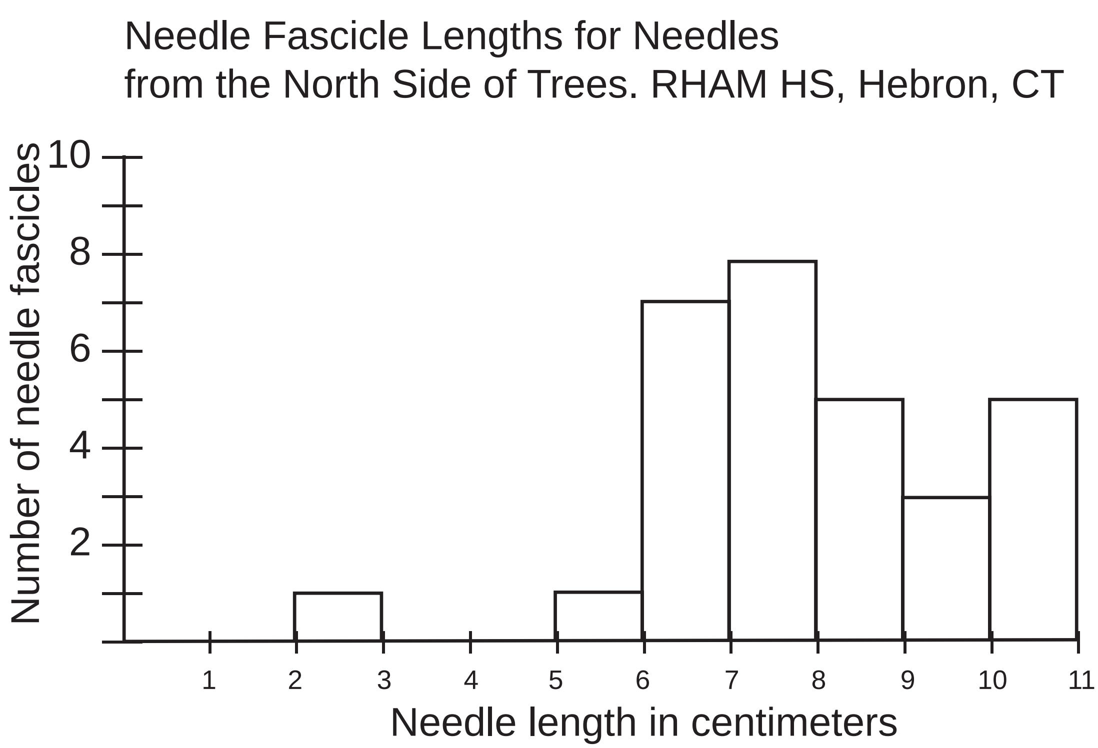
Histograms

The horizontal axis of a histogram is a measurement scale with the horizontal axis broken into intervals of equal size. To construct a histogram, break the range of values of a variable into convenient intervals and write appropriate values on the horizontal axis. The vertical axis lists the count of the data values that fall in each interval.

Histograms resemble stem-and-leaf plots turned on their side. Use stem-and-leaf plots for small data sets and histograms for larger data sets. It is generally not appropriate to place spaces between the bars because the horizontal axis has a measurement scale. Spaces are used, however, if the count for a particular interval is zero.

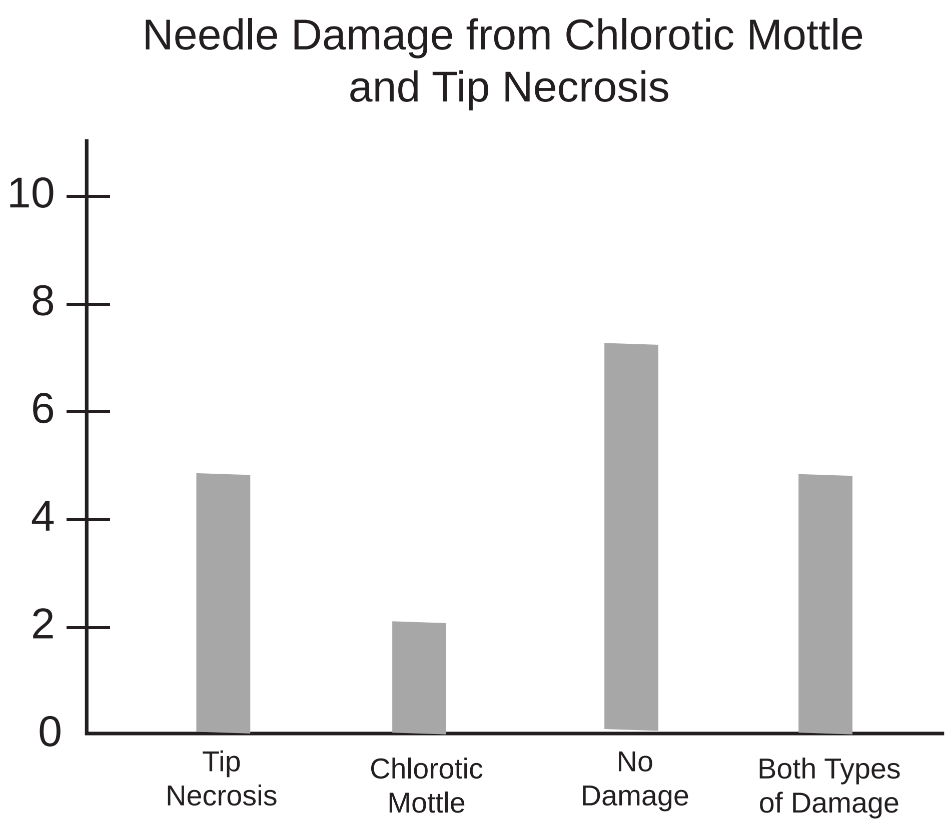
Histograms are sometimes called frequency graphs. Histograms can also be used to display relative frequencies. For relative frequencies the vertical axis is marked in percentages. The percent of observations that fall into each interval is graphed. The percents must add to 100%.

An example of a histogram is shown below. Notice the measurement values on the horizontal axis.



Bar Graphs

Bar graphs are used to compare the amounts of different types of items or categories as opposed to comparing the frequencies of one type of item over a continuum. The horizontal axis of a bar graph need not have any measurement scale but may simply identify the items being compared. Bar graphs usually have spaces between the bars because they are not depicting a measurement scale. The vertical axis records the counts. In the example below, notice the categories on the horizontal axis.



Comparison between the use of histograms and bar graphs

Use a histogram to show the distribution of frequencies or relative frequencies among data values that have been obtained through measurement. Use a bar graph to compare the number of items in different categories.

Answer Key

Histogram exercises:

1. Describe what the difference is between a number that is greater than or equal to 12 but less than 13 and a number that is greater than twelve but less than 13? How would you interpret the temperature situation when someone tells you that the temperature today is going to be in the twenties? Can you think of other examples of the use of greater than or equal to?

   A number that is greater than or equal to includes the lower limit number but not the upper limit number. Temperatures in the twenties imply any temperature equal to or greater than twenty, but not thirty. Age or time examples: he's 12 years old means he's either exactly 12 or some age up to but not including 13; the 90's includes the years from 1990 through 1999.

2. How many needle fascicles in the histogram above were between 7.0 cm and 8.0 cm? (This means equal to or greater than 7.0 and less than 8.0.)

   8 needle fascicles

3. What kind of information can you get about needle fascicle lengths from this histogram? Is there a clear mode of needle lengths on this graph? What is the mode?

   Most of the needles are between 6 cm and 10 cm. The mode in this case is between 7 and 8 cm.

4. What differences are there between the information you can obtain from a stem-and-leaf plot of the needle length data and from a histogram?

   Stem-and-leaf plots show every data value. Histograms do not, but they appear less cluttered.

5. When do you think histograms would be more useful than stem-and-leaf plots.

   Histograms are useful when the data set is larger than is practical to make a stem-and-leaf (greater than about 30).

6. Draw a histogram for the data from the south side of trees at RHAM High School. Describe the shape of the graph.

   The graph shape or distribution appears to be mound-shaped or normal.

7. Compare the histogram from the north side needle lengths with the diagram from the south side needle lengths. How are they similar and different?

   The needle lengths from the south sides of the trees are generally longer than the north side. The mode for the south side is between 8 and 9 cm while the mode for the north side is between 7 and 8 cm. The shape of the distribution on the south side is appears more normally distributed than in the north which lacks a clear shape. There is one low outlier in the north and no outliers in the south.

   Note to the teacher: Whether or not the differences in needle fascicle lengths are *significantly* different or whether they are really the same cannot be determined by these graphs. The box plots of the data might give a clue about that fact.

8. What information does a histogram allow you to read more or less easily than a box plot?

   A histogram tells you the shape of the distribution more readily than a box plot. It also tells you the count or percent of values in various measurement intervals. A box plot tells you the median and quartiles more readily and where 50% of the data lies. You can also use box plots to compare two data sets side by side and to check for overlap in the variation. This gives you a strong clue as to whether or not the data sets are really different from one another or just slightly different due to chance variation in sampling.

B

Bar Graph exercises:

1. When you have completed your graph, write a short paragraph that describes the differences in the average needle length among the four locations.

   The average length of the Tamworth, NH needles are the shortest at 7.3 cm. The average needle length at the Hollis School in NH is the longest at 8.9 cm. The difference between the average needle length for these two schools is 1.6 cm. It's difficult to tell whether or not this is significant without looking at the variation in the needle length data for the two schools. A box plot of the data would help to determine if this apparent difference in average needle length is important. The two ME needle samples differ by only 0.1 cm difference and are closer in length to the Tamworth sample than the Hollis sample.

2. What would happen to the bars if the increments on the vertical axis were increased from 0.5 to 2.0 (if you kept the size of the increments the same height)?

   Since all the data points lie within a 1.6 cm range, it would be difficult to graph the data and difficult to interpret the differences between the average needle lengths.

3. What would happen to the bars if the increments on the vertical axis represented 0.1 cm steps?

   The bars would be much longer and drawn out. This might serve to exaggerate the differences between the data bars.

4. Summarize how a change of scale on your vertical axis could affect your interpretation of the graph?

   The change of scale can artificially amplify a difference when there really isn't much of a difference if the increments are too small. If the increments are too large, then a significant difference may be minimized.

   Graph makers and graph interpreters both have to be aware of the scale and other parts of the graph when using this form of communication to get or give information. Graphs can be manipulated to enhance or detract from a particular interpretation!

5. Do you think location could be an influence on the length of needles? Explain why or why not.

   It would be interesting to look at a map of New England and see where the various schools are found. Why are the Hollis needles longer? Are there any geographic similarities between the three shorter schools? What are the wind patterns? Mountain ranges? Elevation? Distance from coastline? Etc.

7. Write a short paragraph that describes the difference in average needle length that you can observe from your graphs.

   Hollis has the overall longest average needle length for both the north and south quadrants. Tamworth has the shortest average needle length for both quadrants. The difference between the longest and shortest data set for the north needles is 1.3 cm. The difference between the longest and shortest needles in the south is 1.8 cm.

8. Compare North and South Quadrant needle averages. Is there a noticeable difference in needle average from one quadrant to another?

   See answer to # 7. We cannot tell from this data whether or not any differences are significant.

9. If so, brainstorm some factors that might account for this.

   Some data that students look at in the future might show larger differences between the quadrants. The needles on the north side of the tree may receive less sunlight than those on south side. To compensate for this, one adaptation might be that the needles grow longer to increase photosynthetic ability.

10. List three questions that you could ask as you observe this graph.

    Why is there a difference between Hollis and the other schools? Do you see this pattern with data from other schools? Does the location have anything to do with the length of needles? How can we get more information to help us determine if location is significant in estimating needle length?

11. How would you research and answer the question: "Is the pattern of average needle length in the north and south sides of study trees indicated by this clustered bar graph the same pattern that is observed in other Forest Watch study trees?" Would you expect to see this same pattern for other study sites in the Forest Watch Project? Why or why not?

    Look at data from several schools in a wide geographic area. Use more sophisticated statistical analysis that considers the range of needle length data as well as the mean needle length of each school and quadrant.

    AWV on expectations.

12. For each of the following situations, choose which representation: a stem-and-leaf plot, histogram, bar graph, or box plot would provide the most information in the easiest manner. You may use more than one for some! Provide reasons for your answers.

    a. You have a small set of data (n = 20) and you are interested in knowing the median and upper and lower quartiles.

    Use a stem-and-leaf or box plot.

    b. You have a large set of tree diameter data (n = 500) and you would like to know if the data is mound-shaped and whether the mean or the mode is an appropriate representation of the center of the data.

    Use a histogram. You must calculate the mean to compare with the mode.

    c. You have data from a forest that includes the number of trees that show several different injuries: recent evidence of deer chewing on the bark, insects nibbling on the needles and branches broken in a storm. You would like to compare the number of trees with each type of damage.

    Use a bar graph.

    d. You have data representing the heights of 100 trees in a particular area. You would like to know the range of heights in which 50% of the trees fall.

    Use a box plot or a stem-and-leaf.

FREQUENCY TABLES AND BAR GRAPHS

B

ACTIVITY

Background

In this activity, you will learn how to reorganize a large data set into a more easily interpreted table called a frequency table. The information in the frequency table can then be used to make a bar graph. This will provide you with two ways to display the same information. You may prefer using one or the other. One may quickly give you some information that the other does not and vice versa.

Procedure

Two Forest Watch students at Gilmanton School in New Hampshire collected data about chlorotic mottle and tip necrosis damage to needles. The table below shows the data they collected. A "+" indicates the presence of that type of damage. A "-" indicates the absence of that type of damage.

Table 1. Needle Damage Data, Gilmanton School.

Needle number	Chlorotic Mottle	Tip Necrosis
1	-	-
2	-	-
3	-	+
4	+	+
5	-	-
6	-	-
7	-	+
8	-	+
9	-	+
10	-	-
11	-	-
12	-	-
13	-	-
14	+	-
15	-	-
16	-	-
17	-	-
18	-	-
19	-	-
20	-	-

Study the table again. Suppose that you were interesed in analyzing needle damage in more detail. Let's say you want to know:

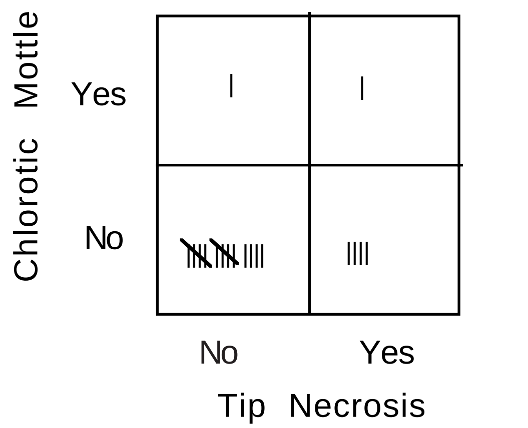
a) the number of needles that had only tip necrosis,
b) the number of needles that had only chlorotic mottle,
c) the number of needles that had both conditions at the same time, and
d) the number of needles with no damage at all?

1. Think of a way that you might represent the data table in a form that would readily help you show the information in a through d above. Draw your diagram in the space below.

2. Share your diagram with another student and ask him or her to tell you what information he or she learns from it. What information about the data can be interpreted easily? Is it the information that you intended?

One way to present the information is to use a frequency table like the one below. Perhaps this is similar to the choice you made.

Incidence of Chlorotic Mottle
and Tip Necrosis

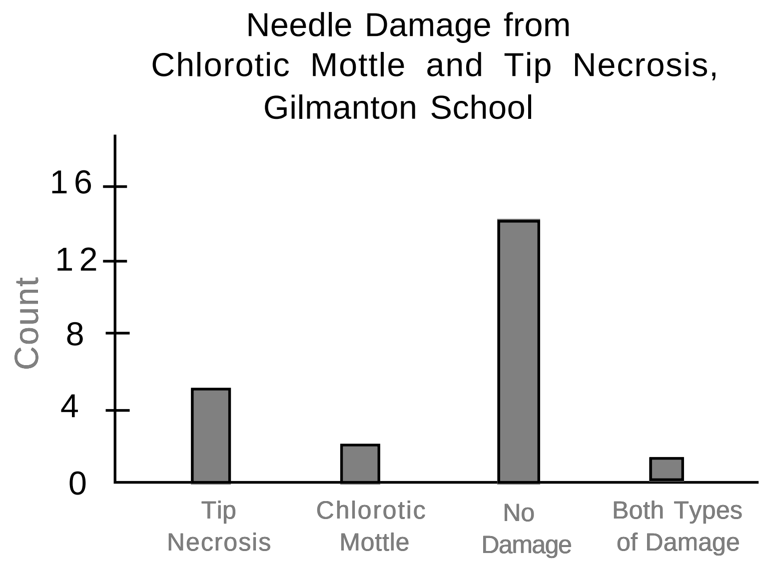


3. Write down the type(s) of damage (if any) indicated in each of the smaller squares in the frequency table:
   i. Upper left square
   ii. Upper right square
   iii. Lower left square
   iv. Lower right square

4. Use the frequency table to answer the questions:
   a) the number of needles that have only tip necrosis,
   b) the number of needles that have only chlorotic mottle,
   c) the number of needles that have both conditions at the same time, and
   d) the number of needles with no damage at all?

5. What do you observe about needle damage to the trees in Gilmanton? How could you find out if your observation is true about most white pine trees in Gilmanton?

Another way to represent the frequency data is to make a bar graph to represent the data about needle damage to trees in Gilmanton. Perhaps this is similar to the choice you made earlier in your representation.



6. Discuss with another student which kinds of information are more easily read from the bar graph than the frequency table and which kinds of information are more easily read from the frequency table?

7. Compare your diagram of needle damage with the bar graph and the frequency table. Is your diagram well labeled? If not, add labels now. What information is more or less easily read from your table than from the bar graph or frequency table?

8. Table 2. shows needle damage data collected by students at H. T. Wing School in Sandwich, Massachusetts, a study site located on Cape Cod.
   a. Make a frequency table that contains the data the Wing School students collected about chlorotic mottle and tip necrosis. Use the same frequency categories as illustrated in the Gilmanton data above so that you can easily answer the same questions about the occurrence of symptoms.
   b. Make a bar graph of the data.

Table 2. Needle Damage Data, H.T. Wing School, MA.

Needle number	Chlorotic mottle	Tip necrosis
1	+	+
2	+	+
3	+	-
4	+	+
5	+	+
6	+	+
7	+	+
8	+	+
9	+	+
10	+	+
11	+	-
12	+	+
13	+	+
14	+	+
15	+	-
16	+	-
17	+	-
18	+	-
19	+	+
20	+	+

   c. Compare the needle damage information between the Gilmanton School and the Wing School. Write a short paragraph that discusses differences and similarities between the number of needles showing: only tip necrosis, only chlorotic mottle, both tip necrosis and chlorotic mottle, and no tip necrosis or chlorotic mottle.

   d. What environmental factors might contribute to these differences? Could the differences occur just by chance?

9. a. Make up your own sample data table with a total of 20 needles. Include some needles with just chlorotic mottle, just tip necrosis, both symptoms present and no symptoms present. Make the table similar to the one above.

   b. Now draw a frequency table of the needle damage using the same categories as above.

   c. Trade your frequency table with another student. See if you can each work backwards and make up raw data charts to match the frequency table. Compare her data table with yours and discuss differences. Do they both accurately represent the information in the frequency table? In what ways are they similar and different?

10. Use data you or your class has collected to make a frequency table and bar graph showing the tip necrosis and chlorotic mottle on needles at your study site. Compare the damage between different trees in your site. Make comparisons with the damage to needles from Gilmanton, NH and Sandwich, MA

FREQUENCY TABLES AND BAR GRAPHS

TEACHER RESOURCE

Introduction

In this activity, students will learn how to reorganize a large data set into a more easily interpreted table called a frequency table. The information in the frequency table can then be used make a bar graph. This will provide students with two ways to display the same information.

After doing this activity, students should be able to easily make and interpret information from a frequency table and a bar graph, understand the benefits of using a bar graph and frequency table to organize data for future analysis and recognize that different data representations make different relationships easier to see. The use of many representations helps one see patterns in data that might not be evident otherwise.

Mathematics Topics

Bar graphs

Frequency tables

Data analysis

Curriculum Standards in Mathematics

Standards are listed by number

1. Mathematics as Problem Solving
2. Mathematics as Communication
3. Mathematics as Reasoning
4. Mathematical Connections
10. Statistics

Materials

- graph paper
- rulers
- values for needle damage data from Gilmanton School (found within activity)

Management Suggestions

Students will probably be able to do this activity with very little direction. They should work in pairs or small groups because they are asked to compare information.

After the students have done the activity, you should discuss the main goals with the class as a whole. That is, students should understand that different data representations make different relationships easier to see. They should represent data in several different ways as they search for patterns and try to find relationships.

Ask students when they would use bar graphs to represent data versus when they would use histograms.

Answer Key

3. Write down the type(s) of damage (if any) indicated in each of the smaller squares in the frequency table:

   i.. chlorotic mottle present but no tip necrosis.

   ii.. chlorotic mottle and tip necrosis.

   iii. neither type of damage.

   iv. tip necrosis but no chlorotic mottle.

4. Use the frequency table to answer the questions:

   a) the number of needles that have only tip necrosis: *4,*

   b) the number of needles that have only chlorotic mottle: *1,*

   c) the number of needles that have both conditions at the same time: *1, and*

   d) the number of needles with no damage at all?: *14.*

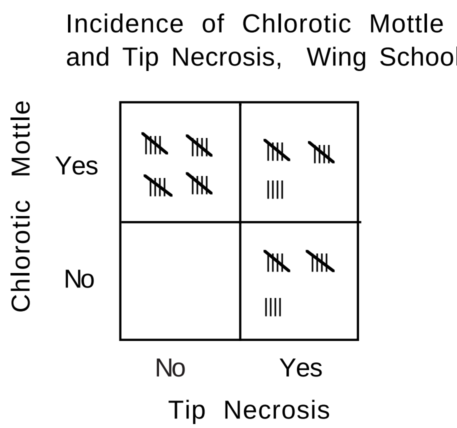
5. What do you observe about needle damage to the trees in Gilmanton? How could you find out if your observation is true about most white pine trees in Gilmanton?

   There seems to be relatively little chlorotic mottle or tip necrosis on needles in Gilmanton. The needles seem relatively healthy. To find out if this is true about most white pines in Gilmanton, you would have to randomly select many more needles from a much larger region.

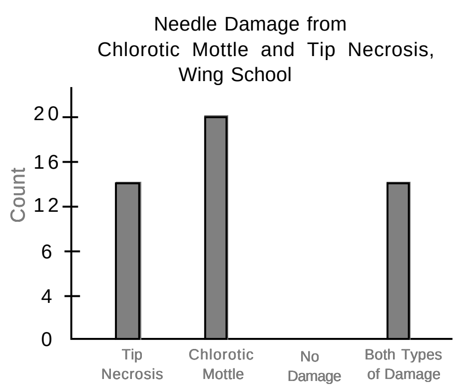
6. Discuss with another student which kinds of information are more easily read from the bar graph than the frequency table and which kinds of information are more easily read from the frequency table?

   Students might think from the frequency table it is easier to determine the number of needles with only chlorotic mottle or only tip necrosis. They may find the numbers easier to access. Others may suggest the comparisons of damage types at one site are easier to make from the bar graph. Large data sets are more easily represented on bar graphs. The main point is that each type of representation draws attention to different features of the data. It is generally helpful to use several representations when analyzing data and looking for patterns and relationships.

8 a. Make a frequency table that contains the data the Wing School students collected about chlorotic mottle and tip necrosis. Use the same frequency categories so that you can easily answer the same questions about the occurrence of symptoms as illustrated in the Gilmanton data above.



b. Make a bar graph of the data.



c. Compare the needle damage information between the Gilmanton School and the Wing School. Write a short paragraph that discusses differences and similarities between the number of needles showing: only tip necrosis, only chlorotic mottle, both tip necrosis and chlorotic mottle, and no tip necrosis nor chlorotic mottle.

There appears to be more damage to needles at Wing School. All of the needles at Wing School showed chlorotic mottle. 70% of the needles at Wing School also showed tip necrosis. No needles were without some form of damage. At Gilmanton School only two needles showed chlorotic mottle and five showed tip necrosis. Only 10% of the needles showed any damage at all. The needles appear to be much healthier at Gilmanton School. Perhaps the trees at Gilmanton School are in a location more protected from damage.

d. What environmental factors might contribute to these differences? Could the differences occur just by chance?

Wing School is on Cape Cod. The White Pine trees are regularly exposed to high winds and salt coming off the ocean. The wind carries the salt from the ocean over the Cape. There may also be blowing sand in the winter. It's possible that the trees on the Cape are in a more polluted area than the trees in the Gilmanton area.

SCATTER PLOTS AND LINES OF BEST FIT

B

ACTIVITY

Background

When analyzing data, the relationship between two sets of data is sometimes displayed using a graph called a scatter plot. The information is represented by individual points on the graph which are not connected the way they are in a line graph. A scatter plot is constructed with the specific purpose of seeing if a relationship exists between the sets of data being plotted. In analyzing the graph, one looks for positive or negative trends in the plotted data to help describe the relationship that might exist.

Scatter plots do not imply any cause and effect relationship between the two variables being compared. For example, even if the results of a scatter plot show a positive relationship between tree heights and DBH, it does not mean that tall trees cause larger DBHs or vice versa, nor that every tall tree will automatically have a large DBH.

Scatter plots are useful as a stepping stone to further scientific study about the relationship between two variables such as tree height and DBH.

Procedure

Part A- Scatter Plots

Study the table of tree measurements below.

Table 1. Diameter at breast height (DBH) and tree height of the five trees at the Gilmanton, NH study site.

Tree Number	Diameter, BH (cm)	Height (m)
371	18.6	6.3
372	14.7	8.9
373	17.1	10.4
374	25.6	10.3
375	24.4	8.8

1. What is the range of the height?
2. What is the range of the DBH?
3. What does the range tell you about the trees?
4. Calculate the average height.
5. Calculate the average DBH.

6. Order the tree height data from least to greatest. Order the DBH from least to greatest.

Some students at Gilmanton School hypothesized that the tallest trees in their sample also had the largest diameter and that shorter trees would have comparatively smaller diameters. Do you think that the trees with the highest height also have the largest diameter? To better visualize this you can create a graph called a scatter plot.

7. Draw a graph of the data from Gilmanton School using the following method:

   Remember, any graph you construct should be well labeled. Always be sure to include labels for both the x and y-axis and the units of measurement (in parentheses next to the label). Make sure you use correct increment ordering and a title.

   a. Label the y-axis (vertical) as **Height(m)** and the x-axis (horizontal) as **DBH(cm).** For the y-axis, use increments of 1 starting with 0. For the x-axis, use increments of 2 starting with 0.
   b. Title the plot: **Tree Height vs. Diameter: Gilmanton, NH, 1996 Data**. When a graph has the comparative "vs." in the title, the first part of the title is always the vertical (y) axis and the second part is always the horizontal (x) axis.

8. Now plot the data from Table 1 as points on the graph that represent the tree height and DBH information for each tree. You should have 5 points on your graph, one for each tree.

Table 2. Tree height and DBH data from Memorial Middle School in Laconia, NH.

Tree Number	Diameter, BH (cm)	Height (m)
206	38.5	12.1
207	36.3	12.3
208	31.6	14.5
209	24.6	10.1
210	38.4	15.9

9. Order the tree height data from least to greatest. Order the DBH from least to greatest.

10. What is the range of tree height and DBH for Memorial's study site trees?

11. Find the average tree height and average DBH for Memorial's trees.

12. Do you predict that you will see a similar trend in data on the graph from Memorial's data as you saw from Gilmanton's data?

    Let's see.

13. Use two different colored pencils (one for each school) to create a scatter plot that contains the data from both the Gilmanton and Memorial data sets. Be sure to label all axes correctly. The y-axis (Height) should have increments of 1 and the x-axis (DBH) should have increments of 2. Be sure to title this scatter plot:

    Tree Height vs. Diameter: Gilmanton and Memorial 1996 Data

14. Compare the ranges for the tree height and DBH for both schools. Compare the average tree height and DBH for the schools. You may make a table to display the data.

15. Brainstorm some reasons why average tree height and DBH might not be the same for both schools. Why might the range of the data sets not be the same?

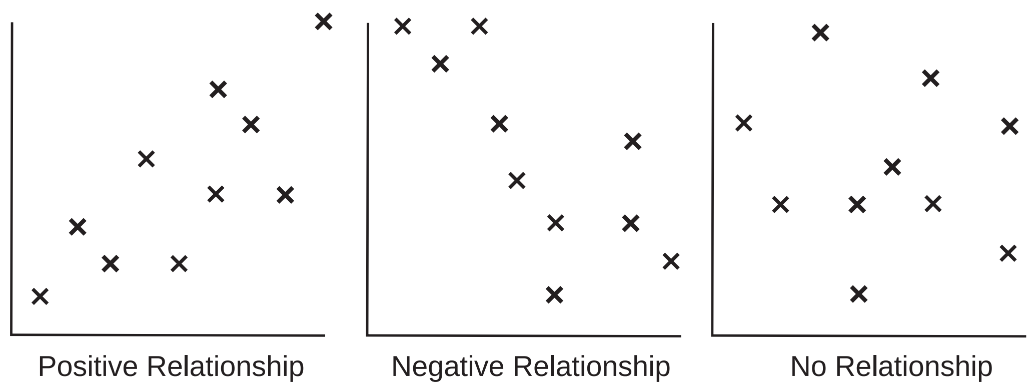
16. List some questions that are raised by these data and that you might be able to answer with more research.

Part B- Relationship between the variables in a scatter plot

Background

As you have probably noticed, the points on a scatter plot usually do not lie on a straight line. But the information displayed on a scatter plot can still be described in terms of noticeable trends in the plotted points. As you look for trends, it is a good idea to plot as many points on the graph as you can. In other words, a larger sample size will help show you a better representation of any trend that may appear in this pictorial representation of your data. The more data you have the greater the ability you have to illustrate what is really going on in your population. Trends in the data can be described as positive or negative relationships. You should also look for clusters of points in the scatter plot as you assess the relationship between the two variables.

Examples:



Positive Relationship -Notice in the left diagram, that the points appear to be scattered relatively close to an imaginary line with a positive slope (starting at the lower left and progressing to the upper right). We can say that there is a positive relationship between the sets of data used in the scatter plot.

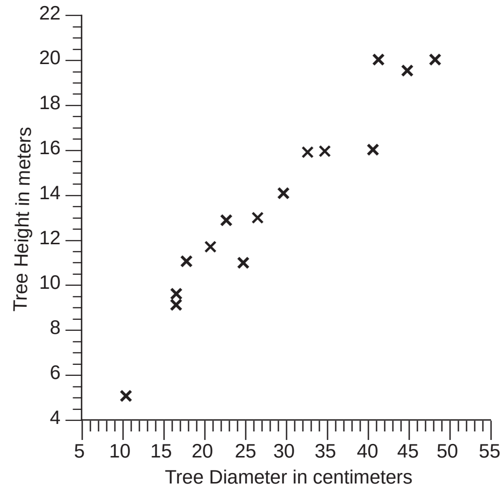
Negative Relationship -Notice in the middle diagram, that the points appear to be scattered relatively close to an imaginary line with a negative slope (starting at the upper left and progressing to the lower right). We can say that there is a negative relationship between the sets of data used in the scatter plot.

No Relationship -When there is no trend in the points, and the points appear to be widely scattered, we can say that there is no relationship between the sets of data used in the scatter plot.

B

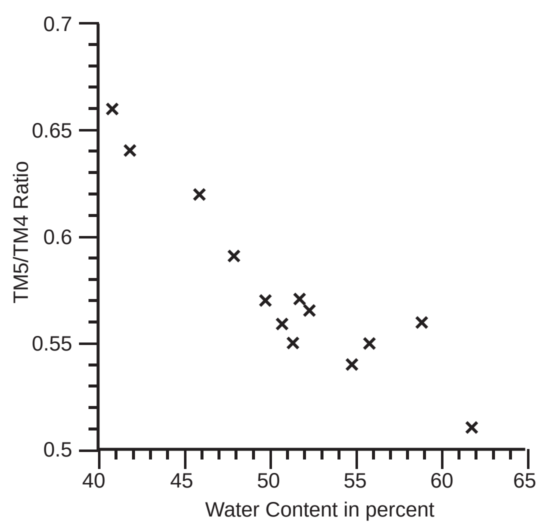
Examples of scatter plots that use data previously collected in the Forest Watch program follow.

The following scatter plot displays a possible relationship between tree diameter and the height of the tree. There is an upward trend in the points; therefore, we can say that there is a positive relationship between the diameter and the height of a given tree. In other words, the scatter plot illustrates that as the diameter of a tree increases, so does the height.



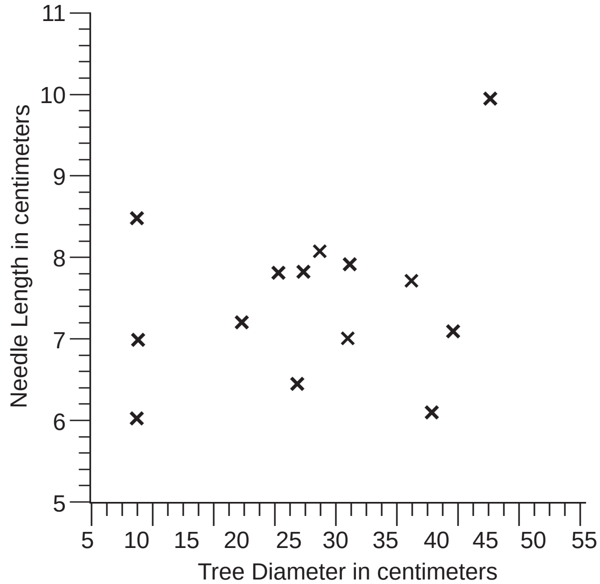
The scatter plot below investigates the possible relationship between needle water content and a spectral parameter, the TM5/TM4 ratio that is an indicator of plant stress. You will become familiar with many different parameters that can be compared as you start to sample your own trees.

Since there is a downward trend in the points, we can say that there is a negative relationship between needle water content and the TM5/TM4 ratio. This is an indication to us, that as the TM5/TM4 ratio decreases, the percent of water content increases. The stress of a plant seems to decrease as the water content increases.



The following scatter plot offers us a way to explore a possible relationship between tree diameter and needle length.

We see that the points follow no trend. Therefore, there is no relationship between tree diameter and needle length. In other words, there appears to be no relationship between the diameter of a tree and the length of the needles on that tree.



Exercises

Examine the following table of tree height information.

Table 3. Tree height at Sant Bani School, Sanbornton, NH.

Year	Average Tree Height (m)
91	11.6
92	12.8
93	13.7
94	14.6
95	15.8

1. Create a scatter plot depicting Sant Bani's Tree Height (vertical axis) vs. Year (horizontal axis). The year represents years since 1990. Remember to include all parts to be labeled and a title.

2. What type of relationship do you find represented on this scatter plot?

3. How much has the average tree height increased in this study site since 1991?

4. What is the average increase per year?

5. Predict the height of the average tree in Sant Bani's study site in 1996?

   If you have the data for the 1996 growing season, you will be able to compare your predicted answer to the actual tree height.

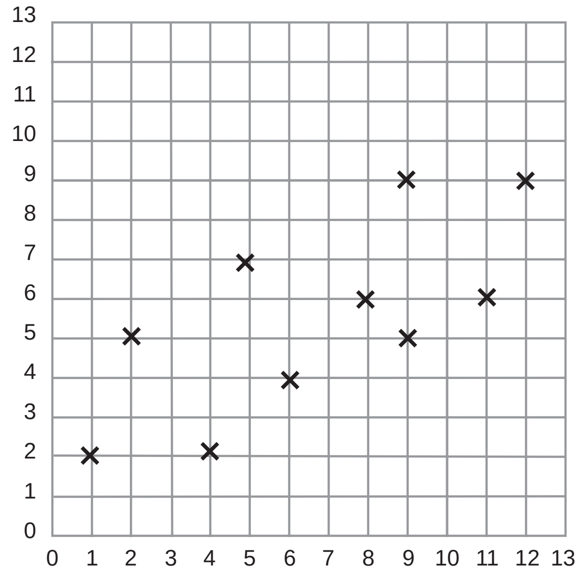
Part C- Fitting a Line

Background

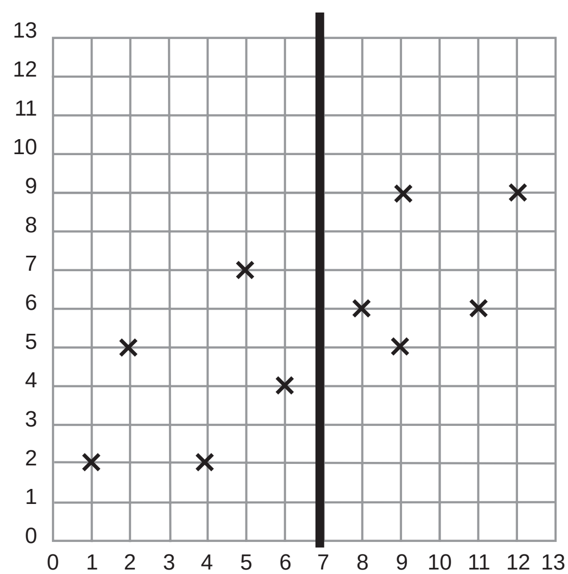
The **best fit line** is a mathematical model of the data on your scatter plot. You can use a mathematical model to represent the data and find values that the data does not give.

Once you have determined that there is a positive or negative relationship, you can determine where a line might be drawn through the points to give you a basis for some predictions. This line is called the *fitted line* or *best fit* line. What follows are step by step instructions for drawing this line.

Look at the example below:



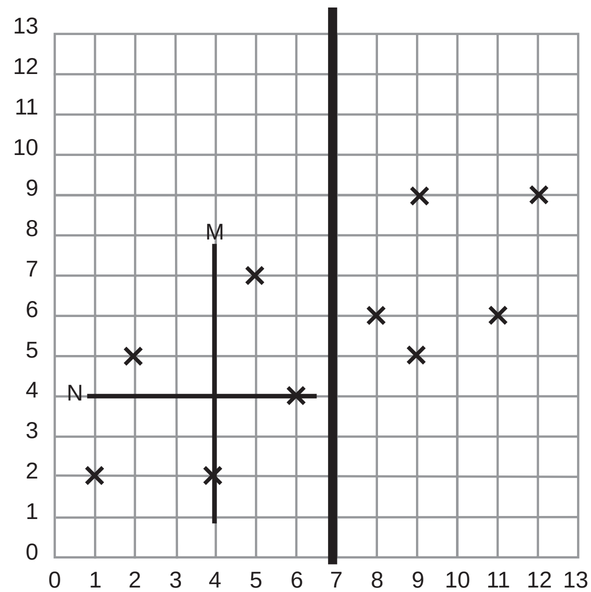
First, count the total number of points in the scatter plot. Draw a vertical line on the plot so that there is an even number of points on either side of the line. In this example, we have 10 points. Therefore, draw the line so that there are five points on either side of the line. (Note: If there is an odd number of points, your line will pass directly through the point in the middle.)



Working with the Left Set of Points

Working with the left set of points, you now want to find the center of the points on this side, in other words, you want to locate the median of the data points. Looking at the points horizontally, count the number of points from left to right. Find the point in the middle. The median in our example is the third point. Draw a vertical line that passes through this point. (Line M in the example) Note: If there is an even number of points on your scatter plot, draw the line between the two middle points.

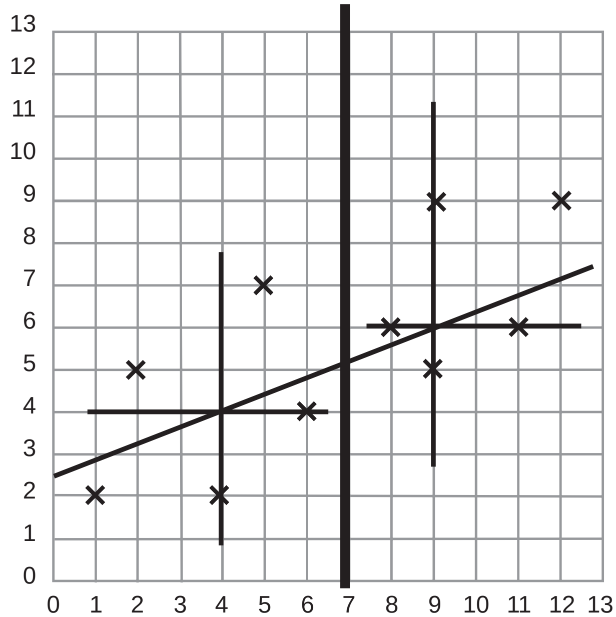
Next, looking at the same set of points vertically, count the number of points from top to bottom. Since the third point is in the middle, draw a horizontal line through this point. (Line N in the example). Note: If there is an even number of points on your scatter plot, draw the line between the two middle points. The intersection of these two short segments indicates the median for the set of data.



Do the same process with the data points to the right of the original line.

To get the fitted line for this set of data, with a straight edge, draw a line that connects the two points of intersection.

The fitted line, also called the best fit line allows you to make some predictions based on the information you have plotted. It is important to note that predictions can be made only within the range of the data.



To get the equation for the best-fit line, use the two points of intersection. From your graph, write the coordinates of the two points of intersection : (4,4) and (9,6). Use these points to find the slope (m) and y intersection (b) in **y= mx + b**

Example:

(x,y)   (x,y)

Points of intersection: (4,4) and (9,6)

Using these two points you can now determine the <u>slope</u> of a best fit line.

The slope of a line is the difference of the vertical axis values ÷ the difference of the horizontal axis values.

Find the slope, if slope = m:

> To do this, use the ordered pair (9,6) and (4,4).
>
> slope = m = (6-4) ÷ (9-4) = 2/5
>
> m=2/5

Find the y intercept, b. Substitute 9 in for x and 6 in for y in:

> **y = (2/5) x + b**
>
> 6 = (2/5) (9) + b
>
> 6 = 18/5 + b

Now solve for b:

6 = 18/5 + b

$b = {}^{30}/_{5} - {}^{18}/_{5} = {}^{12}/_{5}$

$b = {}^{12}/_{5}$

Now substitute your m and b values into y = mx + b for the equation of your best fit line.

$y = ({}^{2}/_{5})\ x + {}^{12}/_{5}$

Repeat this activity using the graph that you generated: Sant Bani's Tree Height vs. Year

1. Draw a vertical line at year 93. This represents the median year of the data set (91,92,93,94,95).
2. Draw a vertical line for the median year value for the left set of data points. This vertical line is the average between 91 and 92 (or 91.5). Draw a horizontal line for the median height value for the left set of data points. This horizontal line is the average between 11.6 and 12.8 (12.2). The intersecting points for the left half of the data should be (91.5, 12.2).
3. Repeat this process for the right set of data points. The average between 94 and 95 (or ______) and 14.6 and 15.8 (or _______). The intersecting points for the right half of the data should be (94.5, 15.2).
4. Using these two points, you can now find the slope, m, of the best fit line.
   The work is started for you.
   slope = m = (15.2 - 12.2) ÷ (94.5-91.5) =
5. You should have calculated m = 1.0. Use the formula: **y = (1)x + b** to solve for b, where b is the y intercept.
6. Write the equation of your best fit line here.

If you want to figure out the average height of the tree sample in a particular year, put the year in for x. If you want to find the year that a tree was a particular height, put the height in for y and solve for x.

7. Assuming that a tree grows at the same rate every year based on the prediction or model represented by the slope of the best fit line, find the year that the tree first started to grow. Use 0 for the height.

8 a. Use the equation of the best fit line to find the value of x when x = 94.

b. What was the data value when x = 94?

9. Explain why the y value of the best fit line when x = 94 is different than the data value at x = 94.

SCATTER PLOTS AND LINES OF BEST FIT

TEACHER RESOURCE

Mathematics Topics

Scatter plot

Correlation

Fitting a line

Management Suggestions

Within the Scatter plot activity, be sure to allow students time to answer the question: What does the range tell you about the trees? Guide the students toward the idea that perhaps the height of a tree might have a relationship to its age. If you have access to the yearly compilation of Forest Watch data, ask students to find the range of average DBH and tree height for all Forest Watch sites.

During the Best Fit Line activity students look at a best fit line that represents the average yearly growth of tree height of one meter. Ask students to refer to the Gilmanton DBH and Tree Height data table at the beginning of the Scatter Plot section. Revisit the students' answer for the question: what does the range tell you about the trees? from the scatter plot section. Then ask the question: "If the best fit line told us that the average tree height growth per year was one meter, then could you estimate the age of Tree #371-375 and #206-210?" They will probably make the association that Tree #371 with a height of 6.3m should be 6.3 years old. Ask: "What could we use to prove the age of the tree?" Core sampling might be one method. This could be done for you by a local forester or Forest Watch staff.

Answer key

Part A- Scatter Plots

1. What is the range of the height?

   4.1 m

2. What is the range of the DBH?

   10.9 cm

3. What does the range tell you about the trees?

   It tells us that there are a variety of sizes of trees, perhaps the trees of different heights have different ages.

4. Calculate the average height.

   8.94 m

5. Calculate the average DBH.

   20.1 cm

6. Order the tree height data from least to greatest. Order the DBH from least to greatest.

   DBH (cm) = 14.7, 17.1, 18.6, 24.4, 25.6   height (m) = 6.3, 8.8, 8.9, 10.3, 10.4

9. Order the tree height data from least to greatest. Order the DBH from least to greatest.

   DBH (cm) = 24.6, 31.6, 36.3, 38.4, 38.5   height (m) = 10.1, 12.1, 12.3, 14.5, 15.9

10. What is the range of tree height and DBH for Memorial's Study site?

    5.8 m, 13.9 cm

11. Find the average tree height and average DBH for Memorial's trees.

    12.9 m, 33.9 cm

12. Do you predict that you will see a similar trend in data on the graph from Memorial's data as you saw from Gilmanton's data?

    One would not expect to see the exact same results, but they may show a similar trend.

14. Compare the ranges for the tree height and DBH for both schools. Compare the average tree height and DBH for the schools.

	Memorial Middle School	Gilmanton School
Range of Tree height (m)	5.8	4.1
Range of DBH (cm)	13.9	10.9
Average Tree height (m)	13.0	8.94
Average DBH (cm)	33.9	20.8

Both the range of height and range of DBH for Memorial were greater than Gilmanton's. The same is true for the average tree height and the average DBH.

15. Brainstorm some reasons why average tree height and DBH might not be the same for both schools. Why might the range of the data sets not be the same?

    The stand of trees that Memorial is sampling from might be a taller stand than Gilmanton's. The taller trees may have larger diameters because they are older or have a larger space in which to grow.

16. List some questions that are raised by this data and that you might be able to answer with more research.

    If the average height of trees at Memorial Middle School is greater than at Gilmanton, how do students obtain their needle samples at midcanopy? Are they still able to use pole pruners? Do they need to borrow a bucket truck?

    How close together are the trees in each of these study sites? Can we determine this from the canopy closure or ground cover information provided in the annual data sets?

    Are the sites similar in biological composition? Are there mostly white pines in the stand? Is there a mix of conifer and deciduous? Do all the trees have generally the same height or is there much variation in height?

B

Part B- Relationships between variables in the scatter plot

2. What type of relationship do you find represented on this scatter plot?

   positive

3. How much has the average tree height increased in this study site since 1991?

   4.2 m

4. What is the average increase per year?

   0.84 m

5. Predict the height of the average tree in Sant Bani's study site in 1996?

   16.6 m

Part C- Fitting a Line

3. Repeat this process for the right set of data points.

   The average between 94 and 95 (or <u>*94.5 cm*</u>*) and 14.6 and 15.8 (or* <u>*15.2 m*</u>*). The intersecting points for the right half of the data should be (94.5, 15.2).*

4. Using these two points, you can now find the slope, m, of the best fit line. The work is started for you.

   slope = m = (15.2 - 12.2) ÷ (94.5-91.5) = 1

   Find the slope: *m = (15.2 - 12.2) ÷ (94.5 - 91.5) = 3/3* *m = 1*

5. You should have gotten m = 1.0. Use the formula: **y = (1)x + b** to solve for b, where b is the y intercept.

   Use the line of best fit formula:

   y = (1) x + b

   15.2 = (1) (94.5) + b

   15.2 = 94.5 + b

   b = 15.2 - 94.5

   b = -79.3

   y = x - 79.3

6. Write the equation of your best fit line here.

   y = x - 79.3

7. Assuming that a tree grows at the same rate every year based on the prediction or model represented by the slope of the best fit line, find the year that the tree first started to grow. Use 0 for the height.

   y = (1)x - 79.3

   0 = (1)x - 79.3

   79.3 = x *1979*

8. a. Use the equation of the best fit line to find the value of x when x = 94.

   y = (1)x - 79.3

   y = 94 - 79.3

   y = 14.7

   b. What was the data value when x = 94?

   Data value from the data table in 1994 was 14.6m.

9. Explain why the y value of the best fit line when x = 94 is different than the data value at x = 94.

   The best fit line allows us to look for trends in the data. 14.7 represents the predicted height in meters in 1994 according to the best fit line. 14.6 is the measurement that Sant Bani students actually measured in 1994.

DATA SETS FOR NEEDLE FASCICLE LENGTHS

RHAM High School, Hebron, CT. and Gilmanton School, Gilmanton, NH.

RHAM High School, Hebron, CT. Needle fascicle lengths measured in centimeters.
North side of study trees. (n = 30)

6.8	6.7	6.7	7.0	6.5	7.1	6.5	6.6	7.5	7.2
8.5	10.5	8.0	9.5	8.7	7.5	6.4	7.7	5.0	2.4
7.4	8.7	8.8	9.7	10.1	9.3	10.0	10.1	10.2	7.0

RHAM High School, Hebron, CT. Needle fascicle lengths measured in centimeters.
South side of study trees. (n = 30)

6.8	6.9	7.2	7.3	7.5	7.4	8.3	7.5	6.9	6.8
9.7	8.4	8.0	8.6	10.5	8.4	7.4	7.9	7.7	8.7
8.2	9.0	9.0	8.5	8.6	8.9	8.9	8.8	8.6	6.3

Gilmanton School, Gilmanton, NH. Needle fascicle lengths measured in millimeters.
North side of study trees. (n = 30)

80	71	71	73	75	70	79	65	74	73
79	88	87	72	67	84	86	87	73	64
50	49	57	60	62	61	60	45	65	56

Gilmanton School, Gilmanton, NH. Needle fascicle lengths measured in millimeters.
South side of study trees. (n = 30)

36	69	69	67	59	60	67	66	65	64
64	64	55	63	54	66	60	62	52	65
58	66	51	74	58	63	69	73	70	64