# The Statistics Teacher Network

www.bio.ri.ccf.org/ASA/stn.html

**Number 57** ASA/NCTM Joint Committee on the Curriculum in Statistics and Probability **Spring 2001**


*Book Review...*

## *Review of Amsco's AP Statistics: Preparing for the Advanced Placement Examination*

When preparing for an Advanced Placement exam, I feel that it is important to consider two issues: covering the content and simulating the format. Students should know the material as well as be comfortable with the test design and style of questioning. So it is that criteria by which I judge the quality of *Amsco's AP Statistics: Preparing for the Advanced Placement Exam* by James F. Bohan (with Beth Chance).

The softcover book opens with an extensive review section that covers all of the topics on the AP Statistics syllabus. The subject matter is covered in eight chapters, arranged in an order that is consistent with many popular textbooks. Each chapter is followed by review questions (both multiple choice and free response) and answers and solutions to all are given at the back of the book. Additionally there are review questions within the text specific to each particular topic. The writing is clear and concise and the author does a nice job of discussing the topics. I found some of the explanations to be better than those in the text I am using for my A.P. Statistics class (David S. Moore's *Basic Practice of Statistics*). Chapter questions are spiraled, with review questions from previous chapters thrown in to keep the student on their toes.

As for format, again this book hits the mark. The style of questioning is consistent with the Advanced Placement exam, with both Multiple Choice and Free Response questions utilized throughout. At the end of the book there are 5 model exams. These exams follow the exact format of the real exam with 40 multiple-choice questions, 5 free response and 1 investigative task. Answers and solutions to all exam questions are provided. For the most part, the problems on the exams are well-written. Not all are what I would deem AP-quality but there are enough excellent questions to provide for good preparation. After taking a few of these tests, a student should be comfortable with the exam design and the style of questioning.

There are also a number of other positive notes about this book. It contains a useful appendix with appropriate formulas, tables, technology resources, websites and a list of reading references. Also, there are technology notes throughout for various statistical calculators and computer software as well as sample output from Minitab, SAS, DataDesk and Fathom.

There are a few negatives worth mentioning. Most notable are the more than 50 typographical errors to deal with. Fortunately, the authors have provided a list of errata on a website (*http://www.calpoly.edu/~bchance/APReview/errata.html*) to assist the reader. Some of the questions require students to perform tedious calculations such as computing a sample standard deviation from a data set, which are not in the spirit of the AP course. Additionally, the

## Also In This Issue...

Chapter 4 multiple-choice question set has too many "E" answers such as "none of the above" and "all of the above". Not all the questions are perfect, but the majority provides worthwhile practice for the student.

On the whole, this is a quality reference that should be helpful to both the student and teacher of AP Statistics. I feel that it is the best AP Statistics review book available. The exposition is well written, thorough and attractively designed. It can be used as both a supplement and as preparation for the exam in the spring. The price is attractive as well (around $14.00). My school purchased a class set this year and my students have found the book to be quite beneficial. I look forward to trying out the model exams on my students in April in preparation for the May 17 exam. I am confident that this review tool will leave them in an excellent position to succeed.

Information on purchasing the book:
ISBN: 1-56765-527-0
Order Code: R686W
To order, call Amsco Customer Service at 1-800-969-8398.
In New York City, call 212-886-6565.
FAX: 212-675-7010.
For an order form or the name of a sales representative in your region, visit *www.amscopub.com*

Dan Shuster
Royal High School
Simi Valley, California
*dshuster@simi.k12.ca.us*

*Book Review*

## *Introduction to Statistics and Data Analysis*

***Peck, Olsen, Devore***
***Duxbury Press, 2001***
***ISBN 0-534-37092-6***

There are excellent textbooks on the market for AP Statistics. We AP Stat teachers usually rely on several – one as the adopted text and others for reference. I believe the ultimate AP Statistics textbook has arrived. ***Introduction to Statistics and Data Analysis*** by Roxy Peck, Chris Olsen, and Jay Devore takes everything that is wonderful in the Devore, Peck textbook *Statistics, The Exploration and Analysis of Data, 3rd edition*, fills in all the holes and develops all the concepts needed for a complete advanced placement course. This book review is, for me, a labor of love; I am using this textbook and continue to be surprised and delighted with how good it really is.

The order of topics presented is the first hint that the authors have a distinctive vision for their textbook. After the obligatory introduction to statistics, they present the experimental design chapter. The chapter is beefed up from a single section in Devore, Peck to a meaty chapter in Peck, Olsen, Devore. All important AP concepts are demonstrated: randomization, blocking, control, replication, bias, confounding variable, and more. There is no longer any need to supplement the material; it's all there. The student now begins the descriptive statistics chapters with an understanding of data collecting, sampling, and experimenting that enriches all of the topics that follow.

The ordering of the chapters for the rest of textbook is the same as other textbooks: exploring data, probability, basic one-and two-sample inferential techniques, and more advanced inferential methodology. But, again, the authors put their own unique touches on the order of topics presented within the chapters. For the chapters on confidence intervals, one-sample and two-sample hypothesis tests, Peck, Olsen, Devore present inferences for proportions before those for means. The student then uses the more familiar z distribution first. It is also notable that the authors quickly move to the t distribution when discussing inferences on means. They make the point early on, in the chapter on confidence intervals, that we seldom know the population standard deviation $\sigma$, and so must estimate it with the sample standard deviation s. As a result, right away students must become familiar with the t distribution and the conditions for its use. There are few examples or exercises that claim to know $\sigma$, so in that sense the textbook is more realistic.

The probability chapter is another example of a hole in the Devore, Peck book that the current textbook has remedied. The chapter discusses in detail basic properties of probability, conditional probability, general rules, and simulation. All topics are thoroughly discussed and presented in exercises; and with an abundance of problems, there is plenty of opportunity for drill and practice here. Devore, Peck had

an outstanding simulation section. Peck, Olsen, Devore have an outstanding probability chapter, including simulation, binomial and geometric distributions, Venn diagrams, discussion of independence and mutually exclusive events, and presentation of the Law of Total Probability and Bayes' Rule.

Another striking feature of *Introduction to Statistics and Data Analysis* is the recurring theme of transforming data. There is a lengthy section on transforming nonlinear data in the chapter on summarizing bivariate data. This section is a significant enhancement of the Devore, Peck section and is another example of a filled-in hole. The authors revisit transformations in the section on checking for normality and transforming nonnormal data. Checking for linearity of the normal probability plot echoes back to the earlier chapter.

The material presented in *Introduction to Statistics and Data Analysis* appears in other textbooks as well. What makes this textbook unique and original? First and foremost to me are the examples and exercises. Almost all of the problems presented come from real world studies and the actual study or paper is cited. We find out that the paper "Objective Measurement of the Stretchability of Mozzarella Cheese" (J. of Texture Studies, (1992): 185-194) reported on an experiment to investigate how the behavior of mozzarella cheese varied with temperature. (Problem 5.4, page 150). We also learn that the eating habits of 12 bats were examined in the article "Foraging Behavior of the Indian False Vampire Bat" (Biotropica (1991): 63-67). These bats consume insects and frogs. (Problem 9.40, page 468). In problem 2.24, page 25, we are told that, based on observing more than 400 drivers in the Atlanta area, two investigators at Georgia State University concluded that people exiting parking spaces did so more slowly when a driver in another car was waiting for the space than when no one was waiting ("Territorial Defense in Parking Lots: Retaliation Against Waiting Drivers," J. Applied Social Psych. (1997): 821-834).

The problems are informative, which one would expect, but they also frequently reveal a sense of humor that we don't usually see in a math textbook. For example, the authors question if it is really the case, as it might seem to an unsuccessful and frustrated angler, that 10% of those fishing reel in 90% of the fish caught? (Problem 3.22, page 80). In response, they cite data from a survey of 911 anglers done during a particular period on the lower Current River in Canada ("Fisherman's Luck", Biometrics (1976): 265-271). We learn in problem 11.41, page 566, that dentists make people nervous (even more so than statisticians). Then there is my favorite (Problem 11.28, page 554): Here is one to sink your teeth into: The authors of the article "Analysis of Food Crushing Sounds During Mastication: Total Sound Level Studies" (J. of Texture Studies (1990): 165-178) studied the nature of sounds generated during eating. Peak loudness (in decibels at 20 cm away) was measured for both open-mouth and closed-mouth chewing of potato chops and of tortilla chips. Forty subjects participated, with ten assigned at random to each combination of conditions (such as closed-mouth, potato chip, and so on).

Technology plays an important role in AP Statistics and is featured throughout *Introduction to Statistics and Data Analysis*. Computer output from Minitab and other software packages appears frequently in the text and in the exercises. Consequently, if students do not have access to computers with statistical software, they can still learn how to read and interpret computer output.

Graphing calculators with statistical capabilities have become an essential tool in AP Statistics. To present calculator techniques, the authors have included graphing calculator explorations at the end of each chapter. There are no calculator keystrokes appearing in the body of the chapter; these show up in the end-of-chapter explorations. These explorations take the topics covered in the chapter and provide students with enrichment activities using the calculator. It is very interesting to note that no particular brand of calculator is featured. The commands are generic and do not refer to any particular calculator. For specific keystrokes, the authors recommend referring to the calculator user's manual. As a result, one has the sense that the authors do not favor one calculator over another. This textbook is not an 800-page advertisement for one particular brand of calculator. The explorations, which are independent of each other, are interesting and informative. The instructor can pick and choose the calculator explorations that are relevant to his or her course.

Also revealed in the calculator explorations are suggestions for communicating calculator output. This is a real benefit for students. They are shown what a good answer looks like when using a calculator to answer a problem. For

example, in the graphing calculator exploration for section 5.3 on linear regression, the authors give a lesson in communicating an answer to a linear regression problem. It is as if Roxy, Chris, and Jay are standing right there telling us "no calculator language, show your steps, define your variables." They state, "We believe that when answering a question, your fundamental task is not only to provide an answer, but to communicate the method by which you got that answer. The calculator...is singularly unable to communicate effectively."(pages 216-7). The discussion that follows shows the student how to write the results of linear regression. We are told to write the regression equation with variable names, not x and y. When using the linear model to predict a value, show the equation with variable names, write the equation with numbers substituted in, and then compute and write the predicted value in equation form, again fully labeled.

Another remarkable feature of *Introduction to Statistics and Data Analysis*, are the summary tables of rules, requirements and formulas for each type of confidence interval and hypothesis test. This is a particularly useful feature. For confidence intervals, we are shown the formula for the interval and the requisite assumptions in an easy to learn list. The information is boxed off from the text, making it stand out on the page. For each hypothesis test, the important components are explained. We see, again boxed off from the text, a summary of the null hypothesis, the test statistic, the alternative hypothesis, p-value interpretation, and the relevant assumptions. In another box nearby, we find the nine steps for hypothesis testing analysis listed. This feature is especially appealing for those students (and teachers) who like a check off list of steps to perform. These boxed summaries in each section will be extremely useful at AP Exam review time. The students can go straight to the boxes for the interval or test they want to review and find everything they need to know and do right there. Every year at the exam reading, we see the importance of stating assumptions, defining the variables, correctly writing the hypotheses, using the appropriate test, and drawing the correct conclusion in the context of the problem. This textbook emphasizes these very issues.

Perhaps the most delightful surprise in *Introduction to Statistics and Data Analysis* occurs at the end of most chapters. It is a section entitled "Interpreting the Results of Statistical Analyses." In each case, the authors give guidelines on what to look for in published data. These are straightforward suggestions that enable students to become more astute readers of statistical information. We see, for example, in the section following confidence intervals, that there are different formats for published confidence intervals. We may see the actual interval, the estimate ± bound on error, or estimate ± standard error. Then we are advised to look for the confidence level associated with the given interval and to ask ourselves if this confidence level is reasonable. Finally, we should wonder if the interval width is appropriate. Has the population characteristic been estimated precisely? Then, the authors present a confidence interval from an actual article and discuss each of the issues raised above. This format is followed in each of the sections called "Interpreting Results of Statistical Analyses." First, we are told what to look for in published data. Then we are presented with an actual study and the authors apply the guidelines in interpreting what is shown. These sections not only provide review and closure for the material just presented, they also show the student how to communicate statistical information. This ability to communicate is critically important in AP Statistics and these end-of-chapter sections effectively present examples of good communication.

As the authors state, *Introduction to Statistics and Data Analysis* is designed with a particular eye toward the syllabus of the Advanced Placement Statistics course and the needs of high school teachers and students. They have succeeded magnificently. This textbook is a wonderful primary text, needing little or no supplementation, or at least a must-have reference text.

Susan Mead
Highland Park High School
Dallas, TX
*susan.mead@worldnet.att.net*

*Statistics in the Classroom...*

# A Statistical Poster for Elementary School: What is the Worth of a Name?

The American Statistics Poster Competition affords students in grades k-12 the opportunity to apply and effect their statistics knowledge through graphs. This article describes an example of the development of a statistics poster for elementary grades.

One type of source of data is through an experiment. Another is via a survey. The activity of interest here is an experiment in which students practice simple addition as well as display their resulting data in two different types of graphs, bar graphs and scatterplots. The title of a statistics poster should be a question that the graphs of the poster answer pictorially with little written description. Suppose that questions of interest for us are how much are our first and last names worth, how does the worth compare to the length of our names, and how do our names compare to each other?

Let's begin by constructing a table of worth (in dollars) per letter of the alphabet. Through a web search for frequency of letters usage of the English language, it is found that E (13.1%) is the most often used letter followed by T (9.0), O (8.2), A (7.8), N (7.3), I (6.8), R (6.6), S (6.5), H (5.9), D (4.1), L (3.6), C (2.9), U (2.8), M (2.6), P (2.2), Y (1.5), W (1.5), G (1.4), B (1.3), V (1.0), K (.4), X (.3), J (.2), Q (.1), Z (.1). Using this distribution, one somewhat arbitrary assignment of dollars to each letter follows.

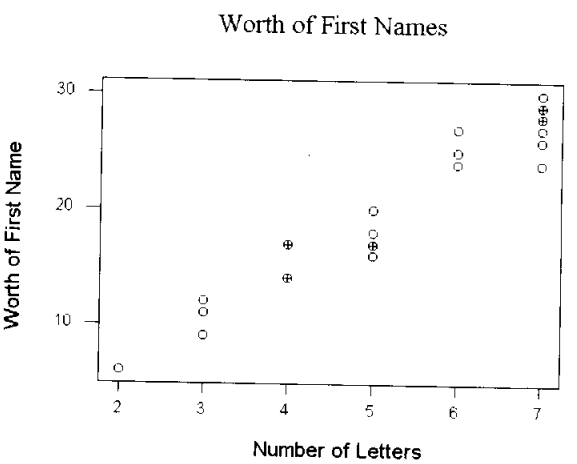
| Letters | Worth ($) |
|---|---|
| E, T, O, A | 5 |
| N, I, R, S, H | 4 |
| D, L, C, F, U, M | 3 |
| P, Y, W, G, B, V | 2 |
| K, X, J, Q, Z | 1 |

### *Data Collection/Graph*

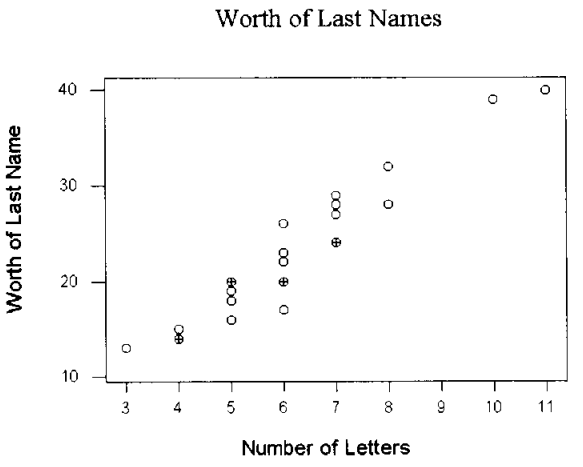
Each student calculates the worth of his or her first name and second name, separately. For example, the author's first name is Jerry and is worth $1 + 5 + 4 + 4 + 2 = $16. His last name is Moreno, worth $3 + 5 + 4 + 5 + 4 + 5=$26.

The first part of the construction is to consider first names only. So that students see the graphs that they will be using in the poster, have them construct the graphs on the board as a class exercise. Each student is to write his or her name on a post-it along with the number of letters in the name and its worth according to the above table. The student then affixes his or her post-it appropriately to a scatterplot on the front board that has Number of Letters labeling the horizontal axis and Worth ($) labeling the vertical. Here's an example for a class of 24 students and the corresponding scatterplot for first names.

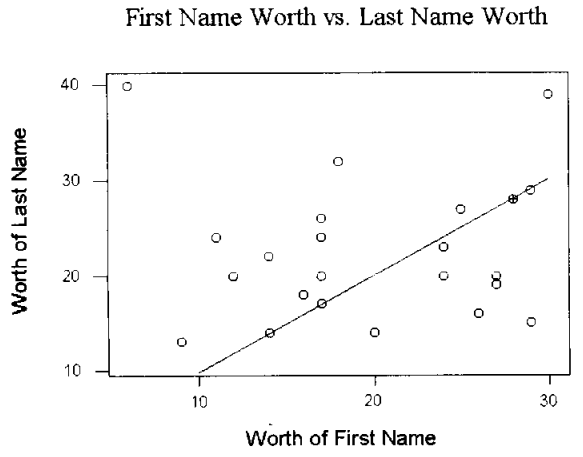
| | | | | | |
|---|---|---|---|---|---|
| Lloyd | 16 | Taryn | 20 | Salmoni | 28 |
| Agnew | 18 | Hill | 14 | Merconi | 28 |
| Nathan | 27 | John | 14 | Dana | 17 |
| Allen | 20 | Howell | 22 | Phelps | 20 |
| Nancie | 25 | Tiffany | 26 | Madison | 28 |
| Campton | 27 | Jewel | 16 | Pulsifer | 28 |
| Patrick | 24 | Leah | 17 | MaryAnn | 27 |
| Cullut | 20 | Joaglin | 24 | Shook | 19 |
| Cameron | 29 | Bob | 9 | Vanessa | 29 |
| Ford | 15 | Lee | 13 | Smetler | 29 |
| Meagan | 24 | TJ | 6 | Tim | 12 |
| Handey | 23 | Lightwaters | 40 | Smith | 20 |
| Holly | 17 | Barry | 17 | Tahmina | 30 |
| Hannah | 26 | Ludwig | 17 | Washington | 39 |
| Lenny | 18 | Joe | 11 | John | 14 |
| Hastings | 32 | Mazzaro | 24 | Wurt | 14 |




The **Worth of First Names** scatterplot will be one of the graphs for the statistics poster. Another will be similar except for last names. Note that if two students have identical data in terms of number of letters in their name and worth, then the graph shows this by a circle with a plus sign inside.

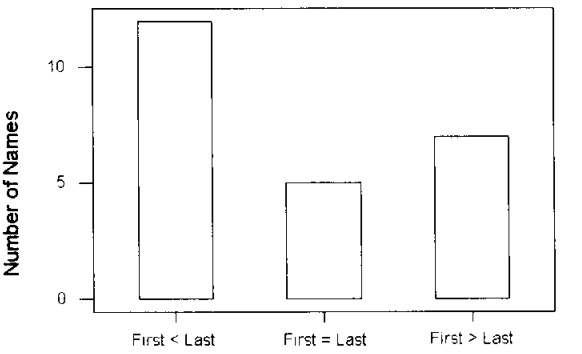



Another scatterplot that could be used is to compare the worth of first names directly with the worth of last names.



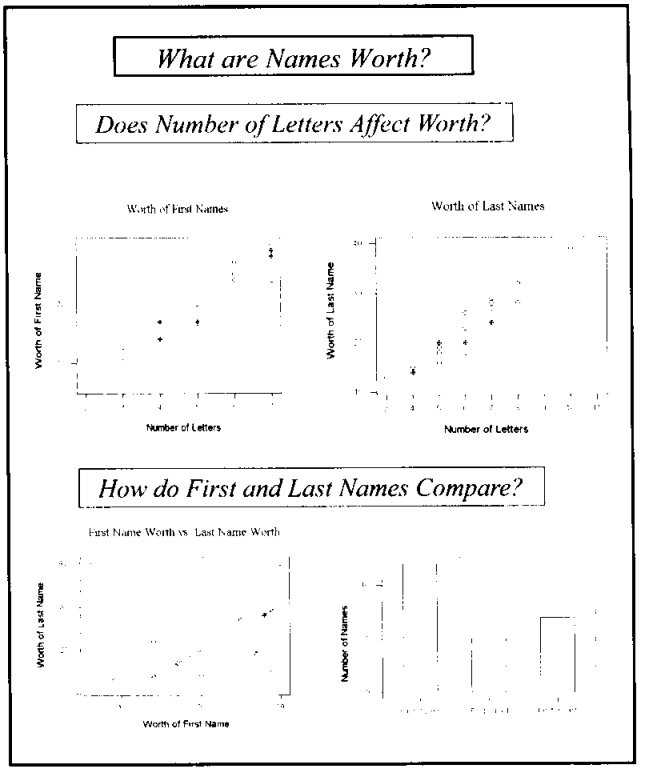

Several questions may be asked in class in drawing conclusions from these scatterplots. In particular, in the **First Name Worth vs. Last Name Worth** graph, draw the 45 degree line on the plot and discuss the meaning of points that lie on the line, above the line, below the line. At least one outlier is apparent in TJ Lightwaters who not only has the lowest worth for first name but owns the largest worth for last name as well. Added information may be found in boxplots plotted along the horizontal for worth of first names alone, and along the vertical for worth of last names alone. Note that TJ would not be an outlier in either of those variables taken alone.

If another type of graph is desired, one choice would be a bar graph that indicates the number of students whose first names were worth less than their last names, those that were of equal worth, and those whose first names were worth more than their last names.




Although there is not a well-defined definition of a statistics poster, it is generally thought of as a visual display containing two or more related graphics (plot, charts, maps, etc.) that summarize a set of data, that look at the data from different points of view, and that answer some specific questions about the data. Regarding the American Statistics Poster Competition, the above definition is excepted for the grades k-3 category in that only one graph is required of them. Assuming that students in upper elementary are working on this poster, how could they create a poster with the graphs of this article?

One suggestion follows. By the way, there is one glaring improvement that should be made on the poster. Do you see it?



What improvements could be made in the poster shown? A glaring one is that the first two graphs have different scales yet are measuring similar variables. In order to compare the graphs, the scales should be identical. What else could be done to improve this poster?

The author extends special thanks to his colleague, Carl Spitznagel, who constructed the graphs and poster.

Jerry Moreno
John Carroll University
*moreno@jcu.edu*

*Announcement*

## *ICOTS VI*

### *International Conference on Teaching Statistics 6 (ICOTS 6)*
### *July 7-12, 2002*
### *Durban, South Africa*

This conference is co-sponsored by the International Association for Statistical Education (IASE), the International Statistical Institute (ISI), and the South African Statistical Association (SASA). The major aim of ICOTS is to provide the opportunity for people from around the world who are involved in statistics education to exchange ideas and experiences, to discuss the latest development in teaching statistics and to expand their network of statistical educators. The conference will include keynote speakers, invited speakers, contributed papers, workshops and forums, demonstration lessons, roundtable sessions, poster sessions, book and software displays, hands-on computer sessions and many opportunities for the communication and exchange of experiences and ideas. As the conference theme for ICOTS-6 is "Developing a statistically literate society," special sessions on statistics literacy are also being planned. For more information see the conference Web site at http://www.beeri.org.il/icots6 or contact, Maria-Gabriella Ottaviani, Chair, International Program Committee, mariagabriella.ottaviani@uniroma1.it

Carol Joyce Blumberg
Department of Mathematics & Statistics
Winona State University
Winona MN 55987-5838
Office Phone/Voice Mail: (507) 457-5589
Fax: (507) 457-5376
email: *cblumberg@winona.edu*