WHAT DO M&M'S, DAHLIAS, SOIL EROSION, AND DATA ANALYSIS ACROSS THE CURRICULUM HAVE IN COMMON?

Jerry Moreno, John Carroll University
Dept. of Mathematics and Computer Science, University Hts. OH 44118

Abstract

The primary purpose of this presentation is to encourage you to become involved in your local school's statistics offerings. Most states and school districts are committed to NCTM's standards documents of 1989 (1) and 2000 (2) regarding incorporating statistics into the K-12 mathematics curriculum. Yet, the implementation of such into the classroom remains painfully slow, apart from AP Statistics whose growth truly has been amazing. I will present my involvement in three classroom activities last year and one school project with the hopes of your saying "I can do that, and I will!"

Keywords: Statistics activities, NCTM Standards, k-12 curriculum, classroom visitations

1. Overview

Most of what I have done in statistical (quantitative) literacy over the past 20 years has been workshops for teachers, k-12. More recently, I have also been working with their students. Admittedly, doing a one-time presentation in a class can be of marginal worth, perhaps it is only a "filler" for a teacher. But it's a start upon which to build a relationship with the teacher so that s/he will replicate your lesson in other contexts throughout the year. Only then will data analysis become an integral part of their curriculum.

Unfortunately these days, no matter how neat and interesting we think statistics is, presenting statistics for statistics sake will not get your foot in the school's door. Teachers and curriculum directors will want to know where what you are doing satisfies standards. So, you should be somewhat familiar with NCTM's Principles and Standards (2), your state's standards (that hopefully are based on those of NCTM), proficiency/graduation tests, and the *Guidelines and Assessment for Instruction in Statistics Education* (GAISE) by ASA (3) that provide a curriculum framework for prek-12 statistics education.

All activities I do whether with teachers or students focus on understanding the concept that data are numbers in a context, that our task is to make sense of the variability that exist in data and that the statistical process, no matter the level, is an investigative one that involves: 1) formulating a question, 2) collecting appropriate data regarding the question, 3) analyzing the data, and 4) interpreting the results in the context of the question. At the elementary school levels, step 3) is often restricted to graphing techniques with the mode as a summary measure. At upper elementary, the median and mean are usually included. A measure of spread is appropriate at the middle school level, particularly the mean absolute deviation; the quadrant count ratio can be used to introduce measuring the degree of relationship between two variables. At the high school level, students are ready to understand the standard deviation as a measure of spread, the correlation coefficient as a measure of relationship, and resampling techniques as an introduction to inference. Normal theory should be restricted to advanced placement statistics.

Allow me to assume that you have relatively easy access to your local school personnel, and that this talk will encourage you to begin becoming involved in helping your school district implement statistics/data analysis into their curriculum.

2. An M&M's Activity in a Primary Grade

Nate Spicuzza is my grandson whose first grade teacher, Mrs. Sunny Doxey, invited me to do a data activity in a 50-minute class period. As wellness programs grow in strength and adoption in school districts throughout the country, the use of M&M's even in classroom activities (lest students "eat their data") may be frowned upon. (Actually, be sure to brief the teacher well beforehand on what you are going to do if it involves any foodstuffs. In doing so with Mrs. Doxey, she alerted me that one of her students was not permitted to have chocolate and that she had an alternative treat for him if the rest of the students were going to eat their M&M's.)

I did two experiments in Nate's class, one involving estimating, creating a dotplot, identifying the mode; the other involved graphing the distribution of a package of M&M's colors using a pie chart, ordering sizes.

2.1 Estimation, dotplot, mode

Each student was given a fun-size bag of M&M's and I asked "How many M&M's are in a bag?" Their estimates went from 12 to 83, almost all were different. We plotted them on a dotplot on the board using post-its and talked about variability as not all of their estimates were the same. A companion question was "Does every bag contain the same number?" Be careful in asking too many questions, no matter how related. Young children have a ton of questions of their own and can easily lose track of the major point(s) being made. We plotted their actual numbers on the same dotplot as their estimates but in a different post-it color. We immediately saw that the variation was much smaller as the numbers went from 17 to 24 (1-17, 3-20, 6-21, 10-22, 4-23, 1-24). We discussed a typical number in the sense of most often as the mode, and even discussed the concept of outlier. To make sure our "outlier person" wasn't picked on, I gave her an extra bag indicating that being different may have advantages. We described the shape as symmetric apart from the outlier, so they learned three new words – mode, outlier, symmetric.

2.2 Graphing the distribution of colors, pie chart

For the second activity, I asked "Does each color occur the same number of times?" They didn't think so as they looked at the contents of their bags lying in front of them. To show their conjecture graphically, I had them arrange their M&M's by color and by order from largest to smallest spread evenly on the inner ring of a paper plate.

The inner ring is typically a little too large for the number of M&M's they may have; spreading them around evenly is a little difficult for some students. That reminds me, don't try to do these activities alone. In my case, 25 energetic kids required Mrs. Doxey, a parent, myself, and even the school principal Mr. Tim Barrett stopped in so we put him to work too as seen in the photograph.

The next step was to take a blue crayon and draw a line from the center of the pie plate to the first blue M&M, and another line from the center to the last blue M&M, and similarly for the other colors, red, yellow, green, orange, and brown. After having colored in the sectors according to the outlined color boundaries, they then proudly showed each other their pie charts, a new graph for them, and decided for sure that the numbers of each color were not the same. Here's Nate's.

If we had time which we didn't, I was going to have them redraw their pie charts in the order in which the magnitude of M&M's are actually produced: blue 24%, orange 20%, green 16%, yellow 14%, brown 13%, and red 13% (4).

2.3 Connections

Regarding connections to the curriculum, Mrs. Doxey had already introduced her students to the bar graph as they had been graphing the number of teeth each child by name had lost since the beginning of the school year. She noted that the pie chart would not be a good graph for that since there would be too many sectors, but that a bar graph could be used for the color distribution of M&M's. She said that she would use the terms mode and outlier during the year. Further, I need to discuss with her some specific exercises that are a part of her current lesson plans that can be thought of in the context of the four components of statistical problem solving.

Clearly, the NCTM Content Standard for Grades PreK-2 on Data Analysis and Probability was addressed. In addition, the Process Standards covered were primarily Communication and Representation.

3. Soil Erosion –
A Fourth/Fifth Grade Science Experiment

Where are data collected on a regular basis in the school curriculum? Science labs! What could be more natural than using the statistical process to inform the scientific method and the scientific method to form a framework for statistical analysis?

I contacted Mrs. Lynn Fagerholm who is the science teacher facilitator for some of the elementary grades in my school district. We discussed the possibility of having her students use quantitative literacy techniques, primarily graphs, to confirm comparative results of data collected in their labs. She was very interested.

3.1 The experiment

I visited her lab on soil erosion in which teams of four students had two troughs, each four feet long, one of which was filled with soil only, and the other with a comparable amount of soil but with grass growing in a very healthy root system. I helped the students try to maintain scientific protocol in collecting their data (a job much easier said than done, by the way); it was quite hectic as there were 42 minutes in which to get all the data collected, and the lab cleaned for the next class. Here are Sarah Wilkes and Brady Burns conducting the turf part of the experiment.

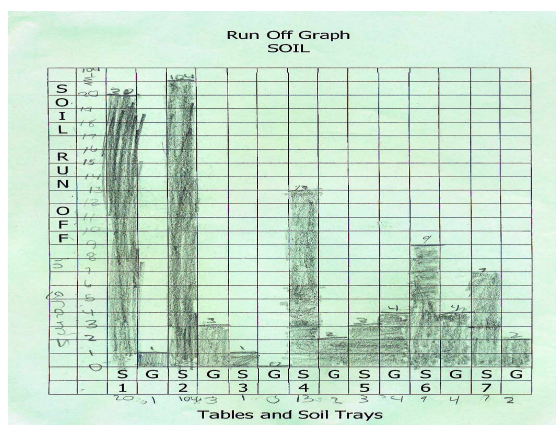
There were many scientific protocol issues for Mrs. Fagerholm's "young scientists," as she calls them, to consider, such as: equal elevations of one end of the troughs; similar nylons tied to the other end to collect the eroded soil; equal amounts of water to pour into the high end being careful that none spills over the sides; appropriate times to stop the experiment; similar squeezing of the nylons to remove the water from the eroded soil; accurate measurements taken of the amount of eroded soil collected in the nylon (they had massed the nylon prior to the experiment, so that the difference represented the amount of soil collected). There were seven teams of three or four students, hence seven sets of data.

3.2 The analysis

The main question was "Does plant growth help prevent soil erosion?" Their experiment produced seven sets of data:

RUN OFF DATA					
Table	SOIL (g)			WATER (mL)	
	No Grass Tray (S)	Grass Tray (G)		No Grass Tray	Grass Tray
1	20	1		450	325
2	104	3		500	275
3	1	0		425	475
4	13	2		450	350
5	3	4		420	500
6	9	4		360	250
7	7	2		400	375

The first thing they did was to graph their results. They decided to use a bar graph.

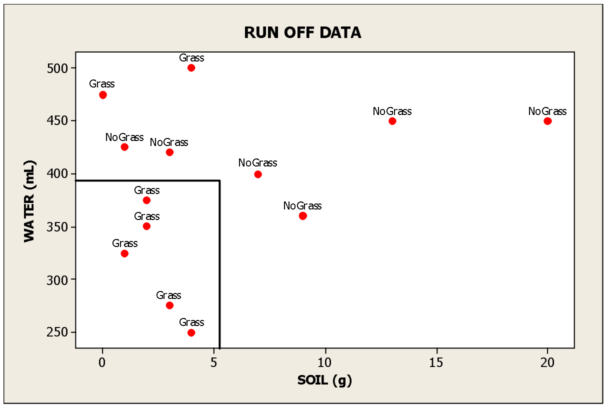


This was their bar graph for the amount of soil run-off for the soil without grass data (denoted by S) and the grass data (G) per table. They constructed a similar graph for the water run-off. Note that the 104 gram data in the No Grass Tray for table 2 prompts science protocol discussion and statistics outlier discussion.

They saw that except for one table, the amount of soil run-off with no grass was higher than with grass. This suggested that the presence of turf matters in helping to prevent soil erosion.

I suggested that they may want to look at the data in another way by aggregating their table data in a dotplot to compare them. Depending on the level of the students, you could talk a bit about inference and introduce Tukey's Quick Test.

Further analysis was done by taking the pairings of (soil, water) into account per table. These data were graphed in a scatterplot identifying which points came from the No Grass Tray and which were generated using a Grass Tray. Note that the outlier No Grass entry has been omitted in the following scatterplot.



Students should detect that the data points falling in the lower left part of the graph are all from the Grass trays (5 of 7 shown inside the superimposed rectangle on the graph). Moreover, the five data points (including the outlier) that are to the right of the rectangle all have substantial water loss and are from No Grass trays.

3.3 Connections

There are many experiments in the upper elementary (and beyond) science labs that are either comparative in nature or can be made so. If the two sets of data are independent, drawing a dotplot, conjecturing a conclusion from it, possibly applying Tukey's Quick Test should be automatically done as a natural part of the scientific method. If the data are paired, then scatterplots with a no effect dividing line if appropriate should become a natural part of arriving at a conclusion.

The NCTM Content Standard for Grades 3-5 in Data Analysis and Probability was covered as well as all of the Process Standards: Problem Solving, Reasoning and Proof, Communication, Connections, and Representation.

4. Statistics, Science, and Dahlias - Another Elementary School Project

Most people would consider it time to retire and relax after having worked two or three quite successful, and quite different, careers - no more stress, no demanding responsibilities, no deadlines – "been there, done that" many times over. But that's not for an absolutely marvelous woman by the name of Roberta "Granny" Paolo. Although gardening was not a part of any of her career paths, it has been a part of her life since childhood. Bypassing the rocking chair, she instead went to her grandson's Loveland Ohio School Board a couple years ago expressing "Our children in primary and elementary school should be gardening as part of their science curriculum." Fortunately for the children of Loveland, the School Board agreed and therein began yet another career for her - "Granny's Garden School." See www.grannysgardenschool.com. Her vision was not to have a couple petunias here, a sunflower there, a manicured landscape bed around the schools' entrances. No indeed. Multitudes of flowers and vegetables and plants, you name it, are growing on practically every square foot of the land around the Loveland elementary and primary schools in beds, buckets, containers of all sorts – a long semicircular hoop frame has morning glories growing over it which serves as an outdoor reading room for children! And it is the children who plant, the children who pick, the children who harvest, while school is in session as well as over the summer (with the help of their parents). The children take the flowers to nursing homes and hospitals in the area. The crops are sold and the proceeds are used in the schools. Throughout it all Granny has a wonderful group of "Granny's Gardeners" who volunteer their time and talent to help her manage this colossal, awe-inspiring year-round educational project.

So what does this have to do with dahlias, you no doubt are asking. Well, the amazing dahlia is one of the children's favorite blooms. They grow dahlias along a couple hundred foot fence. Roberta obtains as many donations as she can for everything. And so I, then president of the Dahlia Society of Ohio, and the rest of the seventeen presidents of the American Dahlia Society Midwest Conference received emails from Roberta asking if we might be able to donate a few dahlia roots. I thought DSO could do that I replied, thinking that to be the end of it. A couple days later, I received another email from Roberta, thanking me for the reply but also mentioning that she had been surfing the web and found that I am a professor of statistics, a member of a school board, and that I have given workshops for teachers and students in data analysis. She asked if I might be interested in possibly incorporating statistics as part of Loveland's third or fourth grade science curriculum. "Would I ever," I replied, thinking I had died and gone to heaven!

4.1 Dahlias, science, statistics

After having thought for a considerable amount of time on what we might be able to do to tie the scientific method, statistics, and dahlias together for third/fourth graders, I suggested that the science question to answer would be "Does compost matter?" Granny allows no fertilizers to be used, only compost. I met with four classes in April, 2005 to discuss scientific protocol and elementary statistics. We discussed what measurements they thought should be taken to answer the question. We decided to plant 16 plants of a certain variety called Baron Katie; eight would be planted in composted soil and the other eight in non-composted soil. The students suggested several variables to measure: weekly heights of the plants, time of first bloom, size of bloom, number of blooms per plant. Their knowledge of statistical measures from their math class focused on the median. To get an idea of what the dahlia experiment would look like, I introduced dotplot and scatterplot graphs with them by having them measure 16 daisies since dahlias weren't available in April. (We used Tukey's Quick Test to compare.)

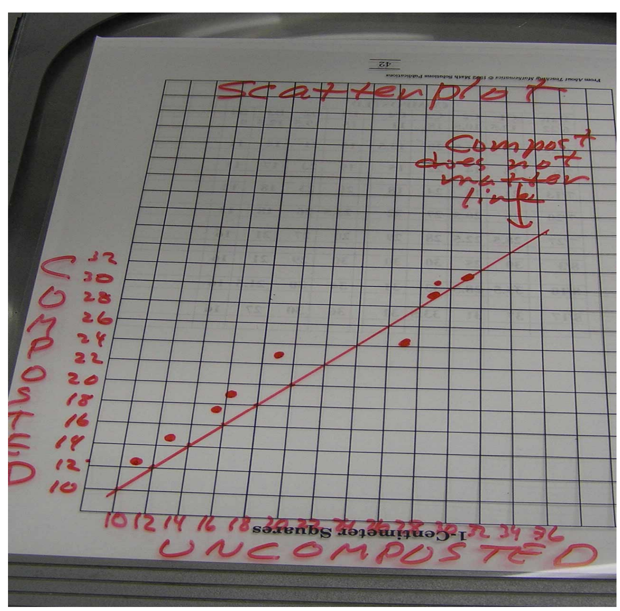
Since it is over a four-hour drive from Cleveland to Loveland, I wasn't able to observe the actual planting of the two experimental beds, nor supervise the measurements over the summer. As it turned out, only height measurements were taken in this first year. I thank Paula Carlsen, one of Granny's Gardeners, and her children for being so diligent in recording the heights of all 16 plants for each of nine consecutive weeks over the summer.

4.2 The data and the analysis

In October, I visited one of the classes of fourth grade science teacher Brandi Carter. As these were not the students I had talked with the previous April, I had to go over the experiment with them. The experimental plots were outside their window so we viewed them from the classroom. We discussed the design of the experiment and the height data, and decided to calculate and compare the median heights of the eight composted plants with the median heights of the eight non-composted plants for each of the nine weeks. A couple students suggested that we use the mean but not all of the students had learned the mean yet. There was another reason not to use the mean. One of the non-composted plants died in the fifth week. Calling its height zero did not cause any problem using the median but using the mean would have introduced a statistical difficulty. We then plotted the nine paired medians by date on a scatterplot with the non-composted scale from 10 inches to 40 inches on the horizontal and a similar composted scale on the vertical. The paired data by week were:

Date	Non-composted	Composted
6/22	11.25	12.5
6/29	13	14.5
7/6	15.5	17.5
7/13	16.5	19
7/20	19.75	22.75
7/27	27	24.75
8/3	28.5	29.5
8/10	29.25	30.5
8/17	31.5	31.0

Although I learned from the fourth grade mathematics teacher that the students were just beginning to learn the coordinate system in two-dimensional graphs, I banked on their having played Battleship since "youngsters" and forged ahead to plot the data in a scatterplot. In three words, they were terrific! Each drew a scatterplot by first labeling the horizontal and vertical axes as mentioned above. Then they drew a diagonal line that they called the "compost doesn't matter line" indicating where compost height = non-composted height.



They analyzed the graph by observing that the first five data points were above the diagonal line indicating that the composted heights were taller than the non-composted ones, then the sixth week produced a data point below the line that prompted one of the students to respond "the non-composted plants had a growth spurt!" The next two weeks were in favor of the composted plants once again with the last week practically a tie. The students concluded that for most of the summer compost seemed to matter a little bit but at the end there was no difference.

We didn't have time to discuss the results in depth. I learned from Paula who supervised the summer data collection that the plants could have been tied a bit better. Also, it wasn't clear to me that the only difference in the soil was the presence or absence of compost. So, I am looking forward to trying the experiment again next year with tighter control of the scientific protocol and experimental conditions.

4.3 Connections

So, a class of fourth grade students began to understand the synergy between the scientific method and statistics, for just as with the scientific method, statistics is an investigative process that begins with a question, collects data to answer the question, organizes the data, analyzes the data by measures and graphs, and interprets the results. And the dahlia is being used in a real gardening experience to enhance the children's science, mathematics and language arts skills in the spirit of Ohio's curricular standards.

The NCTM Content Standard for Grades 3-5 in Data Analysis and Probability was covered as well as all of the Process Standards: Problem Solving, Reasoning and Proof, Communication, Connections, and Representation.

5. Data Analysis Across the Curriculum

As the result of an extensive self-study conducted by the teachers and administrators in a nearby catholic high school in Akron, Ohio, St. Vincent – St. Mary, they identified a major goal: to improve their students' quantitative reasoning. Their strategy to achieve this goal was to put in place a program that implemented data analysis into every subject area. As the headmaster, Mr. David Rathz, knew me, he asked if I would be interested in helping them begin to understand data analysis topics and to give them examples per subject area.

On January 24, 2005, I met with the entire staff of 50 teachers and administrators, and went over a 60-page handout that I wrote and gave to each of them. The idea was to highlight major statistical ideas, have them meet together over the next three months both departmentally and inter-departmentally to apply what they learned to their goal, compile a list of questions on material they did not completely understand, and reconvene as a staff in another workshop in April.

5.1 Agenda for the January workshop

"To understand God's thoughts, we must study statistics, for this is the measure of His purpose."
Florence Nightingale

I: The Importance of Having a Statistically Literate Citizenry

- To read a newspaper more intelligently, to critique advertisements, political claims, health studies
  Examples from the newspaper
- To become a better decision maker – in life, and profession
- Examples of Famous Misuses/Uses of Statistics
  - The 1970 Draft Lottery
  - Alf Landon, FDR, The Literary Digest, Gallup
  - Infant deaths, the thymus
  - The Challenger Disaster
  - Authorship of the Federalist Papers

II: Basic Terminiology

- Definition of Statistics
- Types of Data…qualitative (categorical), quantitative (numerical)
- Types of Graphs…
  - Qualitative…pie chart, bar graph, pictograph
  - Quantitative…dotplot, stemplot, histogram, boxplot, pictograph, scatterplot
- Measures of Center…mode, median, mean
- Measures of Spread…mean absolute deviation, standard deviation
- Measures of Relationship…quadrant count ratio, correlation coef.

III: Projects Across the Curriculum – Data From the Real World

- The 1918 Pandemic
- People Count!, population pyramids

IV: Resources (I provided them with a long list.)

V: Appendix

- Math activity on discovering the triangle inequality
- Data and discussion of the 1970 Draft Lottery
- Data and analysis of the Challenger disaster
- Examples of incorrect pictographs and corrected versions of the data
- A few Ohio Graduation Test (OGT) examples

5.2 Agenda for the April follow-up workshop

April's agenda was based on the questions that they had as a result of their meetings since January:

I: Types of Data and Types of Graphs

- Data…qualitative (categorical), quantitative (numerical)
- Graphs...
    - Qualitative…pie chart, bar graph, segmented bar graph, pictograph
    - Quantitative…dotplot, stemplot, histogram (including population pyramids), boxplot, line graph, pictograph, scatterplot

II: Examples of Activities and Their Analyses Graphically++

- Boxplots and outliers
- Scatterplots and the Quadrant Ratio Count
- Median Fit Lit
- Transformations to Linearity

III: Hands-on Activities

- Is there a relationship between brain weight and body weight in animals? If so, describe it.
- Who wrote the ten letters that were signed Quintus Curtius Snodgrass?
- Investigate whether there is evidence of gender discrimination in promotion.
- How sinister the full moon?
- What is the effect of light and darkness on mustard root growth?

IV: Resources (additional ones to the January workshop list)

- Excellent Data Sets on the web
- Books on good and bad ways of collecting and displaying data, statistical thinking and interpretation

5.3 Review in February 2006

I returned in February of this year to meet with the leaders of the project and review their progress. They showed me and discussed with me projects that were done in almost every high school department. Their progress toward reaching their goal was impressive. The overall project will be assessed in three years.

6. Summary Comment

I hope that you were able to identify with at least one of the projects I've described, and that you have found it interesting enough for me to overhear you say "I can do that, and I will!"

References

(1) *Curriculum and Evaluation Standards for School Mathematics,* National Council of Teachers of Mathematics, Inc., Reston VA, 1989
ISBN #: 0-87353-273-2

(2) *Principles and Standards for School Mathematics,* National Council of Teachers of Mathematics, Inc., Reston VA, 2000 ISBN #: 0-87353-480-8

(3) *Guidelines and Assessment for Instruction in Statistics Education*, American Statistical Association, Alexandria VA 2006
http://www.amstat.org/education/gaise/

(4) Color distribution of M&M's.
http://www.mms.com/us/about/products/milkchocolate