Data Analysis and Generalizations

Developing an Investigation

TEACHER GUIDE

BACKGROUND INFORMATION

The *National Science Education Standards* "Science as Inquiry" standard states that "designing and conducting a scientific investigation requires introduction to the major concepts in the area being investigated, proper equipment, safety precautions, assistance with methodological problems, recommendations for use of technologies, clarification of ideas that guide the inquiry, and scientific knowledge obtained from sources other than the actual investigation. The investigation may also require student clarification of the question, method, controls, and variables; student organization and display of data; student revision of methods and explanations; and a public presentation of the results with a critical response from peers. Regardless of the scientific investigation performed, students must use evidence, apply logic, and construct an argument for their proposed explanations."

You may choose to model the development of an investigation from the solar wind information by using a copy of the completed investigation, Student Activity, "A Closer Look at Solar Wind Speeds," or you may choose to have students develop an investigation idea based on one of the questions they encountered in the "Exploration" section. Students will then organize the data in a meaningful way, and will communicate the information in a coherent way. Students will analyze the information in order to answer their questions and to draw valid conclusions. Throughout this section, students are encouraged to use technology and mathematics to make sense of the information that is being analyzed in light of their question(s). Students will use statistical analysis and graphing technology to understand and organize the data, and to communicate their conclusions to peers. If several explanations or conclusions result from this analysis, students will need to make decisions about which explanation best fits the evidence found in the data.

NATIONAL SCIENCE STANDARDS ADDRESSED

(Source: *National Science Education Standards*)

Grades 9-12

Science As Inquiry

Abilities necessary to do scientific inquiry

Understandings about scientific inquiry

(View a full text of the National Science Education Standards.)

PRINCIPLES AND STANDARDS FOR SCHOOL MATHEMATICS

Grades 9-12

Data Analysis and Probability

Formulate questions that can be addressed with data and collect, organize, and display relevant data to answer them

Select and use appropriate statistical methods to analyze data

Develop and evaluate inferences and predictions that are based on data
Understand and apply basic concepts of probability

(View a full text of the Principles and Standards for School Mathematics.)

MATERIALS

- Student Text, "Solar Wind" (if you did not use this in "Catch a Piece of the Sun")
- Student Text, "Exploring Data"
- Model Student Activity, "A Closer Look at Solar Wind Speeds"
- Student Activity, "A Closer Look"
- Computer access to the LANL Web site, or appropriate data printouts of solar wind summary plots from the LANL Web site

PROCEDURE

1. If you did not use the Student Text, "Solar Wind," as a part of the Briefing Activity, "Catch a Piece of the Sun," distribute copies of it now, and review some of the major descriptions of solar wind that relate to the summary plots from the LANL Web site that the students worked with in "A First Look." These should include the following descriptors of solar wind plasma.
   - Stream speeds vary from 300 to 1000 km/sec.
   - It has a temperature of about 10^5 Kelvins.
   - The density varies between 1 and 10 particles/cm^3.
   - It is a mixture of 95% protons (H^+) and 4% alpha particles (He^{2+}).

   Also review the description of "interplanetary shock," and how these shocks are related to solar wind. Give students time to read the remainder of the text if they have not already done so.

2. Have students recall the observations, inferences, and questions that they formulated in the Student Activity, "A First Look," Part 2. Distribute the Student Text, "Exploring Data." Then, using the information in the text, have students prioritize their questions until they have reduced them to a maximum of three.

3. At this point, you have at least three optional methods for guiding students through the process of developing and carrying out a scientific investigation.
   a) You can **model the scientific investigative process** using data from the LANL Web site and analyzing the results as a whole class as shown in the Model Student Activity, "A Closer Look at Solar Wind Speeds."

   b) You can **assign each student or group of students a research question**, either from their list of questions developed in the Student Activity, "A First Look," Part 2, or from a list of possible research questions found below. Students should then use the Student Activity, "A Closer Look," to design and carry out their investigations.

   c) You may assign students to use the Student Activity, "A Closer Look," to develop their own independent research questions and carry out their own research procedures.

4. If you select "option 3. a)," you will take your students, as a class, through the completed procedure as shown in the Student Activity, "A Closer Look at Solar Wind Speeds." You should emphasize the following as you work your way through the research design with your students:
   a) Ask students what differences they see between the original question and the form in which it is written.
      - The testable relationship (which includes the specific variables and the time span) and
      - The null hypothesis that specifies exactly what variables are to be measured and analyzed but predicts no significant differences.

b) Under the "Materials" list, data tables or spreadsheets were listed because it is important to record data in a useable form when it is collected. You can often study the data tables and detect trends or relationships to be graphed or tested for significant relationships.

Also, have students examine the list of variables to be tested and the list of variables to be held constant. Ask them to give possible reasons for excluding the variables listed as those to be held constant. You may want to have them once again examine the data from the LANL Web site that they used while they were doing the Student Activity, "A First Look."

c) Students should next examine the steps outlined in the "Procedure" section. Ask them if these steps are written clearly and specifically enough so that they could **replicate** the investigation. Replication means to reproduce, which is not exactly the same as to duplicate, which means to copy. When scientists replicate someone else's experiment, it means that they use exactly the same procedure, but with different materials and supplies that are exactly the same as those used in the original investigation. The procedure in a scientific report should make it possible for someone else to do things in the same order and in the same manner as the original investigation. If a replicated investigation yields the same results as the earlier work, it provides confirmation. Conversely, if it yields different results, the earlier work may have been incorrect.

Also, point out that it is not enough to collect the data and look at it for trends. The data must be analyzed statistically before conclusions can be made.

d) Under "Results," have students examine the data table. Ask them how the investigator knew exactly what data needed to be recorded. (The investigator had decided how the data was to be analyzed and knew that enough data had to be collected in order to find the mean speeds for all types of the regimes before it could be compared. In other words, the investigator had to think and plan ahead so that the data taken would be useable to answer the investigative question.) Ask students if any trends of significant differences are apparent by observation. (Some students may identify trends, but it is usually difficult to visually analyze this amount of data.)

e) Under "Analysis of Data," ask students how the investigator knew what kind of analysis to perform on the data. (It sometimes takes a lot of experience in working with investigative data before you know exactly how to analyze it. You must look at the data and decide what is appropriate. Students will probably need your help, or perhaps that of the mathematics teacher, to decide what test to run. Since it was not practical to record all data points shown in the LANL Web site summary panels, the investigator decided to measure the high and low temperatures for each regime during each day that they were flowing. That meant that the means of regime speeds needed to be tested against each other for significance. The more data you have, the more reliable is the mean that results, but the data collection method must also be practical.)

f) Have students examine the graph of "Mean Speeds of Solar Wind Regimes." Point out that the graph has the following elements.

- A title that indicates what variables were graphed—mean speeds and solar wind regimes.
- Each axis has a label. The x-axis is labeled "Regime," and the y-axis is labeled "Solar Wind Speeds," and it includes the units, km/s.
- The y-axis is divided into equal units, and the x-axis is divided equally into different regimes.
- The legend indicates that the height of the bars is relative to the mean speeds of the different regimes.

Ask students why the bar graph (or histogram) is more appropriate for plotting this data than a line graph or scattergram. (Line graphs and scattergrams are appropriate when looking for a relationship between two variables, and there is a continuum of data. Neither of these is the case for this data.)

g) Ask students what the column labels describe in Table 3.

- Mean Pairs indicates which two sets of data were being compared. You can only compare two sets of data at one time.

- The symbol df is the abbreviation for degrees of freedom, which is equal to the number of observations minus restrictions. This is automatically calculated in most computerized statistical programs. Usually, the greater the number of observations (therefore, the number of degrees of freedom), the better chance you have of obtaining a significant difference between variables.
- Significance level (symbolized by α) is the probability that the investigator will arrive at an erroneous conclusion that there is a relationship between variables, when, in fact, there is no such relationship. This investigator chose to test the hypothesis at the 0.05 level, which meant that there was only a 5% chance that any relationship found between the variables would be due to something beside the actual difference in wind speeds. Note that means from only two of the Regime pairs are significantly different at the 0.05 level. *T*-tests are one form of statistical analysis that are appropriate for comparing means of two sets of data. The investigator can use statistical tables that show whether these numbers indicate significance at the 0.05 level of confidence. The *t*-stat column shows the actual calculated value of the *t* distribution for this specific pair of means. The *t*-Critical two-tail column shows the minimum value that the *t*-stat must achieve to be significant at the .05 level.

h) Have students compare the wording of the paragraph under Table 3 and the "Conclusions" paragraph. They should be aware that the "Results" paragraph only stated that there were significant differences found. The wording in the "Conclusion" statement should either accept or reject the null hypothesis. Accepting the null hypothesis would have meant that there were *no* significant differences between mean speeds in the solar wind regimes. Rejecting it, however, means that there *were* significant differences found. Note also the restriction on the conclusions, that there were significant differences found *during the month of December 2002*. You must be very careful not to project significances.

5. If you decide on "option 3. b)," the following are suggestions for possible research questions.
   a) Using data for a given month:
   - Compare the means of minimum and maximum wind speeds for fast, slow, and CME regimes (this is the data used for the model student activity).
   - Compare the means of minimum and maximum temperatures for fast, slow, and CME regimes.
   - Compare the means of minimum and maximum densities for fast, slow, and CME regimes.
   - Compare the means of maximum and maximum He/H ratios for fast, slow, and CME regimes.
   - Compare the means of length of time for fast, slow, and CME regimes.
   - Is there a relationship between increasing or decreasing trends in wind speeds and temperatures?
   - Is there a relationship between increasing or decreasing trends in wind speeds and densities?
   - Is there a relationship between increasing or decreasing trends in wind speeds and He/H ratios?
   - Is there a relationship between increasing or decreasing trends in temperatures and densities?
   - Is there a relationship between increasing or decreasing trends in temperatures and He/H ratios?
   - Is there a relationship between increasing or decreasing trends in densities and He/H ratio?

   b) Comparing the same characteristics for two different months:
   - Are there any replicable changes in wind speeds, temperature, densities, or He/H ratios during interplanetary shocks?
   - Is there a relationship between the length of time that a specific regime "flows" and the magnitude of differences in speed, temperature, density, and/or He/H ratio?
   - Do regime changes follow the same pattern, i.e., fast→slow→CME→slow→fast, or some other pattern?
   - Do regime changes from slow to fast cause increases in speed, temperature, and/or density?
   - Do regime changes from fast to slow cause decreases in speed, temperature, and/or density?

   Use the questions and information in Procedure 4 above to help students complete the Student Activity, "A Closer Look." You may also require students to make an oral and/or visual aid presentation of their findings. You should construct a rubric that includes the factors and descriptors that will be used to evaluate their presentations—either by you or by their peers.

6. If you use "option 3. c)," students should complete the Student Activity, "A Closer Look," to plan and carry out their original questions. Use the questions and information in Procedure 4 above as you guide their investigations. You may also require students to make an oral and/or visual aid presentation of their findings.

You should construct a rubric that includes the factors and descriptors that will be used to evaluate their presentations—either by you or by their peers.

7. Whatever option you choose, you should include a follow-up session to discuss student results, the problems they encountered as they conducted their investigations, how their results have increased their understanding of solar wind regimes, and what further questions they have as a result of their investigations.

TEACHER RESOURCES

http://members.aol.com/johnp71/javastat.html
This site lists Web pages that comprise a powerful, conveniently accessible, multi-platform statistical software package. There are also links to online statistics books, tutorials, downloadable software, and related resources. All of these resources are freely accessible and include the following hot links:

Selecting the right kind of analysis
Calculators, plotters, function integrators, and interactive programming environments
Probability distribution functions: tables, graphs, random number generators
Descriptive statistics, histograms, charts
Confidence intervals, single-population tests
Sample comparisons: t-tests, ANOVAs (analysis of variance), non-parametric comparisons
Contingency tables, cross-tabs, Chi-Square tests
Regression, correlation, least squares curve-fitting, non-parametric correlation
Analysis of survival data
Bayesian methods
Other statistical tests and analyses
Specialized and discipline-specific tests and analyses
Power, sample size, and experimental design