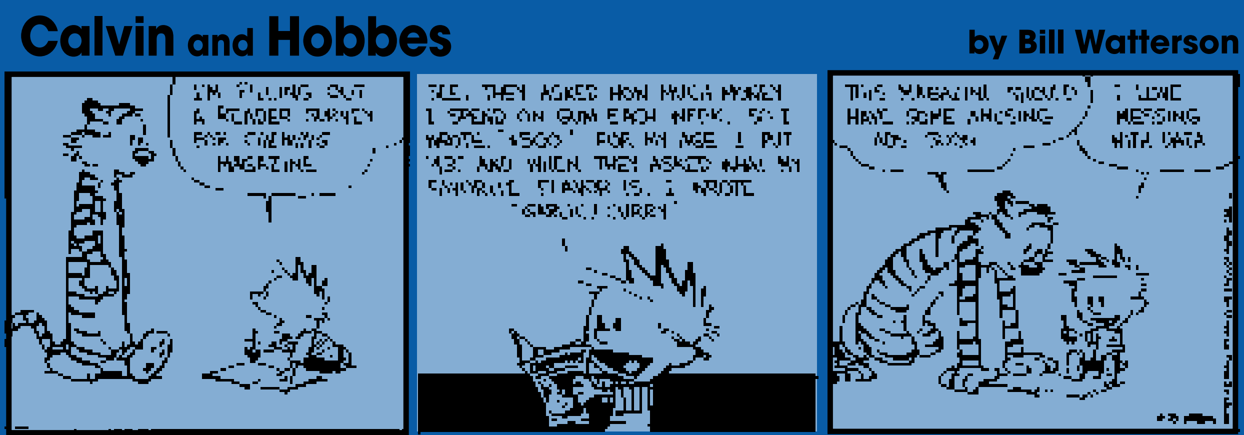

Calvin and Hobbes
by Bill Watterson
I'M FILLING OUT A READER SURVEY FOR CHEWING MAGAZINE
SEE, THEY ASKED HOW MUCH MONEY I SPEND ON GUM EACH WEEK, SO I WROTE, "$500." FOR MY AGE, I PUT "43," AND WHEN THEY ASKED WHAT MY FAVORITE FLAVOR IS, I WROTE "GARLIC CURRY."
THIS MAGAZINE SHOULD HAVE SOME AMUSING ADS SOON.
I LOVE MESSING WITH DATA.

LEARNING GOAL

Students will ask questions and collect, organize, interpret, and transform data related to those questions to communicate information and make predictions and decisions.

Components

With this learning goal in mind, Minnesota students will have the opportunity to:

1 formulate a question(s), determine necessary data, and choose an appropriate method of data collection in order to make sense of a problem or situation

2 collect, organize, and represent data in a variety of ways

3 read, describe, and interpret displays of data

4 formulate, justify, and communicate conclusions, arguments, predictions, decisions, and further investigations based on data

(For more detailed information, see pages 17-25 in this content section.)

Reflections

"A knowledge of statistics and data analysis helps students answer questions, make decisions, and develop predictions using information. Learning statistics can be a vehicle to give meaning to computational skills that are too often isolated from relevant contexts."

Friel & Corwin, 1990

Reflections

" Three critical processes are fundamental to all quantitative data analysis activities at every grade level:
- *Characterizing the overall shape of the data*
- *Summarizing the data*
- *Interpreting the data"*

Russell & Friel, 1989

" Graphing software raises conceptual issues and concerns about the display of data. Since software creates a variety of graphs, students need to understand what graph(s) are appropriate to represent the given data."

Friel & Corwin, 1990

" ...we respect numbers and we cannot help believing them. Yet, more and more of the information we use to buy, elect, advise, acquit and heal has been created not to expand our knowledge but to sell a product or advance a cause."

Crossen, 1994

" Statistics is important in its own right—more important than calculus in most occupations..."

Moore, 1990

We live in a society that bombards us daily with data—weather and crop forecasts, birth and death statistics, sports averages and percentages, neighborhood and school demographics, investigations of racial or gender bias, medical data from clinical trials, and employment and financial statistics. "There are compelling practical reasons to learn statistics" (Moore, 1990). Students who develop skills in dealing with data will be better able to make intelligent decisions and formulate arguments based on quantitative information. These abilities will enhance their social awareness and career opportunities. Data investigation also supports a student's developing mathematical literacy by putting numbers in context, and embedding arithmetic and graphing applications in problem solving situations.

All data investigations **begin with a question** or an "I wonder if…" statement. These experiences should arise from meaningful situations in which students have a need to deal with data to answer their individual or collective questions. Asking good questions is a skill that takes time and teacher assistance to develop. A well-conceived design for collecting data is a major component to a solid investigation. Formulation of questions is an iterative process that might involve some preliminary data collection and exploration.

Students can and should **deal with real data** that they collect or obtain from real-world sources. Real data are "messy" rather than contrived. They are not neatly organized and often deal with too much or too little information. Advances in technological tools make real world data manageable for students and eliminate instructional reliance on artificial data sets.

The **process of data analysis** is both linear and cyclical. Some tend to think of data investigation as a recipe: ask a question, collect data, analyze it, report it. In reality, people often collect data and realize they need to stop and refine the question before continuing. Others summarize the data and identify the need to collect more data. The direction of data investigation is more often "back and forth" than "straight ahead." Encouraging students **to take a "big picture," problem solving view** of data analysis, rather than a step-by-step approach, will help them process and evaluate the usefulness of their work as they progress.

Graphs are not ends in themselves. The primary purpose of a graph or data display is **to facilitate data analysis and communication**. A simple, understandable graph that maintains a balance between technique and communication is the ideal. Student work over time should indicate progression in the level of data analysis—not simply the ability to make fancier, more colorful graphs.

For the same reason, a graph is not complete until students **write a summary or interpretation** of the information displayed. The question, method of data collection, and conclusions reached by an individual or class should be clear to a visitor observing a data display posted in the classroom or hallway.

The well-developed question that initiates a data investigation often leads to more questions, more investigations and sometimes to decisive action. In the classroom, students should be encouraged to **reflect on who would find their data results helpful**. Emphasis in data investigation is less on formality and more on active student participation in using data to interpret their world mathematically. Students should learn to act on the results of their investigations, develop an awareness of the power of data analysis for decision making, and realize how others use data to influence them.

Data investigation can be a **powerful means of connecting mathematics** to the lives of children, to news events, and to other disciplines in the curriculum. It can be the source of challenging mathematical work and rich discourse among students and between students and teachers. It takes time to ask a good question, design a solid data collection plan, explore various displays of the data, and interpret and communicate the information to others. An effective teacher of mathematics knows the process of data analysis takes time and plans appropriately.

Vignette: Data Investigation in the Classroom

In honor of Dental Health Week, Ms. B had managed to get a number of different toothbrushes donated to the classroom from a variety of vendors. She had also encouraged parents, if possible, to purchase and send new, unopened toothbrushes to school for class use.

Ms. B saw an opportunity to make a "real world" connection to mathematics.

Ms. B was seeking a mathematics task for her 4th-5th grade combined classroom in which her students could develop their understanding and skills in data investigation. Her goal was to have the students focus on:

- sampling
- estimating
- predicting

She saw a way to integrate some of the work on Dental Health Week with her mathematics focus.

Students may bring in connections from their life experiences.

MS. B: Now that you have had some time to sort and compare these new toothbrushes, what are your observations?

STUDENT: They aren't all the same price. Some cost more than others.

STUDENT: They are made by different companies and have different designs. Some have lots of little clumps of brushes; some have fewer.

STUDENT: There are toothbrushes for adults, and others for children. Some have different labels, like hard, medium and soft.

(Students continued to share their observations while Ms. B recorded them on chart paper.)

MS. B: I'm going to start a new list and label it "our questions." Do your observations make you more curious about toothbrushes? Are there any questions you are wondering about?

STUDENT: I wonder if you can get rich selling toothbrushes?

Data investigations are often most successful when they are triggered by students' own questions.

STUDENT: Is there one brand of toothbrush that dentists recommend because it works better?

STUDENT: Do they all have the same number of hairs on the brush or are the toothbrushes that cost more bushier?

STUDENT: When you say "hairs," are you referring to the bristles on the brushes?

STUDENT: Oh, yeah, I guess that's what I mean.

(Students continued to share their questions, while Ms. B recorded them. Then as a class, they discussed which questions were most interesting to them. They also discussed what they would have to do to begin to answer different questions: do an experiment, ask people questions, check the Internet for information, etc.)

MS. B: You all seem curious about which toothbrushes have more bristles. What are your predictions?

STUDENT: The adults' toothbrushes have more than the children's brushes.

STUDENT: I think the hard ones are thicker than the soft ones.

STUDENT: I think the expensive ones have more bristles than the cheap ones.

STUDENT: I think all those are right.

(Ms. B gave each student a toothbrush and asked him/her to estimate how many bristles it had. Numbers flew around the room: 100, 750, 300, 24, 1000. She listed the numbers on a large sheet of paper as they were called out.)

Part of the estimation process is to check whether an estimate is reasonable and then reevaluate the validity of estimates that do not seem reasonable.

MS. B: There doesn't seem to be much agreement in your answers. Your estimates range from 24 to 1000. Why do you think your answers are so varied?

STUDENT: Our toothbrushes are all different.

MS. B: How would that affect our estimates?

STUDENT: The adults' toothbrushes are bigger than the children's brushes. They have more rows of bunches.

STUDENT: What's a "bunch?"

Mathematical terms, such as "range," or other terms, such as "bunch," should be discussed.

STUDENT: You know, those little groups. I see that my toothbrush has maybe 10 or 20 bunches. A kid's toothbrush maybe has 5 or 10 bunches. If they have more bunches, then they are going to have more bristles.

MS. B: Are you sure?

STUDENT: I think I'm sure.

Assumptions need to be checked before making calculations.

MS. B: Maybe we should check that out.

MS. B: Now, we think that there may be different numbers of bristles because some people have adults' and some have children's brushes. We think that some brushes have more bunches. Are there any other reasons why we have such different answers?

STUDENT: If there can be different numbers of bunches on a toothbrush, I guess there can be a different number of bristles in a bunch.

MS. B: How many bristles do you think there would be in a bunch?

STUDENT: Maybe 10 to 20. Certainly not 2 or 3 and not 100 or more.

Estimates can be a number or a range.

This student seems clear about the difference between a guess and an estimate.

MS. B: How do you know how many bristles and bunches there are in your toothbrush?

STUDENT: I guessed! I would need to take it out of the wrapping to look at it closer to get a better estimate.

Students will refine their estimates by discussing and defending approaches with their peers.

MS. B: This would seem like a good time to take the toothbrushes out of their packaging. But before you do, please record in your notes the brand name and cost of your toothbrush. Then work in pairs to come up with a revised estimate. I'll write your estimates on the board after you've had enough time to think and talk.

(…buzz in the room as the students negotiate their estimates…)

MS. B: Looks like we have had many revisions. 160…200…300…750. What do you notice about our revised estimates?

STUDENT: Well, most of the high numbers and most of the low numbers have disappeared.

STUDENT: The numbers are closer together.

MS. B: What happened in your groups to get these revised estimates?

STUDENT: We got 200.

Ms. B is asking for strategies, not answers.

MS. B: Remember, I am more interested in what you and your partner were thinking than in your answer.

STUDENT: We thought about how many bunches there could be and how many bristles there could be in a bunch. There has to be at least 20 bunches and at least 10 bristles in each bunch. That's 200.

With two strategies introduced, the class can discuss the advantages of each.

STUDENT: We talked about rows of bunches. I think my brush has 3 rows of bunches and there are about 10 bunches in a row. That's 30 bunches. If there were 10 bristles in a bunch, that makes 300 bristles.

MS. B: So, we've now come up with counting bristles, bunches, and rows. When we use bunches or rows, we get estimates. The estimates have a smaller range than when we guessed the number of bristles at the beginning. How did you come up with a smaller range?

The student has introduced the concept of outliers and checking their reasonableness.

STUDENT: We realized that some answers didn't make sense. There's no way that there could be only 20 bristles.

STUDENT: But we couldn't agree on the number of bunches or the number of rows. We know that the answer is in the 100s, but someone thought it was about 300 and someone thought it was more like 800.

MS. B: How do you think we should figure out exactly how many bristles there are in our toothbrushes?

STUDENT: We could count the bristles.

MS. B: Should we just count the bristles in my toothbrush?

The student recognizes that a sample is more reliable than relying on one piece of data.

STUDENT: No, let's count each of our brushes. Yours might be one of those expensive ones, and I still think they have more bristles than the regular toothbrushes.

MS. B: How could we determine the number of bristles?

The student has developed an algorithm for the total number of bristles based on the prior discussion.

STUDENT: Maybe we could each count the number of bunches and figure out how many bristles there are in a bunch. Then, we would multiply the number of bunches by the number of bristles. That would give us the total number of bristles.

STUDENT: That seems a lot easier than counting all of those bristles.

STUDENT: And probably better than just guessing.

Writing in a math journal often requires the student to explain her/his reasoning and helps the teacher to assess the student's understanding.

MS. B: Okay. Tonight I want each of you to determine how many bristles your toothbrush has. In your math journals, describe how you determined the number of bristles in your toothbrush and how your answer compares to your estimate from this afternoon. In class tomorrow, we will compare bristle counts. Oh, yes, and remember to brush tonight and tomorrow morning, too.

The next morning, there was much conversation about how many bristles each student's brush had. It was very competitive, as if having more bristles were better. Students questioned each other's strategy. All agreed that counting the bristles was far harder than expected. Some students had their brushes in hand as "Exhibit A" for their assignment.

Students' methods vary from sophisticated estimating to concrete counting.

MS. B: How did you determine the number of bunches in your brush?

STUDENT: Easy. I counted them. There were 26.

STUDENT: I didn't have to count. There were 3 rows and 10 in each row, plus two extra at the top and three at the bottom. That's 35.

Some discussion of whether each bunch would have the same number of bristles may strengthen the idea that these are still estimates, highly refined.

MS. B: How did you determine the number of bristles in a bunch?

STUDENT: I estimated. It looked like 25.

STUDENT: I held a bunch down and let them pop up one by one very slowly. They bounced up too fast the first few times. But I think that I'm pretty close at 30 bristles.

STUDENT: I held down a bunch and I had my sister cut off the tops of the bunch—very carefully. Then I counted the little pieces. I think I got them all. I had 43. It's a mess and my brush is smaller now, but I know that my answer is right.

Here the students have collected data and are now organizing it. This isn't as easy as they sometimes think it will be.

MS. B: Let's take a look at our information. How will this data help us answer our question from yesterday? Which brushes have more bristles?

STUDENT: Obviously, look at the number of bristles.

MS. B: But how will you know which brushes have more bristles?

STUDENT: We need to sort them and see if there are any differences. Like, do the adults' toothbrushes have more bristles than the kids' brushes.

MS. B: How could we show this information so that it would be easier to understand?

(Silence for quite awhile.)

STUDENT: We could make a table with the number of bunches in one column and the number of bristles in another.

STUDENT: I think we should also include a column for the number of bristles in a bunch, because they aren't all the same.

STUDENT: We could make a graph.

STUDENT: We could do both.

Evaluating the table prepares students for determining how to graph their data.

Bunches	Bristles in a bunch	Total bristles	Bunches	Bristles in a bunch	Total bristles
28	30	1140	38	36	1368
23	32	736	35	60	2100
21	25	525	40	30	1200
26	40	1040	38	40	1520
50	43	2150	38	55	1870
40	50	2000	50	20	1000
25	25	625	10	60	600
23	30	840	20	40	800
29	68	1972	60	30	1800

Upon seeing a pattern in the number of bristles, the class seeks an explanation.

MS. B: As I look at this table, I see that our answers range from 525 to 2150 bristles. There are a lot more bristles than we initially expected. Why do you think that is so?

STUDENT: The bristles are so small, it doesn't look like 2150 bristles.

Questions like Ms. B's encourage students to look beyond expected patterns.

MS. B: Does anything in the data surprise you?

STUDENT: You were right. All toothbrushes are not the same.

STUDENT: Even when we counted, we got really different numbers.

Students begin to see the relationship or association of variables.

MS. B: Is there anything else you see in this table?

STUDENT: The more bunches there are, the more bristles there are.

STUDENT: Also, the more bristles there are in a bunch, the more bunches in the whole toothbrush, I think—but one has 60 in a bunch and only 10 bunches.

MS. B: Maybe it's time for us to graph these relationships.

The class has been introduced to independent and dependent variables.

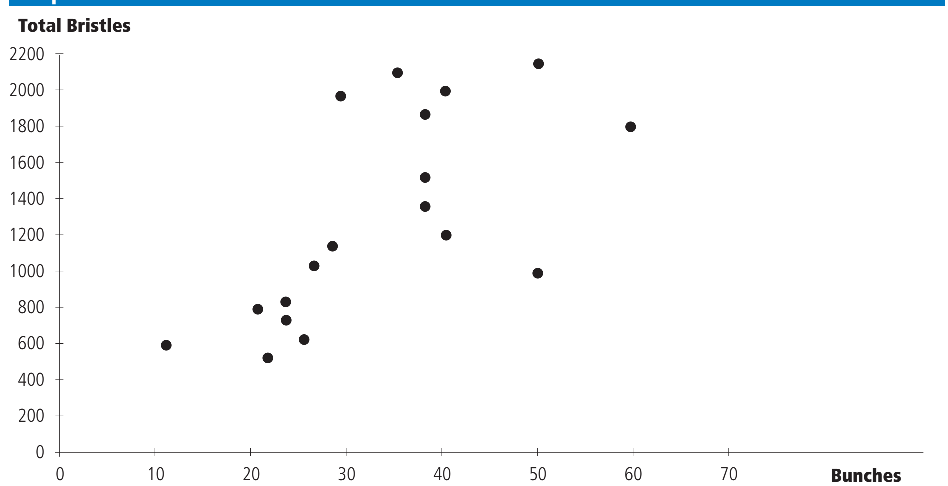
STUDENT: Let's do "bunches" on the horizontal and "total bristles" on the vertical or "bristles in a bunch" on the horizontal and "bunches" on the vertical.

MS. B: Let's try each.

The students will discover which graph is most useful.

(Ms. B has the groups graph the data in both ways. They post their graphs for comparison.)

Graph 1 – Toothbrush Bunches and Total Bristles



Students see graphs as a tool for prediction and begin to understand the concept of correlation.

The size of the sample will affect the reliability of the prediction but only if the data is accurate in the first place.

MS. B: What do you see from the bunches graph (Graph 1)?

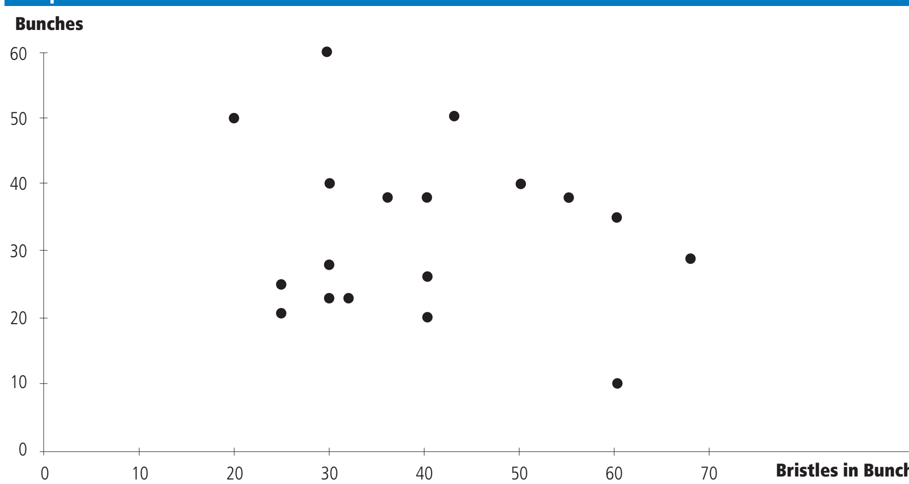
STUDENT: The more bunches, the more bristles.

STUDENT: If I know the number of bunches, I could predict the number of bristles pretty close.

STUDENT: If there are less than 30 bunches, most of them have 600-1200 bristles. If there are more than 30 bunches, it's usually about 1200-2000 bristles. Our answers are spread out a lot. Maybe we need to check our calculations again.

STUDENT: Or maybe we need more toothbrushes to count.

Graph 2 – Bristles in a Bunch and Bunches in Our Toothbrushes



MS. B: What do you see from the bristles graph (Graph 2)?

STUDENT: I don't know.

STUDENT: The points are all over the place.

Not all graphs are useful.

STUDENT: If you know the bristles in a bunch, I don't think you can predict the number of bunches. I don't know how I would use this graph.

"Real world" connections are encouraged by exploring additional uses of the information.

MS. B: How could the data we have collected be useful to us?

STUDENT: If I wanted to make a deluxe brush with, say, 40 bunches (any more would be too big), I would have about 2000 bristles. If I knew how long each bristle was, I would know how much bristle material I would need for a brush. Then I would know what I would need to buy for my factory to make deluxe brushes. I wonder how the bristles are made?

STUDENT: I now know that there's a reasonable range for bristles. You can't have a brush with too few, like 50, or too many bristles, like 5000.

Data investigations tend to be cyclical: a question leads to data collection and analysis which in turn leads to additional questions.

STUDENT: Next time I buy a toothbrush, I'm going to check to see how many bristles there are. As long as I'm going to spend the time brushing, I might as well be using a lot of bristles.

STUDENT: This data might help us look at some of our other questions, such as whether there are more bristles in the expensive toothbrushes. We kept the cost of our toothbrushes in our notes. Maybe we should make a new table.

Reality Check: Data Investigation in the Workplace

When you want a quick, fool-proof way to bake a birthday cake, you can choose from a wide variety of packaged mixes and canned frostings on the supermarket shelf. Have you ever tried Funfetti frosting?

Toni H. helped to develop Funfetti cake and frosting while she was a food technician in the Research and Development (R&D) department at Pillsbury. She also worked on teams that developed other packaged mixes such as pizza, chewy brownies, refrigerated dough, chocolate caramel bundt cake, and microwavable biscuits. Some of these products are big sellers, while others never made it to the supermarket shelf.

It takes team work to solve problems in the workplace.

Most decisions in the workplace are data-driven.

How do new products get developed? In a large company like Pillsbury, no one person does it all. The R&D team, the Sensory Department, the engineers, the marketing people, and the general public all have a role to play. Data is collected every step of the way. The right questions must be asked, and results must be presented clearly and understood by everyone on the team so decisions can be made that will produce a new food product that makes a profit for the company.

New technologies impact changes in our lives. Sometimes we are aware of them, and sometimes we aren't.

The initial idea for a new product usually begins in the Marketing Department. By looking at sales records, they notice a competing company has a product that is successful, or perhaps their own company has a successful product that might be varied. New technology can also lead to ideas for new developments. When microwave ovens became a popular kitchen appliance, new product variations became necessary. Toni said, "For example, biscuit dough was always a popular item, but the traditional mix was too chewy when it was made in a microwave oven, and wouldn't brown, so we needed to develop a new product." Work in developing a biscuit that could be browned and microwavable led to a patent.

Toni graduated from college with a degree in chemistry and a minor in mathematics. She married after college, had nine children, and became adept at cooking. Her combination of science education and experience in the kitchen made her a natural for a job on Pillsbury's R&D team. Although she was not a math major, she used math in her science courses and learned to understand and interpret data in school and on the job.

Data is collected and summarized in a variety of ways using a variety of tools.

When she worked on brownies, Toni's team came up with 30 different recipes with a variety of toppings. Toni said, "We need to use computers to keep records of the recipes, because each recipe has a formula with a certain percentage of each ingredient." The team tastes everything. As soon as they have a number of formulas that look and taste good, they have other R&D team members do "blind" taste tests. For this kind of testing, measured samples of the food are put on small white plates and are identified only by a number. In the initial testing, the tasters just say "yes" or "no." Results are tallied.

Once the R&D team has narrowed the possibilities, the Sensory Department takes over the taste tests. These tests are also "blind," but involve large groups of people in the company. The Sensory Department creates surveys that the tasters fill out. There could be two pages of questions on texture, color, size, aroma — all the different characteristics that the Marketing Department, the R&D team and the Sensory Department feel are important.

A company taster gets a long page of questions, with categories to mark such as "not so good," "good," "very good," and "excellent." Each category is assigned a numerical value. After the results are tabulated, they are returned to the R&D team in percentages, tables, graphs, and/or charts. This information helps the team members to make decisions about changes in the samples. Finally, choices are narrowed and the surveys are refined for use in testing by the general public.

Next, Marketing decides how many varieties will go out for taste testing at malls or supermarkets across the country. It is important to get a representative sample of the company's potential customers. Sometimes Marketing looks at a cross-section of the population, and at other times product testing is aimed at a specific consumer group, such as teenagers. People from the R&D team go along to make sure the product is delivered the way it should be — right temperature, right appearance.

When there is too much data to sort by hand, company employees program computers to sort and summarize it, but team members must still analyze the results and make decisions based on the data.

Here's a situation where the strategy of multiplying the ingredients doesn't always prove successful.

Once again, tasters choose the best food sample and fill out questionnaires. After the answers to the questions are digitized, the Sensory Department runs the results from across the country through a computer. Team members analyze the results carefully to determine which products will be successful. The data must be presented in a clear and understandable form so the R&D Team can decide whether to make other changes, and the Marketing Team can make the final choice and decide how to market the new product across the country.

Finally, an engineer determines how to process the new product in the plant. A product that is developed in the lab often tastes different when it is mixed in large quantities in the plant, so there must be more taste tests until both the lab and plant samples taste the same.

Once the new product is on the supermarket shelves, the Sales Department gets involved. Do the sales confirm the predictions of the taste tests? Is an advertising campaign recommended? Or is it back to the drawing board?

Toni takes a whole new look at food in her job at Pillsbury. She sees product ideas develop from scratch, uses her technical background to prepare a variety of samples for testing, and works with her team to make decisions based on data. The work is exciting and the taste testing is a delicious side benefit!

What investigation skills are needed to make decisions about developing and marketing a new product?

- Formulate the questions.
- Choose a strategy to collect data.
- Collect data and organize the results.
- Analyze and interpret the results.
- Look back. Did you solve the original problem?

(These Reality Checks are included to remind us that it is our responsibility as teachers to prepare students for the real world — specifically for life-work.)

FOCUS

Students entering the primary grades have a natural curiosity about people and things in their world. This curiosity leads to questions about how things fit together or connect. Primary students display their natural need to organize things. They sort, compare, and label objects in a collection according to similarities and differences with respect to a single attribute.

PRIMARY

The focus of instruction at the primary level is to help students develop methods and systems to collect, organize, describe, and interpret information about themselves and their world in ways that reinforce their natural curiosity.

Students entering the intermediate grades participate with their class to formulate questions of interest and to predict the results based on their personal experience. With teacher support they cooperate to collect useful data and organize this information on graphs and diagrams. They are able to quantify and interpret labels, symbols, and information on a classroom graph. Students use information from data displays to answer descriptive questions ("How many?", "How much?") and to make and justify comparisons ("Which is most popular?", "Who has less?", "Who has the same as you?"). They explore results of using different categories or different methods to display information. They compare their initial predictions to actual results and communicate about the data collection process to others.

INTERMEDIATE

The focus of instruction at the intermediate level is to further develop and investigate systematic data collection strategies and methods of organizing and analyzing data based on questions related to the student, family, school, and community.

MIDDLE SCHOOL

Students entering middle school have had various data collection experiences and have begun to develop habits of thinking focused on wondering and predicting. They use information gained in previous data investigations to solve problems or initiate new investigations. Students recognize that the same data set can be represented in more than one way and they can record, organize, and make various data displays using a variety of tools. They can interpret and analyze simple data displays and can communicate information and results clearly. ("The graphs show that the price of coffee has *gone up faster* than the price of…").

The focus of instruction at the middle school level is to develop an awareness of the power of data analysis to answer questions and to make predictions and decisions. Students explore ways in which data can be used or misused to influence the opinions and decisions of groups of people. They develop more precise skills and terminology to analyze data, critique conclusions based on data, and use data to support additional conclusions.

HIGH SCHOOL

Students entering high school can correctly interpret a graph from media sources. They can recognize the distortions or bias in graphs or data displays and relate them to errors in representation or to the source of the data. They pose questions, collect data, and select the most appropriate display(s) from among tables, scatter plots, bar graphs, line graphs, circle graphs, and stem and leaf plots. They have explored box plots and lines of best fit. Students calculate and interpret mean, median, mode, and range and demonstrate an intuitive sense for spread and deviation. They can make an educated guess to estimate the effect of changes in key portions of the data on measures of center. They have used a variety of technologies to create visual representations of data and can summarize data collection methods in written and oral form.

The focus of instruction at the high school graduation standard level is on expanding and building more in-depth procedures for independently posing questions, collecting and interpreting data, and refining decisions based on analysis of data. These students experience a variety of technologies to assist them in displaying and interpreting data, including spreadsheets, databases, scientific calculators, and computer software. They will deepen their critical understanding of data investigation, design of studies, sampling techniques, and the misuses of data. They will communicate their summaries and informal conclusions to appropriate audiences.

BEYOND HS STANDARD

Students working at the high school graduation standard level can identify problems for study, ask questions, collect and display data, and reach informal conclusions based upon the results. These students have some understanding of simulation, sampling, and distribution.

The focus of instruction beyond high school graduation standard level is on designing and conducting statistical experiments to study relevant problems. These students will expand their data analysis skills by applying concepts of variability, reliability, validity, and correlation. They will incorporate probability to make inferences with a specified level of confidence.

Components

LEARNING GOAL

Students will ask questions and collect, organize, interpret, and transform data related to those questions to communicate information and make predictions and decisions.

Components

With this learning goal in mind, Minnesota students will have the opportunity to:

1. formulate a question(s), determine necessary data, and choose an appropriate method of data collection in order to make sense of a problem or situation
2. collect, organize, and represent data in a variety of ways
3. read, describe, and interpret displays of data
4. formulate, justify, and communicate conclusions, arguments, predictions, decisions, and further investigations based on data

1

Students formulate a question(s), determine necessary data, and choose an appropriate method of data collection in order to make sense of a problem or situation.

Learning Goal: Students ask questions and collect, organize, interpret, and transform data related to those questions to communicate information and make predictions and decisions.

PRIMARY K-2

All indicators below assume teacher guidance:

- identify questions based on student's world and natural curiosity
- suggest what data may help to answer the question
- share opinions about expected results
- decide what information is needed, what questions need to be asked, and who to ask

INTERMEDIATE 3-5

- identify a problem of interest from class, family, school or community
- generate and refine possible questions to explore
- create a plan to investigate the question(s)
- use prior experiences to predict the results of the investigation
- collect both categorical (gender, shoe color) and numerical (height, allowance) data

Excerpts from MN Standards Related to Component #1

MN State Standards

Primary Level: K-2

Inquiry:
Data Categorization
- gather information from: media sources, direct observation, interviews, and experiment or investigation to answer a question

Intermediate Level: 3-5

Mathematics:
Chance & Data Handling
- answer questions: collect and organize data

Inquiry: Media Observation & Investigation
- gather information from direct observations or experiments with a variable…frame a question
- gather information from media sources…select a topic and frame a question
- gather information through direct observation and interviews: identify a topic or area for investigation…conduct an interview with follow-up questions or design and conduct a survey

MIDDLE SCHOOL 6-8

- identify a problem of interest that moves beyond the individual or school to the larger community
- work in small groups to clarify questions to focus the investigation and the data collection plan
- refine language to clearly communicate intent of survey questions
- critique questions generated by the class or posed in the media
- design an investigation plan, including appropriate technology, to collect data
- choose an appropriate sample
- identify target audience and create a data investigation plan to support an opinion or argument (e.g., "more people carpool to work")

Mathematics: Chance & Data Handling

- formulate a question and design an appropriate data investigation

Science: Living Systems

- formulate questions to be answered based on systematic observations; design and conduct investigations and field studies

Inquiry: Direct Observation

- gather information from direct observations: frame a question

Inquiry: Accessing Information

- generate a question to be answered or a position to be supported through investigation

Inquiry: Controlled Experiments

- use relevant information to generate a hypothesis or frame a question in a given topic

Middle Level: 6-8

HIGH SCHOOL 9-12

- pose meaningful questions using data from external sources (census, media, graphs, charts, surveys, experiments)
- design investigations to model and solve problems
- apply appropriate sampling techniques in data collection strategies
- use simulations to model the process of data collection
- identify dependent and independent variables

BEYOND HS STANDARD

- study a variety of data analysis design types (e.g., surveys, experiments, polls)
- design a survey study and/or an experimental study
- identify possible sources of bias and error in questions to be asked, data to be collected, and/or measurements to be made
- describe strategies to reduce sources of bias
- confront issues related to validity and reliability of the design to improve accuracy
- formalize an hypothesis statement and design a plan to test it

Mathematics: Chance & Data Analysis

- Investigate a problem of significance: formulate a complex question(s); design a statistical study; collect data

Inquiry: Social Science Processes

- formulate a question about a historical event, issue or interpretation of a concept
- create a plan for collecting and interpreting data
- evaluate research problem for feasibility *(e.g., availability and access to data)*
- gather information through the primary research techniques of observations, interviews, surveys, or experiments

Resource Management: Natural/Managed Systems

- develop and communicate a resource management plan involving natural and managed systems…gather data using appropriate techniques; identify the nature of the interactive and interdependent relationships

Minnesota High School Graduation Standard

2

Students collect, organize, and represent data in a variety of ways.

PRIMARY K-2

- sort objects using at least two attributes
- with teacher guidance, make a plan to collect data
- explore data collection over time (e.g., a class weather graph is displayed on the wall which children add to over a period of time)
- represent data using real objects
- use pictures to represent data
- organize and display data using simple tables, graphs, and diagrams such as tallies, lists, Venn diagrams, bar graphs, picture graphs, line plots

INTERMEDIATE 3-5

- collect data systematically from surveys, experiments, simulations, or existing sources
- modify plan to collect data when necessary
- explore various methods to display data, including tables, charts, maps, timelines, Venn diagrams, picture graphs, bar graphs, and line plots
- choose an appropriate method(s) to display a specific set of data
- if necessary, summarize symbolic information from a graph in a key

Learning Goal: Students ask questions and collect, organize, interpret, and transform data related to those questions to communicate information and make predictions and decisions.

Excerpts from MN Standards Related to Component #2

MN State Standards

Primary Level: K-2

Inquiry: Data Categorization
- record information *(e.g., graphs, diagrams, maps)*
- display information using the appropriate format *(e.g., graphs, diagrams, maps)*

Mathematics: Number Sense
- solve problems and justify thinking…organize data using pictures and charts

Science:
Direct Science Experience
- display information using graphs

Intermediate Level: 3-5

Mathematics:
Chance & Data Handling
- answer questions: collect and organize data, represent data *(e.g., graphs, charts)*
- represent data using at least two graphic forms *(e.g., graphs, tables, charts and pictures)*

Inquiry: Media Observation & Investigation
- gather information from direct observation or experiments with a variable…collect, record and display data
- gather information from media sources; access information from electronic media, print, interviews and other sources

MIDDLE SCHOOL 6-8

- access data from reference sources as basis for some investigations
- explore various methods of displaying data, including cumulative frequency distributions, scatter plots, stem and leaf plots, line graphs, circle graphs, and spreadsheets
- select an appropriate method to display data given the type of data collected and the kinds of comparisons to be made
- use technical tools (graphing calculators and software) to construct graphical displays appropriate to the data
- determine and represent appropriate intervals on graphic displays
- explore sorting data in a database
- use telecommunications to obtain data from populations outside one's community

HIGH SCHOOL 9-12

- use technology to enhance data displays
- recognize when and how to use different graphical displays, including frequency distributions, histograms, box plots, curve fitting
- apply different scales and representations to reveal different characteristics of the data
- select appropriate statistical techniques based on the type of data

BEYOND HS STANDARD

- select appropriate data and effective methods of displaying data for a given purpose and/or audience
- apply data transformations and various graphical representations to reveal different characteristics of the data (e.g., achieve linearity, center around zero, convenient scaling)

Middle Level: 6-8

Mathematics: Chance & Data Handling
- organize raw data and represent it in more than one way

Resource Management: Technology Applications
- use spreadsheets, databases…

Inquiry: Accessing Information
- use electronic media or other available means to access relevant information…

Inquiry: Controlled Experiments
- determine how to record and organize data
- conduct an experiment and record data

Minnesota High School Graduation Standard

Mathematics: Chance & Data Analysis
- represent data appropriately *(e.g., tables, graphs, charts, plots, frequency distributions, data bases)*

Inquiry: Social Science Processes
- gather information through observation, interviews, surveys or experiments

Inquiry: Research Process
- create a plan for collecting and interpreting data
- collect and interpret primary data

3

Students read, describe and interpret displays of data.

Learning Goal: Students ask questions and collect, organize, interpret, and transform data related to those questions to communicate information and make predictions and decisions.

PRIMARY K-2

All indicators below assume teacher guidance:

- describe and compare categories in data displays
- use informal or familiar language to describe the shape of data displays
- identify and describe patterns in data displays
- informally describe mode and range
- identify and use a simple key to interpret information on a data display

(e.g., =10 cats)

INTERMEDIATE 3-5

- describe what information is (and is not) shown in a graph
- use informal language and some standard descriptions to communicate an intuitive feel for what is typical of the data and how the data behaves
- use informal language to describe the shape of the data (e.g., identify bumps, holes, clumps, symmetry, clusters, outliers)
- use terms median, mode and range to describe data
- identify trends and patterns in data collected over time
- explore concrete methods for determining mean of a set of data
- compare and contrast data from more than one graph

Excerpts from MN Standards Related to Component #3

MN State Standards

Primary Level: K-2

Writing & Speaking
- write a report to describe and give information about a person, an object or a situation

Intermediate Level: 3-5

Mathematics: Chance & Data Handling
- understand how to find range, mean and median
- understand information displayed in graphs, tables and charts
- describe patterns, trends, or relationships in data displayed in graphs, tables or charts

MIDDLE SCHOOL 6-8

- describe factual information shown on a graph
- calculate measures of center (mean, median, mode) from a display of numerical data
- determine and explain which measure of center best describes what is typical about the data
- explore measures of variability (range, quartiles, percents, percentiles, outliers) and discuss their effect on the shape of the data
- use scatterplots to explore the simple correlation between two sets of data
- use scatterplots and coordinate graphs of bivariate data to begin exploring the concept of dependent and independent variables
- identify trends in data collected over time and differences across various populations

HIGH SCHOOL 9-12

- download and investigate data sets from the Internet
- recognize and explain inappropriate uses of data
- use formal terminology to describe the shape of data distributions, such as symmetrical, skewed, normal, and bimodal
- use appropriate technology to describe single variable data using numerical measures for center and spread
- calculate and interpret quartiles and percentiles
- use technology to construct lines/curves of best fit

BEYOND HS STANDARD

- download and analyze data sets from the Internet
- interpret and apply common measures of dispersion (e.g. variance, standard deviation)
- interpret and apply the correlation coefficient
- use technology to calculate correlation coefficients for bivariate data (linear, quadratic, exponential, logarithmic, power)
- understand the role of least squares in curve fitting techniques
- apply curve fitting techniques to determine if data is modeled by a linear function or by a type of curve (for instance, logarithmic, exponential, periodic)
- explore data which sorts into categories by using two-way tables, chi-square analyses and/or median smoothing

Mathematics:
Chance & Data Handling:
- calculate basic measures of center and variability *(e.g., mean, median, mode, range, quartiles)*
- analyze data by selecting and applying appropriate data measurement concepts

Read, View, Listen:
Technical Reading
- understand information from visual or graphic data

Mathematics:
Patterns and Functions
- recognize, analyze and generalize patterns found in…data from lists, graphs and tables
- represent and interpret cause and effect relationships using…tables and graphs

Middle Level: 6-8

Mathematics: Chance & Data Analysis
- use appropriate statistics to summarize data

Resource Management: Financial Systems
- create a report based on information obtained from data analysis: describe and display data; analyze effectiveness of past financial actions; recommend future courses of action based on conclusions of data analysis

Mathematics: Technical Applications
- read and interpret information in complex graphs, tables and charts

Minnesota High School Graduation Standard

4

Students formulate, justify, and communicate conclusions, arguments, predictions, decisions, and further investigations based on data.

Learning Goal: Students ask questions and collect, organize, interpret, and transform data related to those questions to communicate information and make predictions and decisions.

PRIMARY K-2

- identify unexpected information in a set of data and suggest possible explanations
- recognize that different interpretations of the data may exist
- summarize data information verbally and in writing
- draw conclusions based on data and explain thinking
- generalize beyond recorded information to answer questions like, "What do you think would happen if you asked this question in Mr. Marshall's classroom?"; "What do you think would happen if we repeated this experiment?"
- reflect on the data collection process: Did the data answer the question? What new questions do we have? Who could use this information?

INTERMEDIATE 3-5

- question data points that don't seem to make sense and suggest explanations
- compare the original predictions to the results of the investigation
- discuss how interpretations of the data might change an existing attitude or opinion
- summarize information, decide how to communicate this information and what, if any, action should be taken
- share information and report conclusions to appropriate audience(s)
- recognize that individual samples may vary
- recognize that a sample of the same size could yield different results

Excerpts from MN Standards Related to Component #4

MN State Standards

Primary Level: K-2

Inquiry: Data Categorization
- explain the answer to the question

Read, Listen, View: Literal Comprehension
- interpret presentations of data *(e.g., charts, tables, graphs)*

Writing & Speaking
- give an informal oral presentation; present an opinion or idea; use reasons or examples to explain it; respond to related questions

Intermediate Level: 3-5

Mathematics: Chance & Data Handling:
- answer questions… communicate results

MIDDLE SCHOOL 6-8

- discuss which possible conclusions are supported by the data
- select the information and determine an appropriate presentation method for a given purpose or a given audience
- discuss the impact the data might have on decision making depending on the slant of the interpretation or the population(s) surveyed
- discuss the degree of confidence students have in their results
- recognize the cyclical nature of investigations (i.e., results tend to lead to new questions)
- use appropriate terminology to communicate an analysis of the data orally and in writing
- take appropriate action based upon the results of the data analysis
- explore possible sources of bias in data investigation plans
- summarize and critique data investigations done by others

Mathematics: Chance & Data Handling:
- critique various representations of data
- predict future results based on experimental results by identifying patterns in data

Science: Physical Systems
- analyze data to support or refute hypothesis

Inquiry: Accessing Information
- evaluate the relevance of the information

Middle Level: 6-8

HIGH SCHOOL 9-12

- review results to determine if the original question has been answered satisfactorily
- identify outliers and justify their inclusion or exclusion in statistical measures based on the possible sources of error
- explore concepts of validity and reliability
- understand that correlation does not imply causation
- identify causes of bias in a statistical design and investigate strategies to remedy bias
- determine whether or not additional data and/or analysis is necessary
- make and test hypotheses inferred by the results of an investigation
- create sampling distributions
- compare variability among sets of different sized samples
- state conclusions about the original question based on correlation coefficients
- use visual, oral and written forms of communication appropriate to the audience to clearly express results
- make practical decisions based on analysis of data

BEYOND HS STANDARD

- evaluate and critique results based upon original hypotheses
- apply confidence intervals to test hypotheses
- reevaluate the original design and the processes used to conduct the study
- reevaluate hypotheses based upon analysis of data
- suggest new hypotheses based upon previous investigations
- distinguish between correlation and causation

Mathematics: Chance & Data Analysis
- understand the following statistical concepts: measures of center, variability and rank; differences between correlation and causation; sampling procedures; line or curve of best fit
- Investigate a problem of significance: formulate a complex question(s); design a statistical study; collect data; represent data appropriately; use appropriate statistics to summarize data; determine whether additional data and analysis are necessary; draw conclusions based on data; communicate the results appropriately for intended audience
- analyze and evaluate the statistical design, survey procedures, and reasonableness of conclusions in a published study or article

Minnesota High School Graduation Standard

Sample Activities

Sample Activity
Primary

Sorting People: Who Fits My Rule

Adapted with permission from: Russell, S. & Corwin, R. (1990). *Used numbers: Sorting: groups and graphs.* Palo Alto, CA: Dale Seymour Publications, pp. 13-19.

"Guess My Rule" is a classification guessing game in which players try to figure out the common characteristic, or attribute, of a set of objects.

The game can be played in many different contexts, but the students first play it in a context very familiar to them: they sort themselves. To play the game, the rule maker (who may be you, a student, or a group of students) decides on a secret "mystery rule" for classifying a particular group of people. For example, classification rules for people might be everyone who is wearing orange, or everyone who is missing a front tooth.

The rule-maker starts the game by giving some examples of people who fit the rule (for example, by having two students who are wearing orange stand up). The guessers then try to find other individuals who might fit the rule: "Can Sandy fit your rule?"

With each guess, the individual named is added either to the group that does fit or to the group that does not fit the rule. Both groups must be clearly visible to the guessers so they can make use of all the evidence, both what does and does not fit, as they try to figure out what the rule might be.

Two guidelines are particularly important to stress during play:

- "Wrong" guesses are clues that are just as important as "right" guesses. Errors are not just mistakes but can be important sources of information.
- When students think they know what the rule is, they "test" their theory by giving another example, not by revealing the rule. Requiring students to add new evidence, rather than making a guess, serves two purposes. It allows students to test their theories without revealing their guess to other students. And it provides more information and more time to think for students who do not yet have a theory of their own.

When students begin choosing rules themselves, they sometimes think of rules that are either too vague (students wearing different colors) or too difficult for other students to guess (students with a piece of thread hanging from their shirts). You can guide and support students in choosing rules that are not so obvious that everyone will see them immediately, but not so hard that no one will be able to figure them out. The students should be clear about who would fit their rule and who would not fit. This eliminates rules like wearing different colors which everyone will probably fit. It's also important to pick a rule about something people can see. That is, one rule for classifying might be "likes baseball," but no one will be able to figure out this rule just by looking.

Since classification is a process used in many disciplines, you can easily adapt the game to other subject areas. Animals, states, historical figures, geometric shapes, types of food, and countless other items can all be classified in different ways.

Being able to see how things go together improves with experience. As you and your students play this game repeatedly in different contexts, you will find yourselves becoming more observant and more flexible in your thinking about similarities and differences.

This activity reinforces many of the key ideas in the K-12 Components at the primary level:

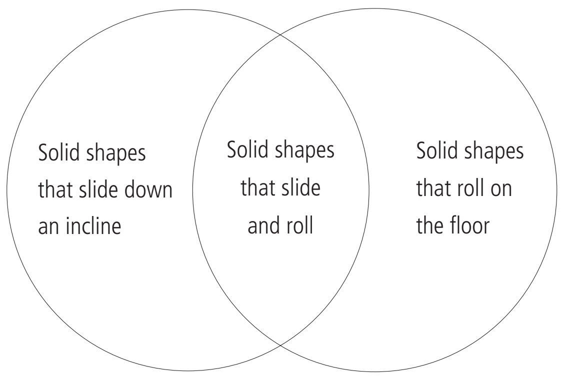
- represent data using real objects
- with teacher guidance, describe and compare categories in data displays
- summarize data information verbally and in writing
- draw conclusions based on data and explain thinking
- recognize that different interpretations of the data may exist

About this Activity

- This unit focuses on collecting, organizing and interpreting categorical data, and emphasizes student work in classification and data analysis.
- This particular activity encourages children to redefine their thinking in response to new data.
- Children feel free to do their best thinking in a safe environment where their ideas are valued.
- The "rules" in the game go a long way to maintaining a thinking curriculum, rather than one based on the "right" answer.

Where do we go from here?

"Guess My Rule" naturally extends to the sorting of concrete objects, and later to the graphical representation of data on charts, graphs, and Venn diagrams. Students first sort on one characteristic, then proceed to sort on two or more characteristics.



Recommended Teaching Resources

Baratta-Lorton, M. (1979). *Mathematics their way.* Menlo Park, CA: Addison-Wesley Publishing Co.

Lindquist, M.M. with Luquire, J., Gardner, A., & Shekaramiz, S. (1993). *Making sense of data: Curriculum and evaluation standards for school mathematics addenda series, grades K-6.* Reston, VA: NCTM.

Russell, S. & Corwin, R. (1990). *Used numbers: Counting: Ourselves and our families.* Palo Alto, CA: Dale Seymour Publications.

Russell, S. & Corwin, R. (1990). *Used numbers: Sorting: Groups and graphs.* Palo Alto, CA: Dale Seymour Publications.

Stenmark, J., Thompson, V., & Cossey, R. (1986). *Family math.* Berkeley, CA: Lawrence Hall of Science.

Sample Activity
Intermediate

Mystery Graphs

Adapted with permission from: Mathematical Sciences Education Board. (1993). *Measuring up: Prototypes for mathematics assessment.* Washington, D.C.: National Academy Press, p. 23-29.

Look at the five graphs below. Each graph shows something about a classroom of fourth graders.

1. Choose the graph that best fits each description. Give the reasons for your choice in each case:

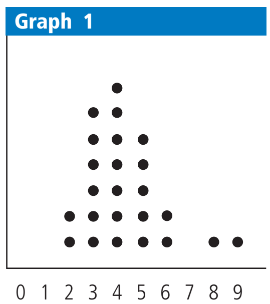
a. The number of cavities that the fourth graders have______

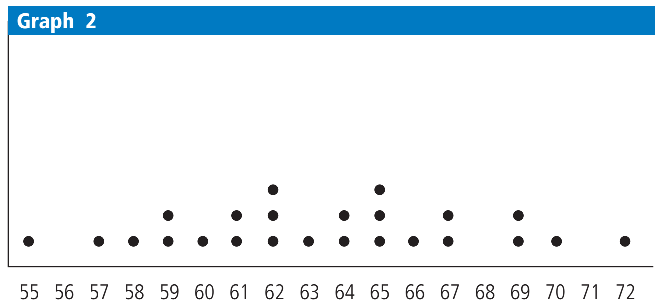
b. The ages of the fourth graders' mothers______

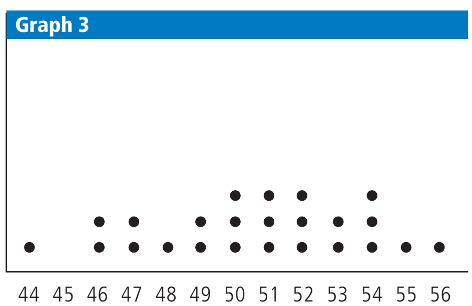
c. The heights of the fourth graders, in inches______

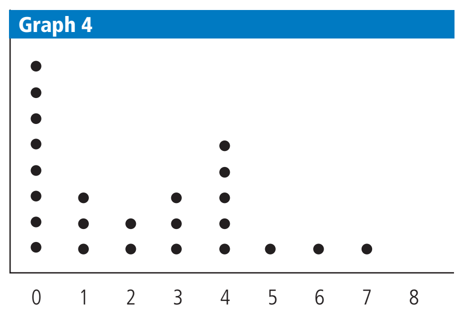
d. The number of people in the fourth graders' families_______

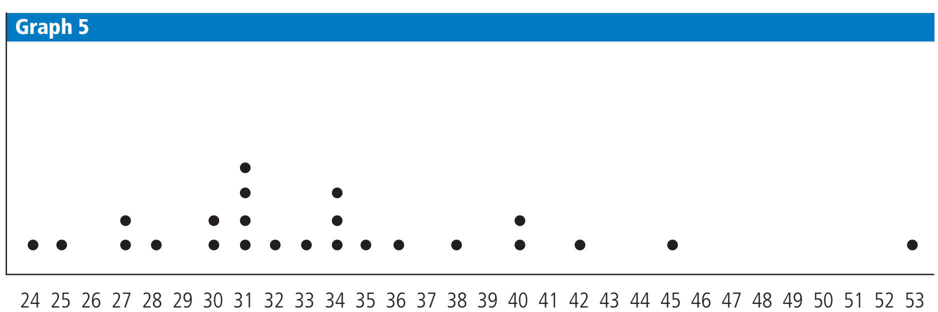
2. Look at the graph you did not select for any of the statements. Make a list of possible things about a classroom of fourth graders this graph might show.











This activity reinforces many of the key ideas in the K-12 Components at the intermediate level:

- use prior experiences to predict the results of the investigation
- describe what information is (and is not) shown in a graph
- use informal language and some standard descriptions to communicate an intuitive feel for what is typical of the data and how the data behaves
- compare and contrast data from more than one graph

About this Activity

- This task involves problem solving, interpreting and comparing data, making inferences, and justifying conclusions.
- This task assumes prior instruction and extensive experience in collecting, organizing and displaying data. It also assumes that children have had experience with sets of data represented in line plots.
- Post a copy of the unselected graph on a large sheet of paper on the wall. Have each group or individual share one possible thing they thought of from part 2. List these on the paper. Go around until all ideas are listed. Leave the list up for students to add to if they think of other things which could have been represented. Come back to it another day as a whole group if items are added.
- Statisticians often use line plots for preliminary work with data distributions. A line plot is a sketch showing the values of the data along a horizontal axis and marks to represent the data points. Line plots ease the job of recording data and allow students to readily identify the "shape of the data" as bumps, clumps, or holes.

Where do we go from here?

1. Extension questions include:

"Suppose Graph 2 really did show heights in inches. Whose heights could they be?"

"Suppose Graph 3 showed the ages of the mothers of students in some grade level in our school. Which grade could that be?"

"What other kinds of data could Graph 1 be showing?"

2. Cut graphs out of the newspaper and remove the title of each graph and the labels on the axes. Encourage the class to create as many appropriate labels/titles as they can for each graph, giving reasons for their choices.

Recommended Teaching Resources

Ann Arbor Public Schools. (1993). *Alternative assessment*. Palo Alto, CA: Dale Seymour Publications.

Burns, M. (1987). *A collection of math lessons from grades 3-6*. New York, NY: Cuisenaire Company of America, Inc.

Lindquist, M.M. with Luquire, J., Gardner, A., & Shekaramiz, S. (1993). *Making sense of data: Curriculum and evaluation standards for school mathematics addenda series, grades K-6*. Reston, VA: NCTM.

Mathematical Science Education Board (MSEB). (1993). *Measuring up: Prototypes for mathematics assessment.* Washington, D.C.: National Academy Press.

Russell, S. & Corwin, R. (1989). *Used numbers: Statistics, the shape of the data*. Palo Alto, CA: Dale Seymour Publications.

Sample Activity
Middle School

Average Nth Grader

Adapted with permission from: Zawojewski, J.S. *Dealing with data and chance: Curriculum and evaluation standards for school mathematics addenda series, grades 5-8*, ©1991 by the National Council of Teachers of Mathematics, pp. 4-10.

What would you like to know about your classmates?

- What does the average Nth grader look like?
- What are his or her favorite sports? music? books? foods?
- How does the average Nth grader spend her or his time?

What does "average" mean?

What data does your group need to gather in order to answer these questions?

Before beginning, certain decisions about conducting the data investigation have to be made by your group, including:

- What is being investigated and why?
- What data need to be collected?
- How will the data be collected and from whom?
- How will the data be organized and analyzed?
- What comparisons will be made?
- Who will find this data useful or interesting?

After you have examined the data, identify patterns. Use the evidence to formulate conjectures and theories. Also view the data from different perspectives in order to address other questions. When you have finished, prepare a detailed report of your group's investigation, including graphs and charts, as well as conjectures and ideas on how the investigation might be extended.

Assessment—portfolio:

Prepare a report describing your group's creation of a database, your use of a spreadsheet, and your analysis of selected data collected in your survey with accompanying charts, graphs, and other displays. Indicate how your project could be extended by asking some "What if…" questions, such as "What if we compared selected parts of these data?" (For example, compare class A with class B, compare homework time with TV time. How would the results differ?)

Another look at the data:

Separate the boys' and girls' data. Look at the male and female mean, male and female mode, and male and female median. Are the female mean, mode, and median identical? Are the male mean, mode, and median identical? Why or why not?

Extend the investigation:

Is our classroom representative of our school? city? state? Is it the same the world over? How could we explore these questions?

Have students find records of the size of people today and in bygone days. Could today's women wear the fashions we see in a museum? Could today's people fit comfortably in the Mayflower? (For example, could they stand up below decks without hitting their heads?) A visit to a historical museum may take on a whole new dimension and lead to some new explorations.

This activity reinforces many of the key ideas in the K-12 Components at the middle school level:

- work in small groups to clarify questions to focus the investigation and the data collection plan
- refine language to clearly communicate intent of survey questions
- design an investigation plan, including appropriate technology, to collect data
- choose an appropriate sample
- identify trends in data collected over time and differences across various populations
- discuss which possible conclusions are supported by the data
- use appropriate terminology to communicate data analysis orally and in writing

About this Activity

- Students should have some knowledge of basic statistical concepts (i.e. mean, median, mode, range). Some class discussion of the meaning of "average" is useful.
- As students work in groups to develop their survey, the teacher may need to help them clarify the questions to be used in order to obtain useful data.
- Students should be introduced to a variety of ways of collecting data and/or representing data.
- This activity may provide an opportunity to show students how to enter data into a database. For example, if each student enters her or his data from this activity into the database, the class could find all the students who like the same favorite food or the amount of time spent watching TV or doing homework each day.
- Students can also prepare spreadsheets on different aspects of the data they wish to analyze.
- Encourage students to look for relationships between the categories and the data. Using their data, can students find an "average" student in their class?
- Students can create a make-believe student having all the "average" characteristics of the class as a whole. This can be the material for a large display or bulletin board, illustrated by graphs and charts highlighting the characteristics of an "average Nth grader."
- If the survey is done at the beginning of the school year and again at the close of the school year, students can make analyses and comparisons to show changes that have occurred individually and collectively in the course of one (short) school year.

Where do we go from here?

Help students formulate new questions and determine further investigations. This might lead to another survey project. Questions such as "What if we compared homework time with TV time?" might lead to new relationships to express or new surveys to make.

Data collected should be referred to at other times during the year. For example, when students are studying functions and relationships, they might use graphing calculators or graphing software to organize and display the data they collected on shoe size or height. They might find the line of best fit.

Hold the data on file to compare year by year. Ask questions such as: "Are we getting taller or heavier over time?"

Recommended Teaching Resources

Landewehr, J.M., Swift, J., & Watkins, A.E. (1987). *Exploring surveys and information from samples: Quantitative literacy series.* Palo Alto, CA: Dale Seymour Publications.

Lappan, G., Fitzgerald, W., Friel, S., Fey, J., & Phillips, E. (1995). *Data around us: Connected mathematics project.* Palo Alto, CA: Dale Seymour Publications.

Zawojewski, J., Brooks, G., Dinkelkamp, L., Goldberg, E., Goldberg, H., Hyde, A., Jackson, T., Landau, M., Martin, H., Nowakowski, J., Paull, S., Shulte, A., Wagreich, P., & Wilmot, B. (1991). *Dealing with data and chance: Curriculum and evaluation standards for school mathematics addenda series, grades 5-8,* Reston, VA: NCTM.

Sample Activity
High School

Heart Disease Investigation

Adapted with permission from: Lovitt, C. (1991). *Maths problem solving and modeling for year 11.* South Melbourne, Australia: Thomas Nelson.

Heart disease, which causes enormous suffering and expense in our community, is a common cause of death in our society. While we don't know the cause of heart disease, we do know many risk factors (i.e., factors which by statistical association seem to be causative factors).

Each of the following factors is known to be connected with heart disease:

- cholesterol and other blood fat levels
- cigarette smoking
- high blood pressure (hypertension)
- lack of physical activity
- psychological stress
- obesity (excess weight)

Statisticians and health theorists are currently working to build a model of risk factors. These activities challenge you to explore the data surrounding the known risk factors of heart disease and to build statistical models of the risks involved.

As a small group, select and complete one or two of the following tasks or design a task of your own with your instructor's approval. Be prepared to present and defend your findings to the class.

1. Design a survey which would allow students in your school to rank the 6 identified factors which are associated with heart disease. Conduct the same survey with a group of adults and compare the results. Analyze and communicate your findings.

2. Using the data provided from an organization such as the American Heart Association, look for interesting patterns. Infer and communicate explanations for these patterns.

Draw a graph of blood pressure vs. age. Plot male and female data on the same axes. What patterns do you see? Pose and explain possible explanations for these patterns. Does it appear that heart attacks and high blood pressure are related? Create a convincing argument to support your answer.

3. Compare data from other countries. Pose a question related to this data. Create a simulation which has more accurate probabilities and run it on the U.S. population to answer this question.

4. What changes would you anticipate if the data represented different groups within American society? For what groups would you be interested in determining the probability of heart disease? Find actual information to evaluate and communicate what the data shows.

This activity reinforces many of the key ideas from the K-12 Components at the High School Graduation Standard Level:

- pose meaningful questions using data from external sources (census, media, graphs, charts, surveys, experiments)
- design investigations to model and solve problems
- apply appropriate sampling techniques in data collection strategies
- review results to determine if original question has been answered satisfactorily
- determine whether or not additional data and analysis is necessary

About this Activity

- This data topic deals with a significant issue in society. Students may choose from a variety of activities rather than all being assigned the same task.
- The choices emphasize active involvement on the part of students as they interact with data.
- The tasks encourage cooperative learning and the use of a variety of resources.
- These choices have multiple solutions and encourage multiple representations and technologies in presenting results.
- Tasks 3 and 4 may encourage students to look at cultural differences from a unique perspective.

Where do we go from here?

Work with students to identify other health issues that can be researched that would be rich sources for investigating and working with data.

What social action could students take based on their analysis of this or other data?

Recommended Teaching Resources

Landewehr, J. & Watkins, A. (1987). *Exploring surveys and information from samples: Quantitative literacy series.* Palo Alto, CA: Dale Seymour Publications.

Lovitt, C. (1991). *Maths problem solving and modeling for year 11.* South Melbourne, Australia: Thomas Nelson.

National Council of Teachers of Mathematics. (1990). Focus issue: Data analysis. *Mathematics Teacher,* 83(2).

National Council of Teachers of Mathematics. (1991). *Applications of secondary school mathematics.* Reston, VA: NCTM.

Schaeffer, R., Watkins, A., Gnanadesikan, M., & Witmer, J. (1996). *Activity based statistics.* New York, NY: Springer-Verlag.

[See NOTES at the end of the "References" in this content section for information on ASA-NCTM newsletters and Internet sources dealing with data topics.]

Sample Activity Beyond High School Standard

What is a Confidence Interval Anyway?

Adapted with permission from: Schaeffer, R., Watkins, A., Gnanadesikan, M., & Witmer, J. (1996). *Activity based statistics.* New York, NY: Springer-Verlag.

Forty people are selected at random and given a test to identify their dominant eye. The person holds a piece of paper about 8-1/2 in. by 11 in. with a 1 in. by 1 in. square cut in the middle at arm's length with both hands. The person looks through the square at a relatively small object across the room. The person then closes one eye. If he or she can still see the object, the open eye is the dominant eye. If not, the closed eye is the dominant eye.

Question: Is this sample of only 40 people large enough that we can come to any conclusion about what percentage of people have a dominant right eye?

Activity 1. Conduct the experiment described above with 40 students. What proportion of your sample was right-eye dominant?

Activity 2. Assume 30% of the population tends to be right eye dominant. This activity is designed to create a confidence interval for samples of size 40 in which we know the success rate is 30%. We can then compare our classroom sample to this confidence interval to see whether the sample would support the claim that 30% of the people are right eye dominant.

a. Use a random digit table to simulate taking a sample of size 40 from a population with 30% "successes." Let the digits 0,1, 2 represent a success and 3, 4, 5, 6, 7, 8, 9 represent a failure. Tally your results (the number of "successes" out of 40 random numbers).

b. Combine your results in a table like that below with other members of your class, repeating the simulations until tally marks from 100 different samples of size 40 have been placed in the frequency column.

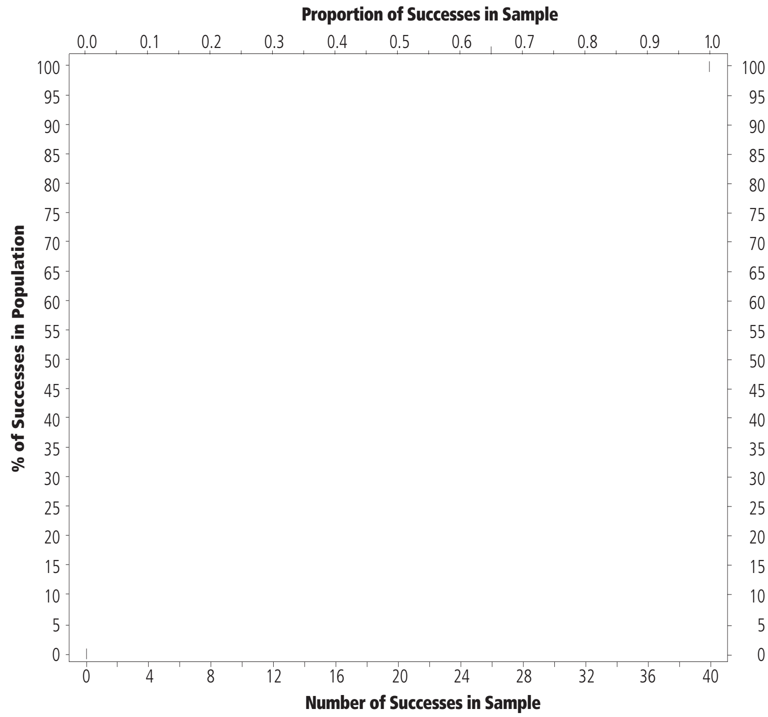
# Successes	Sample Proportion (successes/40)	Frequency Tally	Proportion All Trials
0	0.000		
1	0.025		
2	0.050		
3	0.075		
•	•		
•	•		
•	•		
40	1.000		
		Total 100	1.00

c. Comparing your proportion from Activity 1 to the frequency table above, is it plausible (has a reasonable chance of being the case) that 30% of the population is right-eye dominant? Explain.

d. Complete the following sentences based on your frequency table from part b: Less than 5% of the time, there were __ successes or fewer. Less than 5% of the time there were __ successes or more.

e. On the following chart "90% Box Plots from Samples of Size 40" draw a thin horizontal box aligned with the 30% "Percentage of Successes in Population" on each side. Using "Number of Successes in Sample" on the bottom as a guide, the box should stretch in length between your two answers to part d.

Figure 1 – 90% Box Plots from Samples of Size 40

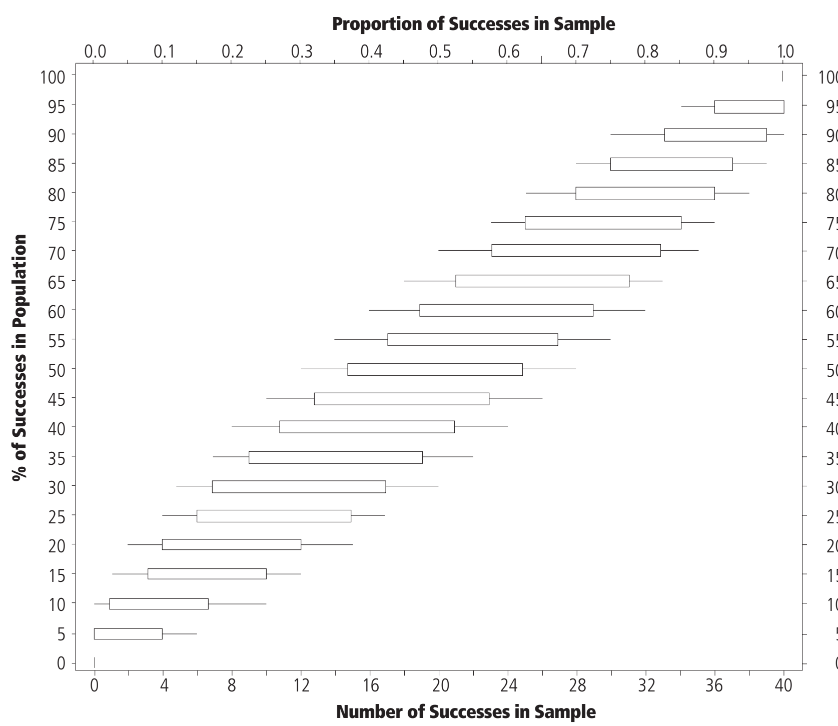


f. If the process for determining the "number of successes" were to be repeated for other percentages, a chart similar to the following would be constructed.

A box in the chart gives the sample proportions that are plausible for that particular population. Other sample proportions are possible, but together they occur less than 10% of the time.

Used with permission from: Landewehr, J. & Watkins, A. (1987*). Quantitative literacy series: Exploring surveys and information from samples.* Palo Alto, CA: Dale Seymour Publications.

Figure 2 – 90% Box Plots from Samples of Size 40



Sample Activity Beyond High School Standard

What is a Confidence Interval Anyway?

(continued)

Activity 3. Use the completed chart in Figure 2 to answer these questions:

a. According to the 1990 U.S. Census, about 30% of people aged 25 to 44 live alone. In a random sample of 40 people aged 25 to 44, would it be plausible to get 20 who lived alone?

b. We select a random sample of size 40 and get a sample proportion of 0.90 successes. Is this sample proportion plausible if the population has 75% successes?

Activity 4. Is the proportion from Activity 1 of students who have a dominant right eye a plausible sample proportion from a population with 60% successes? For which populations is it plausible?

The proportions represented by each box indicate the 90% confidence interval for that percentage.

Observe 40 students at your school. Use your completed chart to find the 90% confidence interval for the percentage of students who carry backpacks.

Polls usually report a "margin of error." Suppose a poll of 40 randomly-selected math majors shows that 20 are female. The poll reports that 50% of math majors are female, with a margin of error of 10%. Use your chart to explain where the 10% came from.

People complain election polls cannot be right because they personally were not asked. Write an explanation about how polls give a good idea of the entire population by asking a relatively small number of voters.

This activity reinforces many of the key ideas in the K-12 Components at the High School Beyond Graduation Standard Level:

- study a variety of data analysis design types (e.g., surveys, experiments, polls)
- identify possible sorts of bias and error in questions to be asked and/or data to be collected/measurements to be made
- evaluate and critique results based upon original hypotheses
- apply confidence intervals to test hypotheses

About this Activity

- In this activity students build a model that can be used to visualize confidence intervals. From this model, students can understand how a sample can give a good indication of the results that would occur had the whole population been observed.
- Students should know how to simulate samples from a given binomial population by using random digits.
- Each student should have a random digit table or access to a calculator or computer program that can draw random samples from a binomial population.
- For Activity 1 you should have exactly 40 people so students may need to bring in results for others outside their class. Proportions will vary.
- Activity 2: Make sure your students use different portions of the random digit table. Check to be sure they see how each column of the chart is calculated. (Sample Proportion is # Successes / 40. Proportion All Trials is Frequency / 100).
- Activity 2e. Discuss how the box plots relate to the table they have just made.
- Remind students that the validity of the confidence interval procedure is based on the assumption of a random sample. People who enroll in a statistics class may for some reason not be representative of the population as far as eye dominance is concerned.

Where do we go from here?

This activity used 90% confidence intervals because it is easy computationally to find the bottom 5% and the top 5% of a distribution. Usually, 95% confidence intervals are reported. Will 95% confidence intervals be longer or shorter than 90% confidence intervals? Explain your answer.

Will the confidence intervals for samples of size 80 be longer or shorter than those for samples of size 40? Design and carry out the simulation needed to answer this question.

Recommended Teaching Resources

Coxford, A., Fey, J., Hirsch, C., Schoen, H., Burrill, G., Hart, E., & Watkins, A., with Messenger, M.J. & Ritsema, B. (1997). *Contemporary mathematics in context: A unified approach, course 1: Core-plus mathematics project.* Chicago, IL: Everyday Learning Corporation.

Interactive Mathematics Project. (1997). "The pollster's dilemma" *Interactive mathematics program, year 4.* Berkeley, CA: Key Curriculum Press.

Landewehr, J. & Watkins, A. (1987). *Exploring surveys and information from samples: Quantitative literacy series.* Palo Alto, CA: Dale Seymour Publications.

Schaeffer, R., Watkins, A., Gnanadesikan, M., & Witmer, J. (1996). *Activity based statistics.* New York, NY: Springer-Verlag.

References

American Statistical Association. (1991). *ASA guidelines for the teaching of statistics K-12.* Alexandria, VA: Author.

Bloom, S. (1994). "Data buddies: Primary-grade mathematicians explore data." *Teaching Children Mathematics,* 1(2), 80-86.

Bohan, H., Irby, B., & Vogel, D. (1995). "Problem solving: Dealing with data in the elementary school." *Teaching Children Mathematics,* 1(5), 256-260.

Burrill, G., Burrill, J., Coffield, P., Davis, G., de Lange, J., Resnick, D., & Siegel, M. (1992). *Data analysis and statistics across the curriculum: Curriculum and evaluation standards for school mathematics addenda series, grades 9 -12*. Reston, VA: NCTM.

Burrill, G. (1994). *Teaching statistics: Guidelines for elementary through high school.* Palo Alto, CA: Dale Seymour Publications.

Crossen, C. (1994). *Tainted truth: The manipulation of fact in America.* New York, NY: Simon & Schuster.

Department of Mathematics and Computer Science of the North Carolina School of Science and Mathematics. (1988). *Data analysis.* [Materials and software]. Reston, VA: NCTM.

Eisenhower National Clearinghouse. (1996). *ENC Focus: Real data resources for teachers.* Washington, D.C.: Author.

Friel, S.N. & Corwin, R.B. (1990). "The statistics standards in K-8 mathematics." *Arithmetic Teacher,* 38(2), 73-77.

Hawkes, J.S. (1993). *Discovering statistics: An adventure in problem solving.* Charleston, SC: Quant Systems.

Hitch, C. & Armstrong, G. (1994). "Daily activities for data analysis." *Arithmetic Teacher,* 41(5), 242-245.

Kader, G. & Perry, M. (1994). "Learning statistics with technology." *Mathematics Teaching in the Middle School,* 1(2), 130-136.

Landewehr, J.M., Swift, J., & Watkins, A.E. (1987). *Exploring surveys and information from samples: Quantitative literacy series.* Palo Alto, CA: Dale Seymour Publications.

Landewehr, J.M., & Watkins, A. (1986). *Exploring data: Quantitative literacy series.* Palo Alto, CA: Dale Seymour Publications.

Lindquist, M.M. with Luquire, J., Gardner, A., & Shekaramiz, S. (1992). *Making sense of data: Curriculum and evaluation standards for school mathematics addenda series, grades K-6.* Reston, VA: NCTM.

Minnesota Department of Children, Families and Learning. (1997). *Graduation standards: Profile of learning*. St. Paul, MN: Author.

Mokros, J. & Russell, S.J. (1995). "Children's concepts of average and representativeness." *Journal for Research in Mathematics Education*, 26(1), 20-39.

Moore, D. (1988). *Statistics: Concepts and controversies.* New York, NY: W.H. Freeman and Company.

Moore, D. (1990). "Uncertainty." In L. Steen (Ed.), *On the shoulders of giants: New approaches to numeracy*. Washington, D.C.: National Academy Press.

National Council of Teachers of Mathematics. (1990). Focus issue: Data analysis. *Mathematics Teacher,* 83(2).

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics.* Reston, VA: Author.

National Research Council. (1990). *Reshaping school mathematics: A philosophy and framework for curriculum.* Washington, D.C.: National Academy Press.

Parker, J. & Widmer, C.C. (1992). "Teaching mathematics with technology: Statistics and graphing." *Arithmetic Teacher,* 39(8), 48-52.

Russell, S.J. & Friel, S. (1989). "Collecting and analyzing real data in the elementary school classroom." In P. Trafton (Ed), *New directions for elementary school mathematics: 1989 yearbook*. Reston, VA: NCTM.

Russell, S.J., Corwin, R.B., Friel, S.J., Mokros, J., & Stone, A. (1990). *Used numbers: Real data in the classroom.* Palo Alto, CA: Dale Seymour Publications.

TERC. (1998). *Investigations in number, data, and space: A K-5 mathematics curriculum.* Palo Alto, CA: Dale Seymour Publications.

Scheaffer, R.L. (1990). "Why data analysis?" *Mathematics Teacher,* 83(2).

Scheaffer, R., Watkins, A., Gnanadesikan, M., & Witmer, J. (1996). *Activity based statistics.* New York, NY: Springer-Verlag.

Schulte, A.P. (Ed.). (1981). *Teaching statistics and probability: 1981 yearbook.* Reston, VA: NCTM.

Shroyer, J. & Fitzgerald, W. (1986). *Mouse and elephant: Measuring growth: Middle grades mathematics project.* Menlo Park, CA: Addison-Wesley Publishing Co.

Stevens, J. (1993). "Generating and analyzing data." *Mathematics Teacher,* 86(6), 475-478, 487-489.

van Reeuwijk, M. (1992). "The standards applied: Teaching data visualization." *Mathematics Teacher,* 85(10), 513-518.

Vos, K. (1993). "Making sense with statistics." *The California Mathematics Council (CMC) ComMuniCator,* (17)4, 4-5.

Zaldivar, R.A. (1997, March 22). "Why all the fuss about the CPI?" *Miami Herald,* pp. 1-2F.

Zawojewski, J., Brooks, G., Dinkelkamp, L., Goldberg, E., Goldberg, H., Hyde, A., Jackson, T., Landau, M., Martin, H., Nowakowski, J., Paull, S., Shulte, A., Wagreich, P., & Wilmot, B. (1991). *Dealing with data and chance: Curriculum and evaluation standards for school mathematics addenda series, grades 5-8.* Reston, VA: NCTM.

The American Statistical Association has a series of brochures including:

Careers in Statistics

What is a Survey?

Surveys and Privacy

Minorities: Looking for a Challenging Career?

What about Statistics!

Women and Statistics

ASA Guide to Colleges and Universities with Departments of Statistics

They are located at
1429 Duke Street, Alexandria, VA 22314-3402
(703) 684-1221, fax: (703) 684-2036 or 2037

e-mail: asainfo@asa.mhs.compuserve.com

ASA also publishes a magazine for students of statistics called *STATS* and a general-interest publication called *CHANCE: New Directions for Statistics and Computing.*

The ASA/NCTM Joint Committee on Curriculum in Statistics and Probability publishes a newsletter, The Statistics Teacher Network.

Teaching Statistics **is an international journal published three times a year by:**

R.S.S. Centre for Statistical Education
University of Nottingham
Nottingham, NG72RD, UK

e-mail: wayne@asa.mhs.compuserve.com

Internet Resources:

Chance Database
http://www.geom.umn.edu/locate/chance
Abstracts and full text articles in the news, a handbook for teaching a chance course and other teaching resources. This database has websites related to teaching statistics.

Journal of Statistics Education
http://www2.ncsu.edu/ncsu/pams/stat/info/jse/homepage.html
An electronic journal on the worldwide web.

Exploring Data—Education Queensland
http://curriculum.qed.qld.gov.au/kla/eda/
A site with activities, worksheets, overhead transparency masters, datasets, and assessments to support the exploration of data. An extensive collection of articles and a page of resources for introductory statistics are provided, including texts, websites, datasets, java applets, and mailing lists.

The **Exploring Data—Math Forum**
http://forum.swarthmore.edu/workshops/usi/dataproject/
This site provides standards, datasets, lessons and websites at the K-4, 5-8, and 9-12 levels.