August 2016 • Issue #470

AMSTATNEWS

The Membership Magazine of the American Statistical Association • *http://magazine.amstat.org*

2016 POSTER AND PROJECT Competition Winners

ALSO:

Criteria for Open Data Portals Offer Best Practices for Websites

ASA Volunteers Create Resource for Teachers: Statistical Commons

Statistics your way

Intuitive.

- Easy-to-learn, predictable command syntax
- Point-and-click menu access for any feature

Powerful.

- Thousands of built-in statistical tools
- Extensive data management features for survival, longitudinal, imputed, time-series, and multilevel data, and much more
- Programmable, so you can add your own custom features

Complete.

- Not sold in modules; everything in one package
- Comprehensive documentation with examples
- Free technical support

From GLM and survival analysis to multilevel models and Bayesian analysis, Stata has you covered.

Learn more online:
stata.com/amstat-ipc

STATA® 14 release

AMSTATNEWS

AUGUST 2016 • ISSUE #470

Amstat News welcomes news items and letters from readers on matters of interest to the association and the profession. Address correspondence to Managing Editor, *Amstat News,* American Statistical Association, 732 North Washington Street, Alexandria VA 22314-1943 USA, or email *amstat@amstat.org*. Items must be received by the first day of the preceding month to ensure appearance in the next issue (for example, June 1 for the July issue). Material can be sent as a Microsoft Word document, PDF, or within an email. Articles will be edited for space. Accompanying artwork will be accepted in graphics file formats only (.jpg, etc.), minimum 300 dpi. No material in WordPerfect will be accepted.

Amstat News (ISSN 0163-9617) is published monthly by the American Statistical Association, 732 North Washington Street, Alexandria VA 22314-1943 USA. **Periodicals postage paid** at Alexandria, Virginia, and additional mailing offices. POSTMASTER: Send address changes to *Amstat News,* 732 North Washington Street, Alexandria VA 22314-1943 USA. Send Canadian address changes to APC, PO Box 503, RPO West Beaver Creek, Rich Hill, ON L4B 4R6. Annual subscriptions are $50 per year for nonmembers. *Amstat News* is the member publication of the ASA. For annual membership rates, see *www.amstat.org/join* or contact ASA Member Services at (888) 231-3473.

American Statistical Association
732 North Washington Street
Alexandria, VA 22314–1943 USA
(703) 684–1221
ASA GENERAL: *asainfo@amstat.org*
ADDRESS CHANGES: *addresschange@amstat.org*
AMSTAT EDITORIAL: *amstat@amstat.org*
ADVERTISING: *advertise@amstat.org*
WEBSITE: *http://magazine.amstat.org*

The American Statistical Association is the world's largest community of statisticians. The ASA supports excellence in the development, application, and dissemination of statistical science through meetings, publications, membership services, education, accreditation, and advocacy. Our members serve in industry, government, and academia in more than 90 countries, advancing research and promoting sound statistical practice to inform public policy and improve human welfare.

features

3 President's Corner

6 Criteria for Open Data Portals Offer Best Practices for Websites

11 ASA Volunteers Create Resource for Teachers: Statistical Commons

12 Benefits of ASA Section Membership

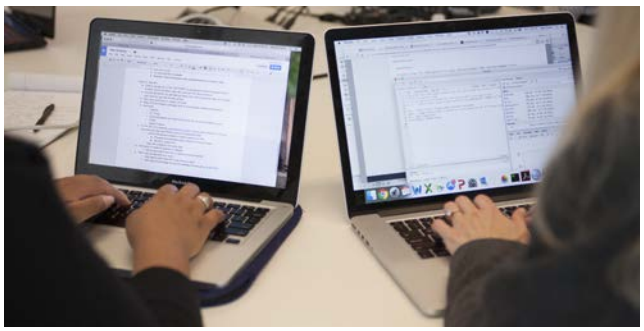
14 ASA Leaders Reminisce: Ronald L. Wasserstein

17 Workshop Focuses on Reproducibility

18 ESRA Award Winner Talks Physics and Statistics

20 Kelly Zou: Mathematics, Statistics, Data Science, and Dreams

24 *Technometrics* Highlights: Big Data Featured in August Special Issue



See **Page 6** to read about the work students in the Data Science for the Public Good Program have done with the Arlington County, Virginia, government on their open data portal.

columns

26 **MASTER'S NOTEBOOK**
Lessons from 34 Years as a Master's Statistician in Industry

This column is written for statisticians with master's degrees and highlights areas of employment that will benefit statisticians at the master's level. Comments and suggestions should be sent to Megan Murphy, *Amstat News* managing editor, at *megan@amstat.org*.

28 **STAT*tr@k***
Communicating Statistics to Nonstatisticians

STAT*tr@k* is a column in *Amstat News* and a website geared toward people who are in a statistics program, recently graduated from a statistics program, or recently entered the job world. To read more articles like this one, visit the website at *http://stattrak.amstat.org*. If you have suggestions for future articles, or would like to submit an article, please email Megan Murphy, *Amstat News* managing editor, at *megan@amstat.org*.

SIDEBAR

The 2002 Joint Statistical Meetings paper titled "People Count: The Social Construction of Statistics" (http://bit.ly/29OavUb) by Joel Best and feedback from Milo Schield gave Lawrence Lesser of The University of Texas at El Paso inspiration to repurpose the Beatles' "Lucy in the Sky with Diamonds." Here, he explores what it means to say statistics are socially constructed.

(THIS IS HOW) STATS ARE LIKE DIAMONDS

Picture statistics in the newspaper
With context all missing or half summarized;
Somebody chose just who would be surveyed
And the order the items arise.
Stats are each chosen and given a setting,
polished and cut for a view:
Look for the angle that reaches your eyes
when they're done.

CHORUS: This is how stats are like diamonds,
This is how stats are like diamonds,
This is how stats are like diamonds, ahhh……

Follow the chain of a long calculation
For students who view stats as math mechanized.
But when you tell them "all stats are constructed,"
Their eyebrows will rise up so high!
Newspaper numbers appear on page 4,
labelled and grouped for display:
Like a nice gem, this story now sparkled,
unlike rocks. (Repeat chorus)

Picture a stat with no error or bias,
Without a confounder that's lurking disguised.
Take care that you see how that stat is presented:
Most jewels have a flaw to apprise.
(Repeat chorus twice)

departments

31 **education**

See **Page 31** to read about the third biennial Electronic Conference on Teaching Statistics (eCOTS).

eCOTS 2016 Draws 600

2016 Poster and Project Competition Winners

Page 32

40 **awards & deadlines**
C. R. and Bhargavi Rao Prize

member news

41 People News

43 Section • Chapter • Committee News

44 Professional Opportunities

Follow us on Twitter
www.twitter.com/AmstatNews

Join the ASA Community
http://community.amstat.org/home

Like us on Facebook
www.facebook.com/AmstatNews

Follow us on Instagram
www.instagram.com/AmstatNews

How a Salsa-Dancing Former Monk from Argentina Is Instilling Confidence and Wonder in Her Statistics Students

For the past few years, I have participated as a mentor in the NSF-funded program Training a New Generation of Statistics Educators (TANGO Stat Ed), which matches university faculty with community college statistics teachers.

The TANGO principal investigators are Michael Posner of Villanova and Monica Dabos of College of the Canyons, a community college in California.

Monica Dabos

While discussing the project and much more with Dabos recently, I learned she has some thought-provoking ideas about fostering creativity, collaboration, and personal development in her students as they learn statistics, so I decided to interview her for this column.

To set the stage for my questions, quotes from two of her students follow. Although the first part of each quote is typical of comments received by any stellar statistics instructor, the second part illustrates the personal development Dabos fosters in her students.

> Before participating in this introduction to statistics course, statistics seemed boring and tedious, but now I view them as exciting, applicable, and maybe even the future solution to many of Earth's problems. … By learning how to be a global citizen, I now have an understanding of how I can impact the world with my time and my career, even if it is just a little part. By "daring to disagree," I know that challenge and criticism are beneficial, and that I should embrace them with an open heart and accepting attitude. This class has certainly made me a better, more intelligent, and more prepared individual for life and my future career as a college student.—*Kevin Joslin*

Jessica Utts

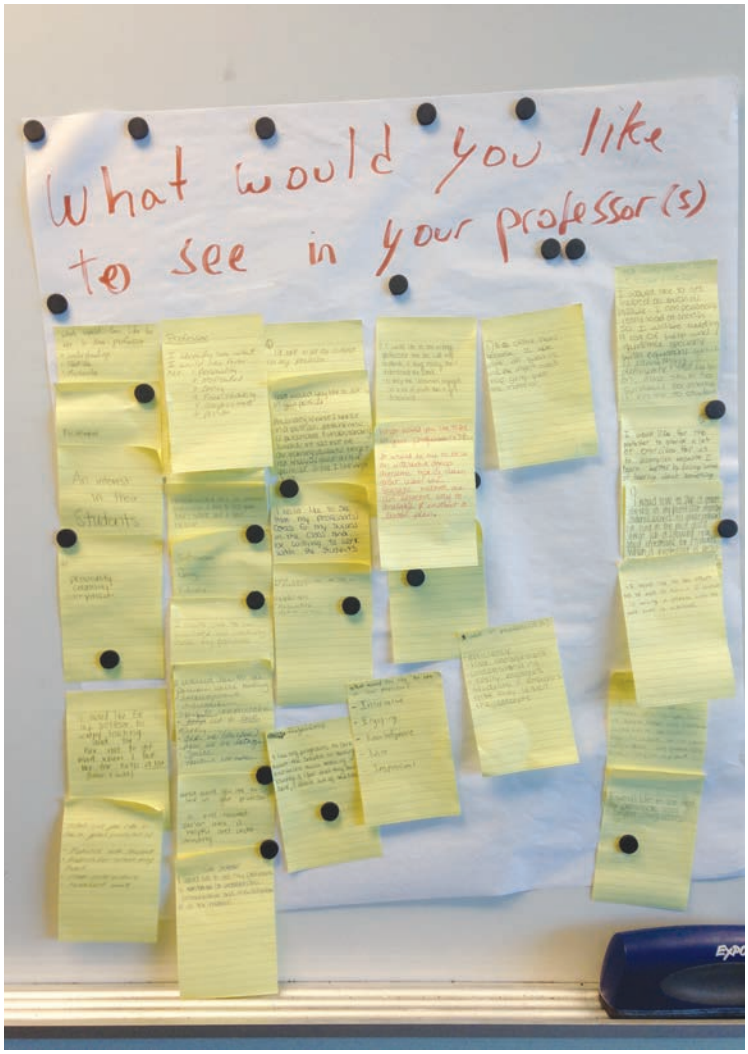
> As this semester draws to a close, I feel I have learned more than the basics of statistics. I have learned that sometimes the unknown can be motivating and empowering. … What I initially thought would be a semester of frustrating confusion has turned out to be one of confidence-boosting personal growth. … Dr. Dabos' unconventional teaching style put me in an uncomfortable space on day one, but out of that space grew confidence and wonder.
> —*Tanya Hauer*

Q: In reading quotes from your students, it's evident that you are interested in helping them grow as individuals as well as becoming better students. Is there something in your background that led to this more holistic approach to teaching?

This is a long story, but one of the aspects I think affects my teaching the most is my prior experience as a renounced monk focusing on inner transformation and meditation. I travelled to India several times, lived in temples, and practiced several types of meditation for more than eight hours every day for 15 years. The experience showed me the value of meditation and introspection and gave me the desire to do service to the community and society at large.

I studied education because of my love for teaching. While I started school with English as a second language at the age of 32, I have earned five degrees, culminating with a PhD in education that focuses on the teaching and learning of statistics. While in school, I went through several hardships that did not stop me from succeeding, but they were certainly blocks. As such, I tend to relate to students' struggles.

Dancing was always part of my life—before, during, and after life in the temple—but it was not until



On the first day of class, Monica Dabos asks her students to fill out yellow sticky notes and post them anonymously on big posterboards. They address six questions, one of which is "What would you like to see in your professor?"

I moved to Santa Barbara, California, that I found an affinity with ballroom dancing, in particular salsa. I was compelled to do something for the community, especially the young people. I saw around me that the only way young people were enjoying dance was in alcohol-fueled environments, and there was nowhere for those under 21 to enjoy dancing. Committed to making a difference, I started an alcohol-free salsa-dancing community in Santa Barbara that has grown to more than 500 members within three years of opening, and it continues growing with a diverse dance community for all ages.

Q: You have had an interesting journey! How does that influence your teaching?
I think my biggest contributions to students are my enthusiasm for statistics, the understanding that we are all human, and the benefits of balance in life and in the classroom. I believe I understand my students, and that this empathy is recognized right away; the students' appreciation is reflected in their commitment to excel. My teaching style stems from a small, yet universal set of guiding principles:

- Desiring to maintain self-balance
- Making room in my life for what I find valuable
- Progressively becoming more introspective in my teaching
- Recognizing when something does not work, and being willing to change
- Recognizing students' real struggles—economically, emotionally, ethnically, or physically

Q: In video interviews I saw with your students, they clearly felt a camaraderie with each other. Can you explain what you do to get students comfortable participating in the class?
Student bonding and group work: On the first day of class, students are asked to fill out yellow sticky notes and post them (anonymously) on big poster boards. They are asked to address the following six questions:

- What would you like to see in your professor(s)?
- What behavior do you expect from your classmates?
- What have I learned about myself as a student?
- What worries me about this class?
- What excites me about this class?
- Other things I want to say

After everyone is finished, the class breaks into groups and they summarize what their classmates had to say. Here is one student's reflection on this activity:

> The activity in class was also enlightening. It was great that you not only asked what the students expected from each other, but also what hopes they had about their professor. We were able to let you and the other students know our thoughts. Also we could see that we all expect the same things from each other. It was a great first class and I look forward to the rest of the semester. — *Sahar Wardak*

We continue to have group activities throughout the semester, but we change groups every day in class (at random) and have "get to know your classmate" time.

Here is an anecdote that illustrates how close the students come to be and how it benefits them. The mother of one of my students died in his arms during the semester. He missed two weeks of class when we were learning hypothesis testing, plus students started to work on final projects. He came back on a review day for an exam. We stopped the class to greet him with hugs and tears; it was a sad day for all of us, but then we got to work. All the students told him that they would help him with the material he had missed and met with him for several hours after class. He took the final exam and took part in the final project and presentation. He got an A on the final exam.

TED Talks: I assign a TED Talk every two weeks during class time or for homework or both. After watching a TED Talk for 18 minutes and reflecting on it for 10 minutes, a half-hour of valuable class time is used, but the next two hours of lessons are so focused and engaged that we cover a lot of material with much greater depth, participation, and appreciation. Here is an example of student feedback on the TED Talks:

> One of the very first TED Talks that we were assigned to watch was Hans Rosling's "The Joy of Statistics." First of all, I had never seen a TED Talk before, and I imagined that it would be quite painful to sit through one. Surprisingly, I felt this talk was great because it made me realize how important statistics really is. Statistics is not just meaningless proportions and averages. Statistics have the ability to tell us a "story" about a population.—*Cristal Munoz*

We also watch TED Talks like "How Not to Be Ignorant of the World" and "Why Smart Statistics Are the Key to Fighting Crime." I make sure that my students are inspired by something big, and then they can do something small in the class like a poster presentation related to cereal and shelf placements.

Fun, real, hands-on activities: Especially in the beginning, we do activities in class that are related to their lives and that are fun. In one such activity, we discuss whether men or women text faster on their phone. Then students have to find a way to measure this. They have to brainstorm in groups all the possible variables that may affect the result of the text speed. They categorize each variable as numerical or categorical, and they then decide whether we can control some of those variables and we conduct a classroom experiment. We talk about if the results would be true for all men or women. That is normally a fun and applicable activity early on in the semester. They are used to teachers saying "No phone in class!" So I say, "Take your phones out; let's do an experiment."

I believe that I understand my students, and that this empathy is recognized right away; the students' appreciation is reflected in their commitment to excel.

Q: I know you are a strong supporter of the ASA's *Guidelines for Assessment and Instruction in Statistics Education* (*GAISE*). How do you think teaching according to the *GAISE* guidelines has changed the students' experience of taking a statistics class, and/or their view of statistics?

Making it real, more applicable, fewer computations. More hands-on experiences help students connect with statistics and think critically about what they are doing. When students know it is real, they see the value, the applicability, and they want to do it. I remember once I asked my students to correct a survey that was used by one of the departments at Santa Barbara City College for collecting data. I told them to comment on what changes the survey needed for improvement and why. Students worked in groups. I told them this was for real, that this survey is used for the GATEWAY program, and they will take their comments seriously as they would like to improve their survey. Students were on it like a mission; they took it so seriously—real application has great appeal to them.

Q: The TANGO Stat Ed project helps connect community college teachers with university professors and seems to be a success. Do you have any recommendations for college professors who might want to interact with community college teachers, or vice versa?

The most important themes across my teaching, my community involvement, and my research projects like TANGO are that people work better when they are not in isolation, when they are not afraid to ask questions, and when they feel free to express themselves because they know their voice counts. No one likes to feel minimized for not understanding or knowing, but when the right environment is created, everyone appreciates the opportunity to help one another grow and concepts of collaboration and respect become tangible results.

Criteria for Open Data Portals Offer Best Practices for Websites

Mark Almanza, Virginia Tech; Madison Arnsbarger, Virginia Tech; Jessica Flynn, Cornell University; Adrienne Rogers, Virginia Tech; Will Sandholtz, University of California Berkeley; Emily Stark, Austin Peay State University; and Iowa State University graduate students Millicent Grant and Samantha Tyner

"Open Data," which is any publicly available information, has become increasingly popular since the 1990s. It has been especially relevant with President Obama's emphasis on the importance of open data for "transparency, participation, and collaboration" (*http://bit.ly/1dLtUiU*). Open data is providing new opportunities for statisticians to engage across science and policy.

Kristanne Littlefield, Chief Digital Officer for Arlington County in Virginia, recently created an open data portal and asked students in the Data Science for the Public Good Program (DSPG) at the Social and Decision Analytics Laboratory (SDAL), a part of the Biocomplexity Institute of Virginia Tech, for feedback. The DSPG team consulted a variety of open data portals, tech websites, and journal articles about effective website design and use to create a comprehensive of a set of criteria extendable to all open data portals. These criteria suggest a number of best practices for user interface and data structure, which any locality creating and maintaining an open data portal should consider. The criteria are divided into two broad categories with the subcategories listed in parentheses:

MORE ONLINE
To read more about the authors and DSPG program, visit *http://bit.ly/298sSV9.*

1. User Interface (organization, navigation, and design)
2. Data (tidiness, accessibility, documentation, and visualization)

Both categories are evaluated on the basis of accessibility to a broad audience, and applicability to a wide range of uses, including statistical analysis.

User Interface

Organization

An open data portal should be structured so users' knowledge or experience level does not limit their use of the site or other resources provided. Essentially, the site should provide users with context so they can intuitively and easily navigate the open data portal (see *http://bit.ly/29AKFVv*). For instance, an open data portal that requires several clicks before reaching a data set could be frustrating and disorienting for the user, especially if it is difficult to return to the homepage. The concept of consistency also applies to file organization. Most websites order lists of files alphabetically; however, the option to sort by date, file type, or other criteria may be included to give users more control over their search.

Navigation

An easily navigable open data portal will encourage better and m ore frequent use of the data. Navigation within an open data portal can be thought of as a GPS in the physical world; the user should either know where they are or know how to return to familiar territory. To achieve this, buttons and other features should be clearly labeled and their functions made predictable. For instance, if the user expects a button to perform a certain action, the result should match the user's expectations. While some features may become easier to use with experience, the open data portal should accommodate the amateur user to appeal to all audiences. A functional and navigable site should also have a "Home" button displayed on each page. This button, like the others, should be clearly and explicitly labeled.

Design

An open data portal's design should reflect its purpose to disseminate information. A simple palette of high-contrast colors should make up the bulk of the site, leaving many colors available to use as accents. Additionally, the site's color palette should remain consistent throughout the site to aid in navigation and appear more unified. An open data portal should be easy to read, with text in a font and color that are easy to look at for an extended time. Black text on a white background is the easiest to read. A designer should think carefully before choosing another option, and should consider the experience of colorblind users.

A well-designed homepage acts as a simple jumping-off point for users, directing them to the data they are seeking. For example, an open data portal with a strong homepage—like San Francisco's at *https://data.sfgov.org* (see Page 8)—displays data categories, a search function, and other useful tools such as the ability to search by government department. Also, the use of accent colors and graphics to differentiate the types of data available can make the site more usable and improve a user's ability to locate information. Legends that assign a unique graphic and color to each type of data help users identify the files they need quickly and easily.

Data Science for the Public Good (DSPG)

The first six authors—**Mark Almanza, Madison Arnsbarger, Jessica Flynn, Emily Stark, Adrienne Rogers**, and **Will Sandholtz—**are undergraduate students and part of the ASA NSF Research Experience for Undergraduate program. The final two authors are graduate students in statistics at Iowa State University. The students are part of the Virginia Tech Data Science for the Public Good program.

The Social and Decision Analytics Laboratory (SDAL), part of the Biocomplexity Institute of Virginia Tech, established the Data Science for the Public Good (DSPG) program to connect aspiring data science scholars to communities that can benefit from their expertise. Students selected through a competitive application process are engaged in a series of hands-on learning experiences while policy makers and government leaders receive data analysis support to inform difficult decisions related to health care, education, and social justice.

The story of each community—its problems, needs, and aspirations—is contained within its data. The DSPG program's over-arching objective is to equip new generations of scientists with the skills they need to bring this story to light for leaders in local government.

The DSPG program teaches student fellows how to sift through vast amounts of information related to public safety, employment, and the provision of services to discover how communities can become more efficient and sustainable. Through the lenses of statistics, social science, and data science research, DSPG students learn to integrate all available data resources.

DSPG fellows also have opportunities to diversify their expertise and form lasting professional connections by taking part in the Social and Decision Analytics Laboratory's data-driven research projects. Research teams are horizontally integrated—combining disciplines such as statistics, data science, and the social and behavioral sciences to address complex issues—and vertically integrated—allowing students to collaborate with project stakeholders at all levels.

These unique fellowship experiences are made possible through the support of several research organizations dedicated to serving the public good: American Statistical Association's NSF Research Experience for Undergraduates (REU); Virginia Tech's Global Forum for Urban and Regional Resilience (GFURR); and sponsored research.

Data

Tidiness

Data tidiness refers to the format and structure of the data, rather than the content. To illustrate, two data tables may contain exactly the same information, but one may be in a format that is tidy while the other is not. Tidy data simplifies the process of profiling and cleaning the data and is essential to exploratory data analysis and visualization. In the *Journal of Statistical Software*, Hadley Wickham offers the following three simple criteria that define tidy data:

1. Each variable forms a column
2. Each observation forms a row
3. Each type of observational unit forms a table

The first two criteria are complementary. In tidy data tables, each row represents a distinct "object," and each column should correspond to some characteristic of that object. Each unique object is an "observation" and each characteristic a "variable." Table 1 is in a tidy format. Each row represents a different person (observation), and the columns record various characteristics (variables) of each person such as date of birth, sex, and height. Table 1 is for illustration only, and is not an example from any existing open data portal.

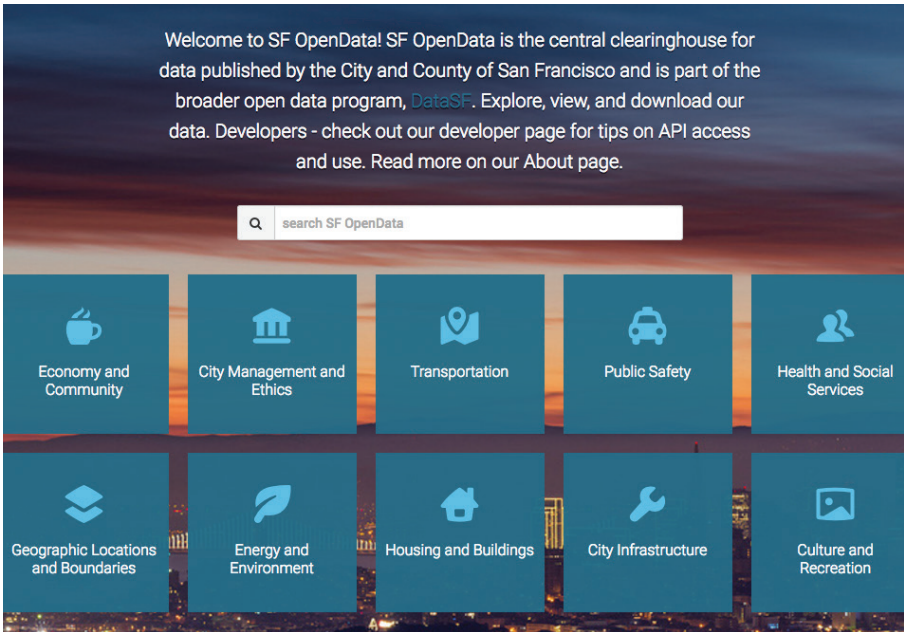
Table 1—Data in a Tidy Format

Name	Sex	Date of Birth	Height (in)
Person 1	F	03/07/1968	66
Person 2	M	11/05/1971	70
Person 3	M	07/17/1964	68
Person 4	F	10/22/1970	67

Most tidiness errors pertain to the table's variables. Wickham summarizes three common errors as the following:

- Column headers are values, not variable names
- Multiple variables are stored in one column
- Variables are stored in both rows and columns

In some messy tables, the data values form the columns—one column for males and one for females—



San Francisco's open data portal has a strong homepage that directs users toward the data they seek and offers useful tools, such as a search function.

MORE ONLINE
These experiences are made possible through the support of the ASA's NSF Research Experience for Undergraduates (REU) program. More about the REU program can be found at *http://bit.ly/29tyaKl*.

rather than having a single column with the variable name as the column header. A tidy format should have one column for the variable, sex, and contain the corresponding value, male or female, in each row. Another common error is to store more than one variable in a single column. In this case, the column headers may convey two values of two variables. For example, if a column contains the range of a set of values, the column should be split into two columns, one for the minimum value and one for the maximum value. Last, variables should never be stored in both rows and columns. Typically, if this occurs, then the top left cell of the table will be empty, though this will not capture all instances of this error.

While it is open for debate whether the responsibility of tidying data is that of the open data portal administrator or the user, tidy data portals may attract more users than portals that require users to tidy the data themselves. Depending on the goal and maturity of the organization supporting the portal, making the data tidy could facilitate better and more frequent use of the content.

Accessibility

Data accessibility is a broad term encompassing the logistics and qualifications of calling a data portal "open." To assess the openness or accessibility of the data sets, it is important to ask the following three questions:

- Can the user access the data?
- What types of data are available?
- Are the available data usable?

Data availability can be hindered in many ways. First, for the data portal to be truly open, there must be zero access restrictions and data must be license free. It is important that users have minimal obstacles to accessing the data such as membership requirements, user fees, and data download limits.

Just as restrictions and licenses hinder the openness of data, poor organization and labeling of data sets may hamper full access as well. Data set titles should answer who, what, when, where, and why. For example, the title "Average Housing Costs" answers what information the data set contains, but the title "County Average Housing Costs 2000–2010" also describes who, when, where, and why with only 15 additional characters. In addition to an explanatory title, the most successful open data portals provide brief descriptions of each data set both on the search page and on the data set specific webpage. This information should describe the source of the data set and explain what information the data set contains.

While it is important and exciting to make data public, it is important to ensure the safety of all constituents when uploading data to an open portal. A major safety concern is personally identifiable information (PIIs). In general, it is important to avoid sharing information such as name, social security number, biometric record, or any information that could be used to identify a specific person. PIIs are determined on a case-by-case basis, but inappropriate sharing can lead to legal repercussions. Be sure to understand the localities' regulations regarding PIIs before sharing data on an open data portal. In general, a unique identifier can be used to anonymize the personal data.

Last, the ability to share data is an essential part of data accessibility. For companies and research groups to use data for innovative apps and studies, the appropriate resources must be sharable. The code for both the website and data sets should be publicly available and an application program interface (API) created. It is important to note that for APIs to be used, the data set must be in a tidy format with appropriate dates and addresses.

Undergraduate student Emily Stark, front right, from Austin Peay State University, types notes during a tutorial on web scraping by Brian Goode at Virginia Tech's National Capital Region location in Arlington, Virginia.

Documentation

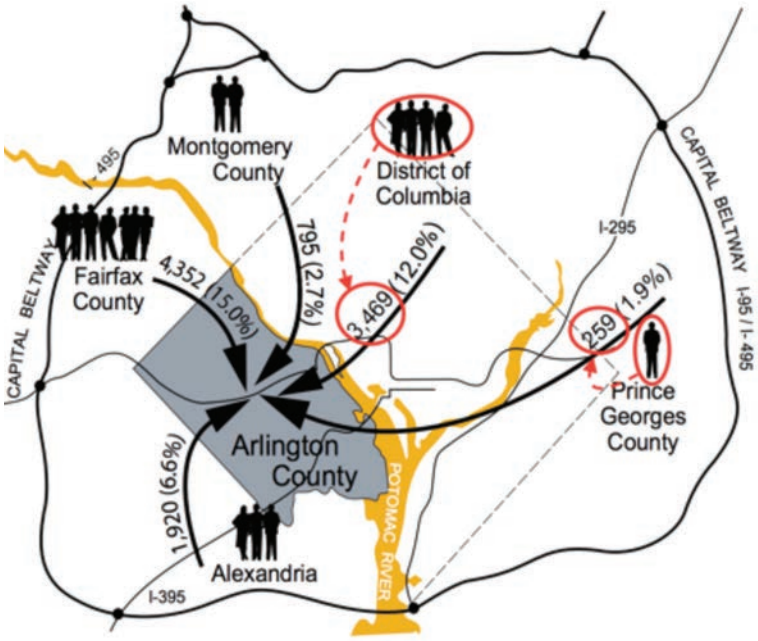
Proper documentation communicates the credibility of an open data portal by answering questions such as where the data came from and how they were collected. Displaying a list of commissions, agencies, and other government organizations that have made their data publicly available on the site may enhance both the attractiveness and credibility of the portal. In addition, the description of each data set should contain information about the origin of the data.

A well-documented open data portal also provides information about the metadata, or data about the data. A comprehensive data dictionary defining variables from every available data set is an extremely valuable feature. A dictionary that is readily accessible and searchable gives users the ability to find variables contained across many data sets and allows them to compile and cross-reference that information. A feature that provides an explanation of the variable by scrolling over the column header helps facilitate a better understanding of the data set. In addition to defining individual variables, it is also valuable to provide a description of the entire data set.

A final aspect of data documentation is how the data are constructed. File type (e.g., csv, html) is a critical component due to software preference and availability and should be indicated clearly. File names should also be informative, and data sets should contain the raw, unaltered data in a tidy format. If summary statistics such as averages and rates are contained in the data, an explicit description of how those values were calculated should be provided. Last, missing values should be left blank and not represented by characters or expressions. This improves users' ability to analyze the data, and to understand what information is available to them.

Visualization

Data visualization refers to any presentation of data that is not in table format, such as maps, plots, and charts. An open data portal may incorporate data visualization in various ways. First, portal administrators may provide visualizations to users. These visualizations are typically complex and tell a specific



This image has added "chartjunk," extra elements that crowd the view and obscure the data. Here, the "junk" are the people added to the visualization.

Further Reading

GovEx. 2016. Open data portal requirements. *http://bit.ly/29tTNK9*

Planning Division of the Department of Community Planning, Housing and Development. 2016. Arlington County PROFILE 2016. *http://bit.ly/29v9lK2*

Quora. 2016. What are some successful outcomes from hackathons? *http://bit.ly/29C7LIr*

Sunlight Foundation. 2016. Open data policy guidelines. *http://bit.ly/29tU53L*

Tufte, Edward. 2001. *The visual display of quantitative information, 2nd Ed.* Graphics Press: Columbia, MD.

Wickham, Hadley. 2014. Tidy data. *Journal of Statistical Software* 59(10).

story about the data that the administrators want the public to see. An open data portal also can provide users the tools to create their own visualizations. Tools like these help users explore the data and ultimately make the open data portal more useful. These tools should be clearly labeled, easy to navigate, and not hidden or several clicks away.

Additionally, user experience will be streamlined if the available tools match the level of measurement of the variable. For example, stock price that is measured at the ratio scale can be displayed using line graphs and scatter plots, whereas bar charts are used to display nominal and ordinal scale data.

Finally, the open data portal may provide links to visualizations created by third parties, who create them by accessing the data through the portal and posting them to separate sites. This can increase the use of a portal by attracting people from third party sites. Many data portals only incorporate the first two methods of visualization, while the most sophisticated portals offer all three.

All methods of visualization on an open data portal should follow the same basic guidelines. All visualizations should be easy to interpret with figure titles, axis labels, and legends in legible font sizes. In addition, visualizations should not distort the data view in a way that misrepresents the data. Common mistakes to avoid include starting the y-axis at the minimum data value instead of at zero or another reasonable value and ignoring standard conventions (e.g., reversing order of axes).

Visualizations should be aesthetically appealing and avoid the use of "chartjunk," a term coined by Edward Tufte to mean extra visual elements that crowd the view and obscure the data. The "junk" items in the image (left) are the people icons added to the graph. The icons are used to represent the proportional flow of people migrating into Arlington County from surrounding areas that are also represented numerically in the same figure. Although the number of people coming from DC is more than 13 times the number of people coming from Prince George's County (3,469 versus 259), there is one person icon for Prince George's County and only four for DC, rather than the 13+ indicated by the actual numbers. This can mislead the viewer.

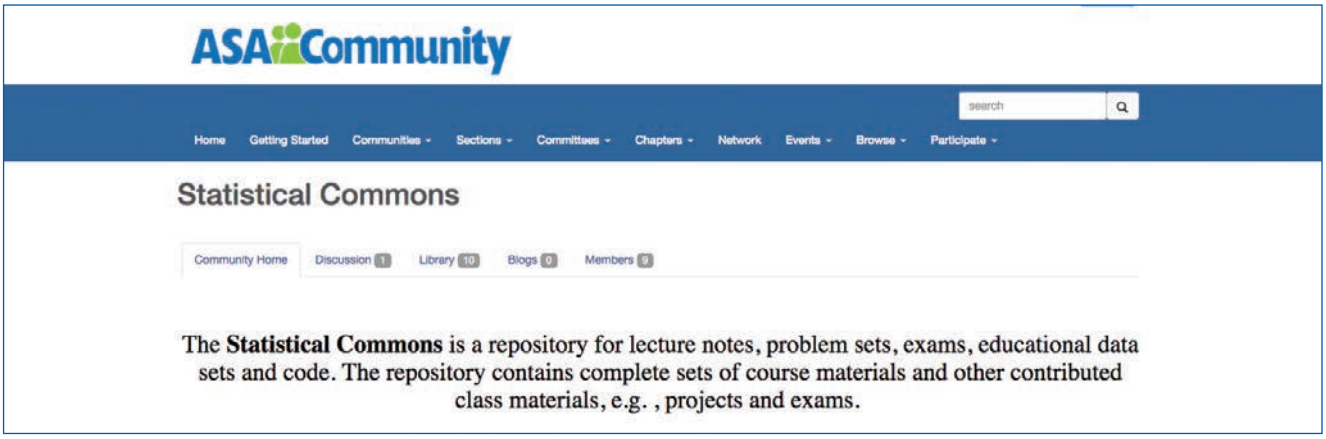
Two ways to represent the data accurately are to remove the people icons and make the numbers more prominent or include a legend indicating how many people each person icon represents with the appropriate number of icons in each region.

Good data visualizations may contain many elements, but adding unnecessary content to a figure does not aid in the interpretation of the data and may be misleading. Effective data visualizations contribute to the overall understanding and utility of the data displayed.

While these criteria cover many topics, they are not exhaustive. Currently, most open data portals excel in just a few of the areas outlined above. Site administrators should seek continuous improvement and evolution over time to enhance the user experience. Users are also encouraged to hone their curiosity toward productive applications of open data. Commitment to open data on behalf of site administrators and users alike guarantees a promising and innovative future. ■

ASA Volunteers Create Resource for Teachers: **Statistical Commons**

David Banks, Duke University



The ASA's Statistical Commons allows teachers to share quizzes, lectures, exams, and data sets online.

Many of us teach. We know the joy of encountering eager and curious minds, and then drilling them on the proper use of a z-table, the memorization of sum of squares formulae, and the rote recitation of the definition of a *p*-value.

But not all aspects of teaching give such satisfaction. Personally, I find it tedious to continually construct exams and quizzes that are fresh and fun. There is an unremitting search to find interesting data sets that achieve specific pedagogic purposes. And I know many new teachers value access to prewritten and battle-tested lecture notes.

To address these problems, and to create new value for members of the ASA, a group of has worked to create a website that allows teachers to share lectures, exams, quizzes, data sets, and so forth.

The website is hosted on the Members Only area of the ASA platform, and can be accessed only by people who pay the regular (nonstudent) membership fee.

Currently, the site contains full course materials for the following:

- A noncalculus introduction to statistics, as developed by Mine Çetinkaya-Rundel
- A noncalculus introduction to biostatistics, as developed by Kari Lock Morgan
- A calculus-based introduction to statistics, as developed by David Banks

The hope is that others will upload similar materials, either full courses or perhaps just portions of courses, such as final exams or notes for especially engaging lectures. And there is no need to restrict this resource to introductory courses; it would be wonderful to add graduate material or anything else that would assist teachers and benefit students.

The site also contains qualifying exams and their solution keys. Most graduate programs require first-year students to pass some such exam, and many provide their students with older exams on which to practice. Since the set of interesting and discriminating questions is large but finite, it makes sense for departments to share these questions broadly and thus reduce the dreary springtime chore of concocting new problems *ab initio*.

This new website is called the Statistical Commons. It can be accessed at *http://bit.ly/29oZd6a*. Instructions for uploading new material are on the website. New material will be lightly reviewed and categorized by moderators before it becomes accessible. ■

Statistical Commons Volunteers

David Banks
Mine Çetinkaya-Rundel
Nick Horton
Donna LaLonde
Kari Lock Morgan
Bill Notz
Rebecca Nugent
Chris Paciorek
Victoria Stodden
Mark Ward

Benefits of ASA Section Membership

Council of Sections Governing Board

Professional growth and networking opportunities are enhanced by participating in any of the 28 ASA sections. Sections are communities developed to further the objectives of the association in a field of statistical methods, theory, or applications. With statistics being a broad and diverse field, a section provides a "home" for statisticians; offers opportunities throughout the year for professional development; and keeps members current through conferences, newsletters, and community discussions. An additional benefit of section membership is the opportunity to network within one's area of interest or expertise. Also, many sections offer mentoring opportunities, organize student competitions, and provide awards. All these benefits can enable career growth and future employment, but specifics of how section membership can benefit you include the following:

Keeping Up to Date

Sections play a critical role in keeping statisticians current throughout the year, and not just at the annual association's conferences. All sections host web communities so their members can stay in touch and share information, and many publish semiannual newsletters or journals that allow their membership to contribute to the development of their scientific interest. Some sections organize their own conferences focused on a sub-discipline to help disseminate new topics, promote learning, and foster collaboration. In addition, certain sections also have their own publications.

The ASA sections are key contributors, helping to support a wide scientific program for the annual Joint Statistical Meetings and other section-sponsored conferences based on section interests. Sections organize invited, topic-contributed, and contributed sessions, as well as poster sessions and roundtable discussions. Additionally, a few sections host a luncheon at the annual meeting with an eminent speaker in their field. Through all these activities, sections ensure their members stay current in a rapidly changing field.

Examples of section-sponsored conferences include the following:

- Health Policy Statistics Section sponsors the International Conference on Health Policy Statistics
- Mental Health Statistics Section sponsors the Thomas R. Ten Have Symposium on Statistics in Mental Health
- Nonparametric Statistics Section co-sponsors the International Society for Non-Parametric Statistics Conference and the Workshop on Nonparametric Statistics for Big Data
- Statistics and the Environment Section co-sponsors the biennial ENVR/EnviBayes Workshop

Examples of section-affiliated journals include the *Journal of Business and Economic Statistics* and the *Journal of Nonparametric Statistics*. Also, the Section on Statistical Learning and Data Science has its own journal dedicated to applications and theory of data science—*Statistical Analysis and Data Mining*.

Education

Education is an important part of a section's activities, giving members economic options to gain new knowledge, stay current, and grow professionally. Many sections sponsor continuing education courses at annual conferences and webinars throughout the year to facilitate professional development in their area of interest. Members of ASA sections are sometimes called upon to support efforts toward establishing national education guidelines or providing commentary on topics of importance (e.g., Statistical Education Section's recent article, "Barriers and Threats for Statistics in Secondary Education"). In another example of such collaboration, several statistics societies and Springer collaborated to establish a new online encyclopedia. Also, sections encourage their membership to contribute to topic-relevant definitions and content.

Examples of recent webinars include "Enhancing Reproducibility in Statistical Analysis Using R Markdown," sponsored by the Section for Statistical Programmers and Analysts, and "Mobile Phone-Based Data Collection and Analysis," sponsored by the Mental Health Statistics Section.

Also, sections such as the Biopharmaceutical Section conduct and archive podcasts (see *http://bit.ly/29HDl7W*). And the Section on the Teaching of Statistics in the Health Sciences (TSHS) promotes sharing teaching materials through the TSHS Teaching Resources Portal at *www.causeweb.org/tshs*.

Leadership Opportunities

Section members have the opportunity to gain leadership and organizational experience by serving on sections' executive boards. These boards contain elected and appointed positions in which members can help shape the section's value to its members. It is a stepping stone for those both early in their careers and those who want to gain more visibility in the ASA.

Examples of leadership roles include section chairs, who define and lead the section directive, and program chairs, who help identify educational opportunities for their members through section

webinars and annual conferences such as JSM and the Conference on Statistical Practice.

Advocacy

Section members are able to contribute to discussions about funding, policy issues, and industry regulations through section-organized activities that enable members to have an impact beyond their current roles and to shape our society for the better. For example, members of the Health Policy Statistics Section interacted directly with a patient-centered outcomes research institute (PCORI) representative at the Joint Statistical Meetings in Chicago to learn about funding opportunities and other PCORI initiatives.

Another example is the Statistics in Defense and National Security Section, which promotes the statistical profession and best statistical practices within the defense and national security governmental and affiliated organizations on an ongoing basis.

Similarly, the Statistics and the Environment Section connects statisticians in academia to applied statisticians and environmental scientists working directly with regulations and policy to have an immediate and practical effect on environmental regulation and policy.

Networking

Sections help facilitate discussion among members from different backgrounds (e.g., industry, government, and academia) and among multiple societies with a shared purpose or interest. Section business meetings and mixers held at the Joint Statistical Meetings are an excellent way to meet those working in a similar field or sharing related interests. Small group settings such as these are an excellent place for discussion and networking. Several sections also produce newsletters covering section-relevant topics, maintain an electronic mailing list to facilitate communication, and advertise job openings to connect section members to employers.

A few examples of section-sponsored activities include the following:

- The FDA/Advanced Medical Devices and Diagnostics Statistical Issues Conference (sponsored by the Medical Devices and Diagnostics Section) brings together leading authorities from the FDA, industry, and academia.
- The Physical and Engineering Sciences Section sponsors the Marquardt Memorial Industrial Speakers Program to foster communication between industrial and academic statisticians.

The Biometrics Section cooperates with the Biometric Society, Statistics Section of the American Public Health Association, and Section U of the American Association for the Advancement of Science to exchange ideas and build bridges.

Member-Specific Initiatives

Members are recognized for exceptional scientific work and outstanding service through section-sponsored awards. Multiple sections provide major awards to recognize members doing remarkable work. Other sections convene an ASA Fellows nomination committee to assist deserving members with preparing nominations. Certain sections award members for exceptional mentoring efforts. Sections provide important career opportunities and enhanced professional visibility by enabling their members to disseminate their work or serve as a session chair or discussant at a conference or workshop. Many sections promote specific mentoring initiatives for early-career and junior statisticians. From time-to-time, members have the opportunity to contribute to additional initiatives, or even propose their own section-sponsored initiatives.

Examples of section-sponsored or co-sponsored awards include the following:

- Mitchell Prize and Savage Award, co-sponsored by the Bayesian Statistical Science Section
- Julius Shiskin Memorial Award for Economic Statistics, Nathan Mantel Award for Lifetime Contributions to Statistics in Epidemiology Award, and Young Investigator awards, sponsored by the Statistics in Epidemiology Section
- Jeanne E. Griffith Mentoring Award, sponsored by the Government Statistics Section
- David P. Byar Young Investigator Award and funding mechanisms for training such as postdoctoral career-development awards, sponsored by the Biometrics Section

An example of a section-sponsored initiative is the Statistical Consulting Section's initiative to develop online resources and tools helpful for statistical consultants.

Student Benefits

Many sections offer opportunities for their student membership. Multiple sections sponsor student paper competitions to encourage students to submit their best work and present at JSM or other meetings. Many of these competitions offer awards for travel costs. A few sections provide education grants (e.g., Survey Research Methods Section) or scholarships (e.g., Natrella scholarship by the Quality and Productivity Section). Additionally, there are initiatives to help and engage under-represented groups of students. Examples of student travel awards include those sponsored by the Bayesian Statistical Science Section and Biometrics Section. ■

HOW TO JOIN

For a full list of sections, see *www.amstat.org/sections/sectionlist.cfm*. ASA members can join online through the "ASA Members Only" tab at *www.amstat.org*, where one can "Add a Chapter or Section." Annual dues are quite small (ranging from $5 to $12), with the option to try out a new section for a year for free. Join an ASA section today!

ASA LEADERS REMINISCE

Ronald L. Wasserstein

Ronald (Ron) L. Wasserstein assumed the ASA's top staff leadership post in August 2007. In this role, he provides executive leadership and management for the association and is responsible for ensuring that the ASA fulfills its mission of promoting the practice and profession of statistics. He also is an official ASA spokesperson and responsible for a staff of 35 at the ASA's headquarters in Alexandria, Virginia.

Prior to joining the ASA, Wasserstein was a faculty member in the mathematics and statistics department and an administrator at Washburn University in Topeka, Kansas. During the last seven years of his 23-year tenure at Washburn, he served as the university's vice president for academic affairs.

Wasserstein has extensive experience applying statistics to real problems. He has served as a consultant to the Kansas Department of Revenue, Missouri Department of Revenue, and Illinois Department of Revenue. He also has served as an expert witness in several legal cases, and he was a member of the Governor's Tax Equity Task Force in Kansas in 1995. In 2012, he testified as an outside witness in a House of Representatives hearing on the FY13 Commerce, Justice, Science (CJS), and Related Agencies Appropriations Bill.

Wasserstein is a longtime member of the ASA, having joined the association in 1983, and has been active as an ASA volunteer for more than 20 years. He twice served as president of the Kansas-Western Missouri Chapter and served as chair of two ASA sections—the Statistical Education and Statistical Consulting sections. He also chaired the Council of Chapters Governing Board in 2006 and was a member of the ASA Board of Directors from 2001–2003.

Wasserstein is a Fellow of the ASA and American Association for the Advancement of Science. He was presented the John Ritchie Alumni Award and Muriel Clarke Student Life Award from Washburn University and the Manning Distinguished Service Award from the North American Association of Summer Schools. He also was president of the Kappa Mu Epsilon National Mathematics Honor Society from 2009–2013 and served on the MATHCOUNTS Board of Directors from 2007–2014.

Wasserstein and his wife, Sherry, live in northern Virginia and enjoy movies, live theater, books, and doting on their children and grandchildren.

James Cochran

In the 20th installment of the Amstat News *series of interviews with ASA presidents and executive directors, we feature a discussion with current ASA Executive Director Ronald L. Wasserstein.*

Q You are the only person I know of who has served as chair of both the ASA Section on Statistical Education and the ASA Section on Statistical Consulting. What similarities do you see in the missions of these two sections?

A My earliest involvement with the ASA was with what was then called the Subsection on Statistical Consulting Education. Back then, sections had subsections, and this was a subsection of the Statistics Education Section. Just a few years later, a constitutional revision of the ASA was undertaken and subsections became sections. By then, I was deeply tied into both the Statistical Education Section and the Statistical Consulting Section. The area of commonality for me is that a fair amount of time in consulting is spent educating the consultee in statistics. Of course, at least an equal amount of time in consulting involves statisticians getting educated by their consultees about their scientific domain.

Q You have written several times on lotteries for the *Huffington Post* and have discussed lotteries in the media several times—including an interview during which Melissa Lee stated she thought you were likely the first statistician CNBC had interviewed on air. What motivated you to initiate this series of blog posts, and what has been the reaction to them?

A I started thinking, teaching, and writing about lotteries back in the late 1980s, when Powerball came to Kansas. I had bright students telling me they were buying lottery tickets and winning more than they were spending. I started to look for ways to help people understand what's going on under the hood with lotteries, and I've been speaking on the topic ever since. I'm grateful to the *Huffington Post* for the opportunity to blog on their site about the lottery, as it has tremendous reach. I was motivated by all the talk about lotteries as Powerball reorganized itself to have larger jackpots by greatly reducing the probability of winning. As for the "Fast Money" interview, it was interesting and fun to work with that team to explain how buying more tickets indeed increased one's chances of winning, but not by enough to matter.

ASA Executive Director Ron Wasserstein attends the JSM 2014 opening mixer with Ingram Olkin and Sastry Pantula.

Q What has surprised you most in your role as a member of the steering committee for The World of Statistics?

A The World of Statistics is a follow up to the International Year of Statistics, or Statistics2013. I was surprised and delighted in 2013 to see the worldwide enthusiastic response to promoting the field of statistics through Statistics2013. Thousands of organizations and individuals got involved. It was one of the most encouraging things I've ever been involved with in my professional life.

Q The American Statistical Association created the "*This* is Statistics" website and campaign to increase awareness of statistics careers among students, parents, educators, and counselors. How has this campaign been received?

A I am proud of the ASA for its commitment to promoting the profession of statistics to young people. Through the "*This* is Statistics" national outreach program, the ASA is making high-school and college students aware of careers in statistics and letting them know that even if they don't want to be a statistician, understanding the basics of statistical thinking is important for navigating 21st century life. I'm proud of the ASA because it has made a substantial investment that does not directly benefit the association, but has great potential to impact the profession.

Q The ASA has been very active in trying to educate the public and the scientific community about the correct way to interpret and think about *p*-values. What effects do these efforts appear to be having?

A We are extremely pleased with the response to the *p*-values statement. The article in *The American Statistician* has been downloaded well over 120,000 times, and the subject is being discussed in forums across the country and, I presume, around the world. I'm excited that the ASA plans to further invest time and resources in some follow-up activities on this subject. Readers of *Amstat News* will be hearing more about this very soon, I hope.

Q You served for several years on the faculty at Washburn University. What course or courses did you most enjoy teaching and why?

A I loved teaching statistics at any level, but especially the intro level. I enjoyed being with students and seeing them grow, not just in my class,

Members of the Wasserstein family (from left: Ron, Peterson, Abner, Rose, and Sherry) enjoys a day at the tidal basin in Washington, DC, during the 2016 blooming of the cherry trees.

but as they progressed through the university. Even when I served in administration, I tried as often as possible to be teaching a course each semester. I do *not* miss grading, but I miss most everything else about teaching statistics.

Q Are you involved in any new statistics initiative that you find to be particularly exciting?

A One of the most exciting projects I'm currently involved with is the launching of the International Prize in Statistics. An award of international stature for our profession is long overdue and now, thanks to the cooperation of five major statistical societies—the ASA, International Biometric Society, International Statistical Institute, Institute of Mathematical Statistics, and Royal Statistical Society—we will be selecting and announcing the first award winner this October and presenting the award at the ISI World Statistics Congress next year. Details are at the website *statprize.org*.

Q You became executive director of the ASA in August 2007. What has been the most gratifying aspect of the nine years you have served as the ASA's top executive?

A I am not exaggerating when I say that every day, in some way, I'm reminded of what a wonderful opportunity and privilege it is to be the executive director of this organization. I'm gratified by the incredible volunteer commitment of our members. We estimate that more than 2,000 members volunteer for the ASA in some capacity each year. We couldn't successfully promote the practice and profession of statistics without them. I'm grateful for the leadership of ASA Board members, who work with great vision and diligence to advance the association—and do so collegially. ASA presidents have particularly been a pleasure to work with. These amazing leaders have been inspirational mentors to me, and I am thankful for each of them. I am a very fortunate person indeed!

I've enjoyed two wonderful careers, first at Washburn University, and now here at ASA. My favorite career, though, remains being a husband, father, and grandfather. I'm blessed with wonderful children who make me laugh and think. I'm grateful for our two newest family members, two boys adopted from Haiti, for all the excitement and joy they bring to the entire Wasserstein tribe. And, of course, none of these careers would have been possible without the loyal love of my wife of 40 years. Her capacity to care about others astounds me. Our lives have been quite a ride, but no matter how rough the road, she's always up for another trip. ■

Workshop Focuses on Reproducibility

Michelle Schwalbe, Director of the Committee on Applied and Theoretical Statistics of the National Academies of Sciences, Engineering, and Medicine, and Claire Ji Former Mirzayan, Fellow at the National Academies of Sciences, Engineering, and Medicine

The reproducibility of scientific research is a growing concern among the scientific community. A number of high-profile articles in scientific journals and the popular press have called attention to the issue, noting that difficulties in reproducing scientific results waste resources, slow scientific progress, and could erode the public's trust of science.

Among the many factors that contribute to poor rates of reproducibility, several are of particular interest to the statistical community, including insufficient training in experimental design, poor data management and analysis, and inappropriate understanding and interpretation of statistical concepts. The impact of these factors varies among scientific disciplines, because each field has a different research culture and a unique set of needs.

A February 2015 workshop—organized by the Committee on Applied and Theoretical Statistics (CATS) of the National Academies of Sciences, Engineering, and Medicine and sponsored by the National Science Foundation (NSF)—focused on three reproducibility issues from a statistical perspective: the extent of reproducibility; the causes of reproducibility failures; and the potential remedies for these failures.

To address these issues, workshop discussions considered, among other items, the definitions of reproducibility and associated terms, how to improve our understanding of scientific discovery through pursuing reproducibility, re-evaluating the threshold for statistical significance, enhancing and clarifying protocols, facilitating the sharing of statistical tools, and enhancing education and training.

In addition to probing the statistical aspects of reproducibility, the workshop also brought to light a broad array of efforts to address the overall problem. In particular, the computing research community, data-sharing clouds, and open-source software packages have been developed in an attempt to improve reproducibility and sharing. Companies and organizations such as the Science Exchange for biomedical sciences and the Center for Open Sciences for social sciences are establishing mechanisms to improve reproducibility by starting peer review of study designs and methodologies as early as possible in the research process, even before data are collected. Funding agencies such as the National Institutes of Health and NSF are piloting new ways of doing grant review to promote transparency, and they are considering funding opportunities for replication studies.

A summary report from the 2015 workshop, *Statistical Challenges in Assessing and Fostering the Reproducibility of Scientific Results*, is available for free download at *http://bit.ly/1pl2D0r*. Visit the workshop webpage at *http://bit.ly/1FOBoQC* for videos and other materials from the workshop.

For more information about this workshop and other CATS activities, contact Michelle Schwalbe at *mschwalbe@nas.edu*. ■

ESRA Award Winner Talks Physics and Statistics

Morteza Marzjarani

The Excellence in Statistical Reporting Award (ESRA) Committee selected Natalie Wolchover as the honoree of its 2016 award. Wolchover is a senior writer at Quanta Magazine*, covering the physical sciences. Previously, she wrote for* LiveScience, Popular Science, Seed, *and* Make *magazines. She has a bachelor's in physics from Tufts University, studied graduate-level physics at the University of California at Berkeley, and co-authored several academic papers in nonlinear optics. The ESRA committee was especially impressed with the number and breadth of Wolchover's articles related to the discipline of statistics.*

Tell us about yourself and what professional experience you have had so far.

As a writer for *Quanta*, I cover research at the bleeding edge of physics and occasionally dip my toes in mathematics, computer science, and other fields. (Statistics plays into all of these, of course.) I love my job because it allows me to spend my days learning. I was born in London, England, and grew up going back and forth between there and the podunk town of Blanco, Texas. From a young age, I was shaped by those opposites, while also balancing seemingly opposing loves of science and literature. I planned to be a physicist; I studied physics at Tufts and started in on a PhD at Berkeley. But, eventually my writerly side protested and I left the program to pursue science writing. I started up a blog and began freelancing while also tutoring college physics students [so I could] earn extra cash. My clips helped me land an internship in New York, which led to my first staff job, which led to my second. I now live with my wife and two cats in Brooklyn. I'm enjoying city life, but—always seeking a balance of opposites—also longing for that big Texas sky.

Do statistics enter into your physics reporting?

Yes, all the time. Beyond just explaining sigma values and that kind of thing, I try to get across that the laws of physics are, fundamentally, an attempt to describe statistics. Both at the quantum scale of particles and the cosmic scale of the universe, everything we know about nature is in the form of statistics. On the one hand, there are the statistics of how particles morph and scatter in collisions, and on the other hand, [there are] the statistical distribution of galaxies and other cosmic structures. These are sets of data points in no particular order.

The Standard Model of particle physics doesn't have an explanation for why, when two protons collide at the Large Hadron Collider, the scattering process produces two photons one time and a Higgs boson the next. The order of these scattering processes doesn't matter—and in fact quantum mechanics says the order is impossible to know or predict. All you can do is collide particles over and over again and tally up the number of times you get photons versus the number of times you get Higgs bosons. The Standard Model is essentially a theory of these statistical likelihoods.

Similarly, cosmologists can't tell you why two superclusters are found a certain distance apart while two others are some other distance apart. They just gather statistics on the correlations between structures and then develop a theory of the universe's birth that accounts for these statistics. So it's really interesting that our fundamental theories are that way. And there's a deep issue where, in order to have a complete theory of particle physics and cosmology (and ultimately these have to converge), we would actually need infinite statistics. This might suggest we need a new approach.

Your award-winning article concerned the Tracy-Widom distribution. Can you give a teaser?

Tracy-Widom is a statistical distribution that crops up frequently throughout math and physics—as the distribution of largest eigenvalues of random matrices and variations along the interface of a liquid crystal, for instance. It looks like a lopsided, left-leaning

Natalie Wolchover won the 2016 ESRA award for an article concerning the Tracy-Widom distribution.

Gaussian distribution. Like the bell curve, the Tracy-Widom distribution also exhibits "universality," emerging as the common macroscopic behavior of diverse microscopic systems, but in its case, researchers are only beginning to understand why it is so universal and what the underlying cause is.

My article focused on some fascinating work in that direction by the statistical physicist Satya Majumdar and colleagues. By studying the extreme tails of the distribution, they found it bears striking similarities to a third-order phase transition—and as readers will learn, that's a major clue as to why it arises.

One wonderful thing about reporting the article was that I had the opportunity to speak to Leo Kadanoff, the great statistical physicist who, decades ago, coined the term "universality" in his study of phase transitions and who sadly died in October. He told me he had always suspected there was a connection between the universality of phase transitions and universality in random matrix theory, and this seems to be bearing out. The puzzle isn't solved yet, but I definitely plan to stay on top of the story.

Do you plan to write more about statistics?

I'm writing a statistics-related article right now, in fact! I keep coming back to this theme of emergent universality in many-body systems. It's a fascinating and mysterious topic, one that transcends boundaries between disciplines. I, like everybody, love becoming aware of connections between seemingly disparate phenomena, and that's what universality is all about. It is a hard topic to write about, though, because when a certain pattern or law is universal, meaning it manifests itself in the behavior of many disparate systems, it can be a struggle to give a sense of that breadth while also tying all the threads together and telling a cohesive story. I'm in the thick of that struggle at the moment.

What do you think about ESRA?

I'm thrilled to have won this award and very grateful to the American Statistical Association for their recognition. The prospect of winning awards is, of course, not what motivates me to write—I write because I love to learn and think about the world, and to articulate what I'm learning and thinking about—but it's nonetheless incredibly gratifying to receive such positive feedback. I love that the award comes with a traveling cup. All my friends are insisting we have to do crazy things with it, Stanley Cup-style, but don't worry—I plan to keep it on my desk at work and look to it for inspiration in times of writer's block!

What are your interests other than writing?

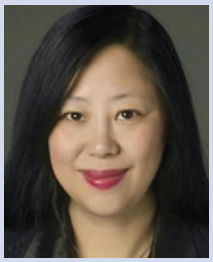
Above all, reading—I love nothing better than a good book or *New Yorker* article. My partner is an art historian and we see a lot of art together. I love hanging out with family, friends, my beloved cats, animals in general. Also, playing squash, cooking, traveling. Taking walks in [the] woods and swimming in rivers. It would be disingenuous not to also mention Netflix. ■

BIOPHARMACEUTICAL COMMUNITY LEADERSHIP SERIES

Kelly Zou: Mathematics, Statistics, Data Science, and Dreams

Alex Dmitrienko

This is the second interview in the leadership series launched by the Biopharmaceutical Section. The first interview was with Gregory Enas (see http://bit.ly/29ElBtY). The series includes interviews with statistical leaders among the biopharmaceutical community.



Kelly H. Zou, a PStat and ASA Fellow, is senior director and analytic science lead of real-world data and analytics at Pfizer Inc. She earned both her MA and PhD degrees in statistics from the University of Rochester and completed her postdoctoral fellowship training at Harvard Medical School. Her undergraduate major was mathematics, along with a minor in physics. Previously, she was associate professor at Harvard Medical School, associate director at Barclays Capital, and senior director and statistics lead at Pfizer Inc. She has authored more than 130 articles and several statistical monographs.

Kelly has been on several professional journal and magazine editorial boards and is the incoming chair-elect and past secretary of the Health Policy Statistics Section; chair and past vice chair of the Statistical Partnerships among Academe, Industry, and Government Committee; and chair of the Committee of Presidents of Statistical Societies Awards Committee.

On Leadership

What is your particular leadership style?

I entered an election for the first time when I was in High School Affiliated to Fudan University in Shanghai, China. Our homeroom teacher and a wonderful mentor, Mr. Dehong Wang, encouraged the students to volunteer by taking charge of student activities. There were multiple roles to be filled, and I entered the race to become a league representative. I discussed with my boarding-school roommates, wrote a speech with solid suggestions about how to improve our organization and connectivity among students, and gave a speech in front of all my classmates. Then the ballots poured in …

Probability became an extremely useful tool that my classmates can still recall tallying up and counting the ballots one-by-one on a blackboard. It was a gut-wrenching moment between the two fierce candidates. After being declared the winner, I helped other class officers on events such as celebrating the New Year, other national holidays, and traditional cultural events such as the Moon Festival.

Fast-forward to 2016. My most recent election was for the incoming chair-elect position of the Health Policy Statistics Section of the ASA. I am very grateful that our ASA members have put their faith in my hands by casting their votes of confidence. Several goals have entered my plan for our section members, which are how to encourage professional statistical activities, improve networking opportunities, enhance communication skills, and provide influences in the policy arena using sound statistical methodology and applications in health care and related fields.

I would like to bring up a useful article (see *http://on.inc.com/1mCcIgW*) that has echoed well in my experience on the topic of statistical leadership. It summarizes the following nine important traits that define leadership:

- Awareness
- Decisiveness
- Empathy
- Accountability
- Confidence

- Optimism
- Honesty
- Focus
- Inspiration

In my view, an effective leader exhibits a combination of all these qualities. Although these traits would embody a universally admired figurehead, you get the general picture that this leader is an inspiring visionary who is also compassionate and responsible toward the betterment of our society.

I wish for there to be visionaries with determination, innovation, and compliance in the statistical profession and quantitative analytics community, in general. My favorite author, Lu Xun (1881–1936), once wrote, "Hope can be neither affirmed nor denied. Hope is like a path in the countryside: originally there was no path—yet, as people are walking all the time in the same spot, a way appears."

A key obstacle for those who are technically focused is that excellent workers may not easily evolve into visionary leaders. Many of us may wonder why it is not so easy to break and then surpass this artificial ceiling by thinking strategically. For example, it can be a dilemma how to juggle and balance multiple tasks in work and in life. The personalities on a team may be varied, and a leader must understand the purpose of a complex task, the main paths taken to achieve the success, and to understand the needs and limitations of the team members.

In turn, a great leader can function as a shepherd, who guides passengers to form a critical mass. Faced with employees with diverse goals, interests, skill sets, and backgrounds, an effective leader takes charge, but collaborates effectively. In my company, for example, the "own it" mentality is highly encouraged for all leaders and employees alike.

I have been fortunate enough to thrive along two quite different career tracks. For about 10 years, I progressed from postdoctoral fellow to associate professor at Harvard. For nearly just as long, I have been working in the industry as first associate director at Barclays and then senior director at Pfizer. Although these are well-known academic and private institutions, my experience has been with these selective few. While at Harvard, I was a member of the faculty taskforce the Joint Committee on the Status of Women (JCSW). The JCSW discussed extensively, particularly in terms of the glass ceiling for female employees and faculty.

I tend to be an extrovert in terms of my personality, except when I am reading a protocol, designing a statistical analysis plan, or writing a manuscript. I generally solicit opinions well in advance, but I also give colleagues or committee members time to discuss on a smaller scale within sub-teams before finalizing the overall task or project.

Since I have been exposed to multinational and multifunctional settings, I also see things from different angles and cultural perspectives. Everyone has a different and innate purpose and mission in life, and everyone must come up with and come to terms with his or her own definition of happiness. As long as life is fulfilling, it will make one's life content.

On Aspirations and Goals

What are your views on the pursuit of the American dream?

There may not be a universally accepted magical equation that will solve all career-related puzzles, dilemmas, and challenges. In fact, there is hardly a common panacea to guarantee the attainment of happiness if everyone has his or her own version or definition of happiness. A relatively young Steve Jobs once famously asked the vice president at Pepsi-Cola, John Sculley, "Do you want to sell sugared water for the rest of your life? Or do you want to come with me and change the world?" For some diligent workers, selling sugar water is precisely the heart of their pride and joy, and there is nothing wrong with that, either.

As an extrovert and people-oriented person, I get integrated into the environment and culture fairly quickly and seamlessly. I can easily engage in a meaningful chat with new acquaintances through common aspirations and interests. In America, which is a multicultural melting pot, I have not only showed acceptance and gratitude, but also respect and curiosity. I am also realistic and tend to look forward.

I like to set reachable goals and identify the best quality of each collaborator. For example, when planning for a major project, a grant application, or a manuscript, I pay attention to the guidelines, instructions, authorship arrangement, and deadlines. Thus, others will know what to expect of the entire team well in advance. Then, I hope to motivate and spark the creativity in each of my coauthors or teammates, so that the project glues the entire team together. Efficiency is as important as perfection, and I always strive for utmost excellence and efficiency.

MORE ONLINE
This summer, the biopharmaceutical blog will feature several prominent statisticians, including Erik Pulkstenis and Christy Chuang-Stein. The blog can be viewed at *http://bit.ly/1U9SrDP*.

Over the years, my coauthors and I have been grateful for the various opportunities that we have had by winning multiple poster awards from the Biopharmaceutical Section and the Scientific and Public Affairs Advisory Committee of the ASA. The coauthors not only developed and published new methodology, but also learned how to present complex statistical work and analyses in visual and succinct ways. After all, solving new problems is a fun way of extending a perpetual learner and scholar to a technically savvy statistician or data scientist in the biomedical field.

However, having aspirations by pursuing the "American dream" may mean different things for different people. Some of us may aim to be leaders, some may aspire to be excellent applied statisticians and data scientists, and some may find theoretical work highly rewarding.

In the Asian culture in which I grew up, for example, we wish for our family members to be completely free of any major illness or catastrophe. To achieve such a balance, we may choose to maintain "the doctrine of the mean" and value harmony, rather than constantly riding the ebb and flow of unexpected tides. Along one's career journey, a bamboo ceiling may be encountered.

To be successful, one must first become an expert in a particular subject matter or a specific topic area, without being easily replaceable due to the lack of knowledge. Another suggestion is to acquire the useful skill of delegating tasks and not being hands-on all of the time. If an assignment is not within one's comfort zone, then it is imperative to examine whether the task adds value and, if so, learn more about it and grow from the learning experience.

In reality, however, implicit biases may exist, but must be eliminated. For example, there are the so-called glass ceiling and bamboo ceiling. According to the U.S. Department of Labor, a glass ceiling is "the unseen, yet unbreachable, barrier that keeps minorities and women from rising to the upper rungs of the corporate ladder, regardless of their qualifications or achievements" (see *http://bit.ly/29yijFE*). The same report states, "Breaking the glass ceiling is an economic priority that this nation can no longer afford to ignore. It is an economic imperative driven by recent dramatic shifts in three areas that are fundamental to business success: (1) changes in the demographics of the labor force; (2) changes in the demographics of national consumer markets; and (3) the rapid globalization of the marketplace."

On Mentorship

What do you think about the idea that a good leader is, first and foremost, a good mentor? When is it appropriate to take a mentorship mentality with a leadership role, and when not?

The key characteristics of an effective mentor are listening and problem solving skills. Nowadays, soft skills can enhance technical prowess. These are all crucial in business acumen and in leadership qualities in the field of statistics, data science, and quantitative analytics.

The past chair of the ASA SPAIG Committee and incoming ASA president, Dr. Barry D. Nussbaum, has encouraged our committee to foster active networking and partners across different career sectors. He spoke highly of the impact of mentorship among seasoned and upcoming ASA members. One of the most endearing mentors to many of us was Professor Ingram Olkin (see *http://bit.ly/2a9faO0*).

Well, mentorship has influenced my career path in a profound way. One day when I was about 21 years old, a mathematics professor made a comment, "Kelly, you seem to enjoy probability so much. Have you thought of studying statistics?" Until that moment, I had never thought of statistics outside the framework of mathematical statistics. I was intrigued and excited. At that point, an entire world of probability density functions unfolded themselves in front of my eyes and mind. I dived into the library to devour professors Norman L. Johnson, Samuel Kotz, and N. Balakrishnan's continuous distributions. They tied so well into my undergraduate mathematics research on Fourier series and differential equations. If I had not met my mentor, I would certainly not have had the exposure to the beauty of statistics in the realm of mathematics, other than regarding it as one of the courses I took.

For those of you who feel the same passion toward mathematical statistics, please read professor Edward Frenkel's book, *Love and Math: The Heart of Hidden Reality*, which won the 2015 Euler Prize from the Mathematical Association of America.

Last year, to help with young statisticians, I signed up as a mentor during the last Joint Statistical Meetings (JSM), met with my designated JSM mentee on the night of the Opening Mixer, and attended the JSM Mentoring Workshop. There was great rapport between us as mentor and mentee, respectively, and we discussed her career goals and current

statistical pursuits throughout the mixer. By being a mentor outside the immediate scope of work-related projects, it is more rewarding to listen attentively and provide useful suggestions to the mentee.

My company constantly encourages talent development through mentorship paring. In fact, there is a mentorship-matching process in place within our company, which may be utilized by any colleague. I have been quite active in participating in the mentorship process, organized by the Global Asian Alliance (GAA) within my company. A thoughtful and caring GAA mentor from another part of the company became a guiding light. Subsequently, through a friendly "on-boarding buddy" process for a newly hired data scientist, I became a mentor through the formal GAA-sponsored mentor-mentee pairing process and training. I was quite honored to serve as a GAA mentorship panelist.

Generally, mentors and mentees are encouraged to meet periodically, such as once a month. An important component of mentorship is that mentees should think of the goals they hope to achieve and how mentors could help accordingly. It may be useful to jot down a short list of topics on which the mentees may have to seek mentors' input. In our company, besides technical and applied skills, individual development plans become a blueprint for short-term and fairly long-term aspirations. It may be beneficial to share in confidence some of these goals with the mentor. Furthermore, mentees broaden their horizons by seeking out different mentors with skill sets, technical and interpersonal, that the mentees may discuss and draw inspiration from.

An important component of mentorship is that mentees should think of the goals they hope to achieve and how mentors could help. ... It may be useful to jot down a short list of topics on which the mentees may have to seek mentors' input.

On Life Outside of Work

What do you do with your free time? What are your hobbies? Do any of your hobbies or activities give you a unique perspective for how you conduct yourself at work?

While I was a high-school student, I was a television show producer and a director of the "You and I High School Students" series, which was developed by and broadcasted from the Shanghai TV Station. In addition, I am an oil-painting enthusiast, particularly impressionistic artwork.

Of all interests outside work, I can say that I am a classical music enthusiast. My most favorite composer is Giacomo Puccini, and my most admired tenor, now a baritone, is Plácido Domingo.

I am an avid movie buff ever since I was invited to be on a movie critiquing team during my junior-high-school years, and I especially love watching cartoon movies and TV series via computer animation. If I hadn't studied math as a major and physics as a minor before pursuing my doctoral education in statistics, I would probably have gone into the multimedia, broadcasting, film, theater, or television business. From time to time, quoting humorously and metaphorically from movie scenes or stage plays becomes a staple of my conversation style.

I recall that my language teacher asked me to join the movie critiquing team when I was a junior-high-school student, and I wrote two pieces a week after watching two movies. I was a playwright who directed several theater plays on stage, as well as comedies.

Of course, there is always a "math geek" lurking. Tracing back historically and through my own PhD adviser and mentor, professor W. Jackson Hall, I have literally become one of the academic descendants or offspring of one of the greatest German mathematicians of the 19th century, Carl Friedrich Gauss (1777–1855).

For example, music notes make me think of white clouds floating in the air, full of Greek symbols. When music and statistics intersect, the creativity is enjoyable! I admire the professor band "The Imposteriors" (*http://bit.ly/2a9fw7i*).

My lifelong goals as a mathematics-statistics-data science whiz exist across space and time, transcending boundaries and limits. Beyond statistics, I have always hoped for many layers, facets, and dimensions within one's lifetime. Like Brownian motion, little unexpected events would bring out so many facets and colors. One life full of adventures can still be sufficiently satisfying.

In closing, when I have a little free time away from data and deadlines, I am still a dreamer … ■

TECHNOMETRICS HIGHLIGHTS

Big Data Featured in August Special Issue

Peihua Qiu, *Technometrics* Editor

The August 2016 issue of *Technometrics* is a special issue on Big Data analysis. A total of 11 papers are included, covering a wide range of topics in describing, analyzing, and computing Big Data.

The first five papers propose numerical algorithms that can analyze Big Data fast. In "Orthogonalizing EM: A Design-Based Least Squares Algorithm" by **Shifeng Xiong, Bin Dai, Jared Huling,** and **Peter Z. G. Qian**, an efficient iterative algorithm intended for various least squares problems, based on a design of experiments perspective, is proposed. The algorithm, called orthogonalizing EM (OEM), works for ordinary least squares and can be extended easily to penalized least squares. The main idea of the procedure is to orthogonalize a design matrix by adding new rows and then solve the original problem by embedding the augmented design in a missing data framework.

In "Speeding Up Neighborhood Search in Local Gaussian Process Prediction" by **Robert B. Gramacy** and **Benjamin Haaland**, the authors suggested an algorithm for speeding up neighborhood search in local Gaussian process prediction that is commonly used in various nonlinear and nonparametric prediction problems, particularly when deployed as emulators for computer experiments.

"A Bootstrap Metropolis-Hastings Algorithm for Bayesian Analysis of Big Data" by **Faming Liang, Jinsu Kim,** and **Qifan Song** proposes a so-called bootstrap Metropolis-Hastings (BMH) algorithm that provides a general framework to tame powerful MCMC methods for Big Data analysis. The major idea of the algorithm is to replace the full data log-likelihood by a Monte Carlo average of the log-likelihoods calculated in parallel from multiple bootstrap samples.

"Compressing an Ensemble with Statistical Models: An Algorithm for Global 3D Spatio-Temporal Temperature" by **Stefano Castruccio** and **Marc G. Genton** suggests an algorithm for compressing 3D spatio-temporal temperature using a statistics-based approach that explicitly accounts for the space-time dependence of the data.

"Partitioning a Large Simulation As It Runs" by **Kary Myers, Earl Lawrence, Michael Fugate, Claire McKay Bowen, Lawrence Ticknor, Jon Woodring, Joanne Wendelberger**, and **Jim Ahrens** is about analysis of data streams in which data are generated sequentially and data storage, transferring, and analysis are all challenging. The authors suggest a so-called online *in situ* method for identifying a reduced set of time steps of the data and data analysis results to save in the storage facility to significantly reduce the data transfer and storage requirements.

The next two papers concern machine learning methods for handling Big Data. "High-Performance Kernel Machines with Implicit Distributed Optimization and Randomization" by **Vikas Sindhwani** and **Haim Avron** proposes a framework for massive-scale training of kernel-based statistical models, based on combining distributed convex optimization with randomization techniques.

"Statistical Learning of Neuronal Functional Connectivity" by **Chunming Zhang, Yi Chai, Xiao Guo, Muhong Gao, David Devilbiss**, and **Zhengjun Zhang** looks at identifying the network structure of a neuron ensemble beyond the standard measure of pairwise correlations, which is critical for understanding how information is transferred within such a neural population. The spike train data poses a significant challenge to conventional statistical methods due to not only the

complexity, massive size, and large scale, but also the high dimensionality. The authors proposed a novel structural information enhanced (SIE) regularization method for estimating the conditional intensities under the generalized linear model (GLM) framework to better capture the functional connectivity among neurons.

The last four papers cover specific Big Data problems. "Measuring Influence of Users in Twitter Ecosystems Using a Counting Process Modeling Framework" by **Donggeng Xia, Shawn Mankad**, and **George Michailidis** focuses on analyzing data extracted from social media platforms such as Twitter that are both large in scale and complex in nature, since they contain both unstructured text and structured data such as time stamps and interactions between users. The authors develop a modeling framework using multivariate interacting counting processes to capture the detailed actions users undertake on such platforms, namely posting original content and reposting and/or mentioning other users' postings.

Profile monitoring is an important problem in manufacturing industries. "Discovering the Nature of Variation in Nonlinear Profile Data" by **Zhenyu Shi, Daniel W. Apley**, and **George C. Runger** proposes a method for exploratory analysis of a sample of profiles to discover the nature of any profile-to-profile variation present over the sample.

"Variable Selection in a Log-Linear Birnbaum-Saunders regression model for High-Dimensional Survival Data via the Elastic-Net and Stochastic EM" by **Yukun Zhang, Xuewen Lu**, and **Anthony F. Desmond** proposes a simultaneous parameter estimation and variable selection procedure in a log-linear Birnbaum-Saunders regression model for analyzing high-dimensional survival data.

"Online Updating of Statistical Inference in the Big Data Setting" by **Elizabeth D. Schifano**, **Jing Wu, Chun Wang, Jun Yan,** and **Ming-Hui Chen** develops iterative estimating algorithms and statistical inferences for linear models and estimate equations for analyzing Big Data arising from online analytical processing, where large amounts of data arrive in streams and require a fast analysis without storage/access to the historical data. ■

MASTER'S NOTEBOOK

Lessons from 34 Years as a Master's Statistician in Industry

Nestor Rohowsky provides statistical support to the pharmaceutical, biotechnology, and medical device industries as an employee and, as of 1995, as the owner of Integrated Data Consultation Services, Inc. He earned a bachelor's in mathematics from New York University and a master's in mathematics—with a statistics concentration—from Hunter College of the City University of New York.

May 2016 marked the completion of 34 full years of providing statistical services to the pharmaceutical, biotechnology, and medical device industries. During the first 14 years, I moved up the corporate "food chain" from an entry-level position, where I did those jobs my more senior colleagues passed on to me, to a director-level position, where I had high-level decision making interactions with other groups within and outside the company and personnel assignment and budgetary responsibilities within my department. Then, a little more than 20 years ago—being fed up with mergers, acquisitions, and general "right-sizing" involving companies I worked for—I decided to try my hand at independent statistical consulting. I consider it one of the best career moves I have ever made.

You may be thinking, "Big deal! A lot of people have done that!" Yes, a lot of people have had career paths similar to mine, but there's a difference between me and the majority of others. I did it without benefit of a doctoral degree. No PhD, no DrPH, no ScD. My highest level of completed formal education is a master's degree. Was that how I planned things to happen? Absolutely not.

After earning my master's, I spent a summer interning at a large New Jersey-headquartered pharmaceutical company, after which I entered a PhD program in statistics. It was there that the events that ultimately shaped my future occurred. First, I discovered that no matter how hard I studied, I could not fathom the intricacies of measure theory—a prerequisite for the degree at this university. Later, I lost some time recovering from a fairly major surgical procedure, and, last but not least, my wife let me know we were expecting a child (this, alas, did not actually happen until a couple of years later). This combination of events led me to conclude that I needed to find employment to pay bills and support my growing family. School would have to wait. I'd go back when things settled down.

It was then I landed the entry-level position. The company was a smallish unit of a large foreign-owned corporation. I spent more than three years there, going from doing what others didn't want to do to being lead statistician on several projects. I helped contribute to two successful new drug applications (NDAs), but toward the end of my stay, I wanted new challenges. I wasn't convinced I would find those with my employer, and so began looking to make a change.

I moved to another small unit of a different large company, where I had my first taste of management and decision making. I discovered I wasn't yet ready for the hard calls managers sometimes have to make, but loved the interaction with other departments. Another successful NDA was the result. I made several more moves within the industry and, at each one, I took it upon myself to learn skills traditionally not considered statistical in nature. I wanted to understand how other groups I interacted with—clinical, regulatory, data management, and sales and marketing—approached and resolved problems. I attended a number of professional meetings intended for clinical and/or regulatory specialists and insisted on participating in the planning and preparation of manuscripts and promotional materials.

In late summer of 1995, I began my own statistics consulting company. Almost all my colleagues told me it wouldn't last, because running such a business is difficult from the get-go and providing "expert" consultation without a doctorate was the kiss of death. It's now nearly 21 years later and I'm still in business, and, in fact, I'm thriving, whereas many others who have tried—the majority with doctoral degrees—have failed. Why? I think it's because I learned to listen to my clients, and I try to keep things as simple as I can without compromising statistical integrity. I'm also willing to do less "sexy" tasks (data entry comes to mind, ugh).

The biggest "coup" of my early consulting career was landing a 15-month project with Columbia Presbyterian Medical Center in New York City. During my interview, I asked why they didn't use statisticians from Columbia University—after all, they were two parts of a common institution. The answer stuck with me and guides me to this day: "We need someone who gets down to the nuts and bolts of a project; listens to us and asks good questions; and is interested in results, rather than constantly trying to

tweak things to satisfy their own professional curiosity." About a year after completing that project, they called asking me to help with something new. Unfortunately, I was busy with other clients by then and couldn't devote the time and energy necessary to do a proper job.

Why have I told you all this? The reason is simple. A master's-level statistician can make it and do well for him or herself. First, study hard. Try to land work related to your studies, either through your school or on your own, to gain real experience. Remember that reality and what's in textbooks are often different. Once in the workplace, observe what is going on around you. Sherlock Holmes was right when he said Dr. Watson and others could look at something without actually seeing what was there! Ask questions, keeping in mind there really are no stupid questions. Be curious about the process as a whole, not just your part in it.

In my case, I learned how to program and acquired the basics of data management. I went on monitoring trips with CRAs to witness how they interacted with physicians and nurses at study sites and when discussing data in source documents.

Attend professional meetings—not only the large JSM types, but also smaller, more focused ones—to keep current with new ideas and developments in your field. Take some risks. You'll fail some of the time (my second job was as a manager, a position I was woefully unprepared for), but if you learn from your failures and missteps, you'll likely succeed as well. Eventually, you may find that your successes outnumber failures, both in quantity and quality.

Finally, try your hand at public speaking (much harder than one may think) and performing nontraditional roles. I recently taught undergraduate university-level math and statistics courses to business majors. I also have presented at professional meetings geared toward clinical trials and will continue to do so. As for nontraditional activities, I have provided statistical support to PhD candidates—not in statistics, but in other disciplines in which statistics play a major role.

My advice to you is if you can, earn the doctorate. But if you can't, it doesn't have to be the end of the world. You can still be successful. You'll just need to be a little more creative and assertive, work a bit harder, and not take no for an answer! ■

STAT*tr@k*

Communicating Statistics to Nonstatisticians

Kim Love is the owner and lead collaborator at K. R. Love Quantitative Consulting and Collaboration. She earned her PhD from Virginia Tech in 2007 and, prior to starting her practice, served as the associate director of the University of Georgia Statistical Consulting Center. Visit her blog about starting her private practice at *www.krloveqcc.com/blog*.

One of the biggest challenges faced by any collaborative statistician is communicating statistical information to those with less knowledge of statistics. Many of us with a formal education in statistics receive extensive training in theory, methods, and application; however, even with a PhD in statistics, it is not uncommon to have taken one or no courses that focus on communicating this knowledge to those who can benefit from it. In other words, many of us leave school with little understanding of how to put our skills into effective practice.

What do our nonstatistician colleagues need from us to get the most out of our interactions? As posited by Janice Derr in her 1999 textbook *Statistical Consulting: A Guide to Effective Communication*, there are five dimensions of quality that nonstatisticians evaluate when collaborating with statisticians:

1. Availability of support
2. Responsiveness of support
3. Timeliness of support
4. Completeness of support
5. Pleasantness of support

Note that none of these dimensions directly incorporates correctness, technical savvy, or methodological awareness. It's not that those are unimportant; they are extremely important, and a collaborative statistician will not last long without solid abilities in those areas. It's simply that most nonstatisticians are unable to evaluate those aspects of a collaboration and have to make the assumption that their statistical collaborator possesses those skills.

The question then becomes, how can we improve our communication skills when working with nonstatisticians so they will understand and appreciate our expertise? The following recommendations are based on my personal experience and the advice of other statisticians in the consulting and collaboration community. I refer to nonstatisticians as "clients" in these recommendations, but that term is not limited to what one might view as a traditional consulting client; it could be a boss, a coworker, or even a friend who asks for help with a quantitative problem.

Focus on the client and project at hand, rather than general statistical concepts.

During my time at the University of Georgia Statistical Consulting Center, I supervised many students who were just learning to become collaborative statisticians. Across the board, when these students were initially challenged to explain a statistical method to a client, they provided equations full of Greek letters and other mathematical notation. While each client's needs should be evaluated individually, for many clients, this tends to add to their confusion about a method rather than mitigate it.

Here is my personal hierarchy, from greatest to least chance of success, of techniques to explain statistical methods to most clients:

1. Plain English

   Example: "Across your group of students, for every additional point a student scores on the entrance exam, the final achievement score increases by an average of about half a point."

2. Equations with your client's variables written out in words

   Example: "Estimated Average Achievement Score = 0.32 + (0.54 x Entrance Score)"

3. Equations with mathematical notation

   Examples: "$\hat{y} = 0.32 + 0.54x_1$"; "$y = \beta_0 + \beta_1 X_1 + \varepsilon$"

Well-labeled figures are always helpful when it comes to understanding, and should be used in tandem with these techniques when possible. Don't label figures using the cryptic, abbreviated variable names we often use as programming shortcuts; this is a barrier to a client who does not think like a programmer and who would need to continuously remind him or herself of the meaning of those labels.

Some addendums: Some clients require mathematical notation for their eventual research output, including those who are publishing in quantitative academic journals in their fields. Many of my clients require a combination of these methods to both understand the concepts and be able to provide a final product that meets the requirements of their stakeholders.

Some clients do have an interest in learning more general information about statistics. When a client asks a general question (e.g., What is power?), it is still helpful to explain it in a way that is specific to that client and his or her research. For example, for a client in agriculture, "Power is the probability that you will choose a sample of lettuce plants for your study that will result in a statistically significant difference between your two lettuce strains, assuming a difference exists." The wrong explanation would be "Power is the probability of rejecting the null hypothesis when the alternative hypothesis is true."

Actively improve your communication; communication skills can be learned.

Why do so many new statisticians respond to clients' difficulties in understanding with Greek letters and nonproject-specific explanations? The short answer is this is how most of us learned statistics. This notation and general conceptual discussion is useful in a classroom—it's a shorthand language we have in common so we can learn advanced concepts quickly. However, its usefulness diminishes greatly when working with someone who does not share our background in quantitative sciences. The good news is, contrary to what many of us believe (at least in practice), communication abilities can be practiced and learned.

One of the best ways to improve communication in client interactions is through the use of video review. Think of it as collecting data on client interactions (with the client's approval, of course). Set the recording device in a location where both the client and statistical collaborator are visible on the video, record the interaction, and then choose some or all of the video to watch—preferably with a trusted colleague. A video by Eric Vance from the Laboratory for Interdisciplinary Statistical Analysis at *http://bit.ly/29OK28X* is a helpful resource for implementing video coaching and feedback sessions.

One of the best ways to improve communication in client interactions is through the use of video review. Think of it as collecting data on client interactions.

Focus on the following during video review:

- **Was your communication effective?** Did you provide explanations that were appropriate for your client, and did he/she seem to respond to them? Did you ask the necessary questions when you didn't understand something? Listen to the verbal exchange, but also examine body language.
- **What was the relationship between you and the client like?** Was it collaborative or hierarchical? Was it constructive or combative?
- **Did you follow a proper structure for the interaction (see the next recommendation)?** Make a checklist.

Most importantly, leave your review session with one or two goals for future client interactions.

Statisticians also can improve communication outside of live client sessions. New collaborators can practice communication skills using role play scenarios, in which one person acts as a client and the other acts as a statistical collaborator (and these can be video reviewed as well). Workshops and continuing education opportunities to help improve communication throughout a statistician's career also are available at ASA conferences (such as the Conference on Statistical Practice or the Joint Statistical Meetings) and through other professional organizations.

Structure your interactions and their outcomes.

One of the keys to practicing communication effectively is to have a well-thought-out plan for your interactions. Several prominent collaborative statisticians have presented structures for interactions, including Derr (again in her textbook, *Statistical Consulting: A Guide to Effective Communication*) and Doug Zahn, who developed the POWER process

for client interactions (see *http://bit.ly/29DIuAG*). While it is probably more important to have a structure in the first place than to adopt a specific meeting structure, these structures do have a number of elements in common. There is an initial period to prepare for the meeting; terms for the interaction and a mutual agenda are agreed upon between the statistical collaborator and the client; there is a work session in which information is exchanged productively, with opportunities to question and enhance understanding on both sides of the table; and time is allowed at the end to review the interaction and agree on steps going forward.

While it will take some time and experience to implement a meeting structure smoothly and with appropriate flexibility, the benefits to the participants are well worth it. When there is a structure in place and the statistician no longer needs to concentrate on the meeting logistics, it becomes much easier to focus on communication and gauge effectiveness. Note that this idea can be extended to written reports and other outcomes outside of immediate interactions—having a general, flexible structure in place to organize statistical information is immensely helpful to making that information understandable.

MORE ONLINE
Visit the web version of this article at *http://stattrak.amstat.org* to add your comments about communication.

Gauge your clients' knowledge and communication needs on an individual basis.

Not all nonstatisticians who need statistical expertise are the same. That may sound like common sense, it is easy to begin treating all clients as if they are the same over time. While I have seen many attempts to categorize clients—some serious and some humorous—they fall short of describing the variety and nuances of nonstatisticians who seek statistical collaboration.

With respect to background knowledge, the simplest approach is to ask clients what kind of experience they have had with statistics. It's important to do this in a respectful manner (see my next recommendation), as immediately firing off a barrage of questions related to specific statistical techniques and courses can be intimidating. Instead, clarify that you are asking because you want to make sure you use appropriate vocabulary and provide proper explanations. Also, invite them to ask questions any time you are not being clear.

Communication needs are a bit subtler, and I've emphasized just a few points in the following bullets:

- **Atmosphere:** Some clients prefer an atmosphere that includes friendly conversation, while others prefer a more polished, down-to-business environment. I usually try to gauge this based on my client's demeanor after the introduction—is the discussion moving to the parking situation outside? Or is the client already placing material on the table and starting to tell me about some of the issues involved in his or her project?
- **Attitudes:** Unfortunately, some clients have preconceived negative perceptions about statistics; they have had little exposure and believe they have poor abilities. They may be nervous or even fearful about meeting with a statistician. Make an effort to be particularly patient when you recognize a client has an emotional reaction to working with statistics, and that client will be more receptive to what is being communicated.
- **Directness:** Some clients are very good about asking questions and guiding the direction of the meeting. Others are quieter and prefer to be invited to contribute. Never be afraid to ask a client if he or she understands something or is happy with the pacing and direction of the meeting.

Practice respect.

Respect takes many forms during client interactions. General politeness (such as showing up on time, remaining focused on your interaction, and not interrupting your client) is one way to demonstrate respect. Some of what I have already discussed in my other recommendations also go toward demonstrating respect and are more specific to collaboration: acknowledge your client as an individual; structure your interactions so they are efficient; and be patient with your clients' hang-ups when needed.

There is also an issue of professional respect, which cannot be understated—clients collaborate with us because they acknowledge we have statistical expertise they do not. It is important to recognize in return that our clients have expertise in their areas that we do not have. Just like clients do not become experts in statistics after an hour-long meeting, we do not become experts in their fields or areas of expertise during that meeting. Allow and encourage clients to contribute to the interaction. This not only results in open lines of communication and positive relationships, but even improves the technical aspects of statistical work, as it provides a more complete view of clients' research problems and objectives.

Mentor others.

One of the greatest contributions experienced statisticians can make to the field is to share their experiences. How do you explain complicated statistical procedures to nonstatisticians (think random effects, ordinal responses, and computationally intensive techniques)? How did you learn to expand your communication abilities? ■

eCOTS 2016 Draws 600

Participants of the regional eCOTS meeting at Florida Atlantic University

The third biennial Electronic Conference on Teaching Statistics (eCOTS), hosted by the Consortium for the Advancement of Undergraduate Statistics Education (CAUSE), took place online May 16–20 with the theme "Changing with Technology." With topics including teaching a flipped or blended course, teaching simulation-based inference, teaching with R/RStudio, teaching with technology, and teaching data science, the conference drew nearly 600 registrants.

The conference included an extensive program, consisting of the following:

- Keynote addresses by Andrew Gelman on introductory statistics and Michael Jordan on a data science class for first-year students
- Three two-hour workshops disseminating National Science Foundation–funded projects on passion-driven statistics, simulation-based inference, and flipping the introductory statistics course
- Five one-hour panel discussions about hot topics in statistics education related to technology, including an invited panel about teaching data science with Nick Horton, Jeff Leek, Deb Nolan, and Hunter Glanz and an invited panel about teaching with simulation-based inference with Nicola Justice, Robin Lock, Allan Rossman, and Chris Wild
- Eighteen half-hour breakout sessions
- Thirty-four virtual posters (five-minute videos) about a variety of topics
- Eight birds-of-a-feather small-group open discussions

eCOTS 2016 Organizers

Kari Lock Morgan, *chair*
Jackie Miller, *poster session chair*
Beth Chance
Monica Dabos
Jennifer Kaplan
Megan Mocko
Dennis Pearl
Randall Pruim

All conference material was recorded and is publicly available for free viewing at *www.causeweb.org/cause/ecots/ecots16/program*. The recordings from eCOTS 2014 and 2012 also are available online for free viewing.

The conference also included, for the first time, eight regional face-to-face mini-conferences happening at different locations throughout the country: Southern California, Southern Florida, Northern Kentucky, Middle Tennessee, Northern Florida, Central Ohio, Boston, and Northern Georgia. The companion conferences were designed to foster regional statistics education communities and leverage and integrate with the online programming of eCOTS. All reported successful meetings and new collaborations.

eCOTS occurs in May of even-numbered years—alternating with its face-to-face sister conference, USCOTS (the United States Conference on Teaching Statistics), in May of odd-numbered years. Like eCOTS, USCOTS focuses on undergraduate statistics education and is hosted by CAUSE. Save the date now for USCOTS 2017, which will take place at the Penn Stater Hotel and Conference Center in State College, Pennsylvania, May 18–20, with the theme "Show Me the Data!" ■

2016 POSTER AND PROJECT Competition Winners

The American Statistical Association is pleased to announce the winners of the 2016 Poster Competition and Project Competition. First-place winners received $300, a plaque, a plaque for their school, and grade-appropriate graphing calculators for the students and advisers provided by Texas Instruments. Second-place winners received $200 and a plaque; third-place winners received $100 and a plaque; and honorable mentions received plaques.

The poster and project competitions are directed by the ASA/NCTM Joint Committee on Curriculum in Statistics and Probability. The 2016 poster competition leader was Rodney Jee of Discover Financial Services. Daren Starnes of The Lawrenceville School served as the head project competition leader, with Nathan Kidwell of Dubuque Senior High School as associate project competition leader.

K–12 posters are due every year on April 1. Projects (written reports) for grades 7–12 are due every year on June 1. Visit *http://bit.ly/2a1DpkO* for details about the competitions—including previous winners, entry forms, instructional webinars, and the rubrics used for judging the posters and projects.

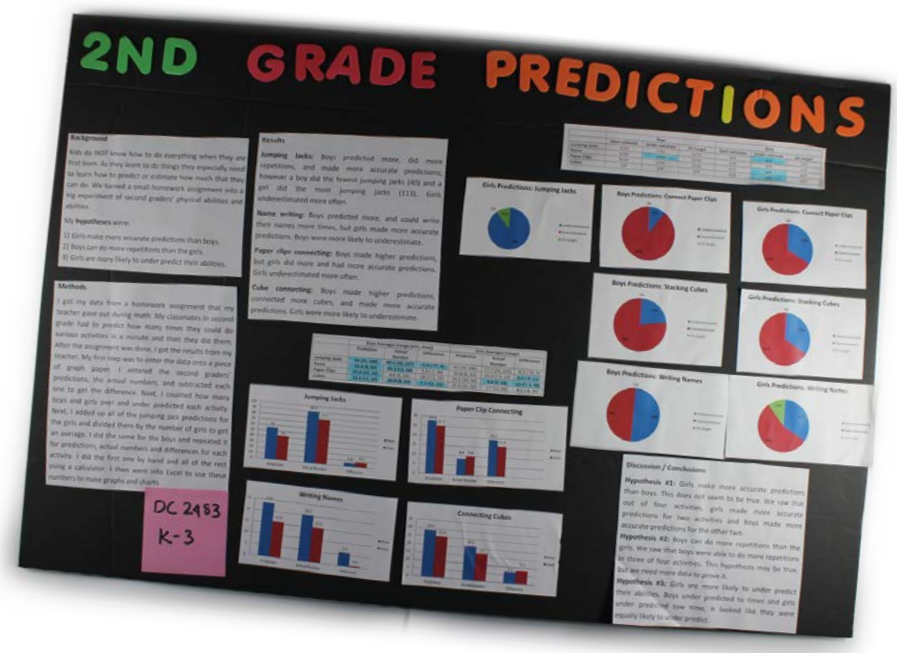
GET INVOLVED

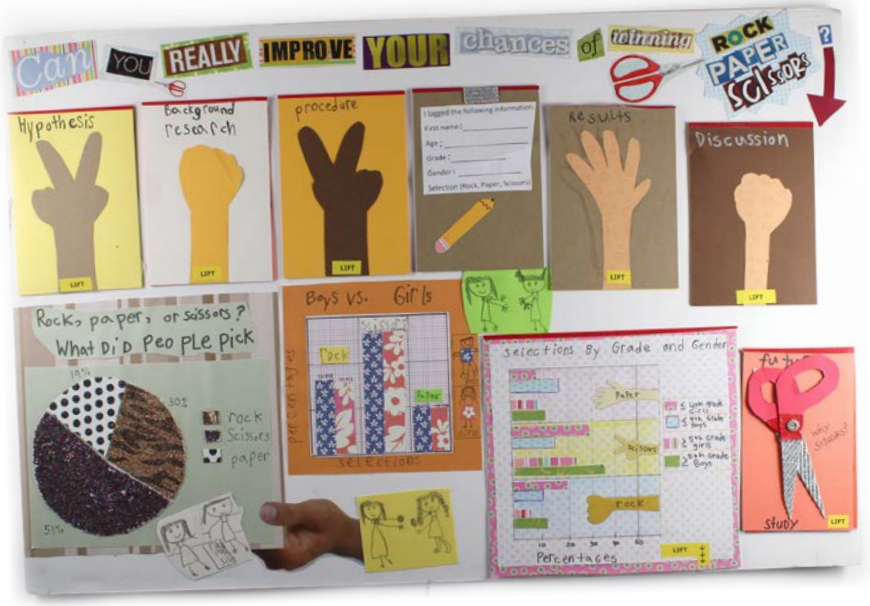
For information about how you can start a regional poster competition or mentor students in your area, see the article appearing in the July 2011 issue of *Amstat News* at *http://magazine.amstat.org/blog/2011/07/01/poster-comp-how-to*. You can download a flyer about the ASA poster and project competitions and other K–12 statistics education programs and resources to share with your local schools at the bottom of *www.amstat.org/education*. For additional information or questions regarding how to get involved in the poster or project competitions, please contact Rebecca Nichols at *rebecca@amstat.org*.

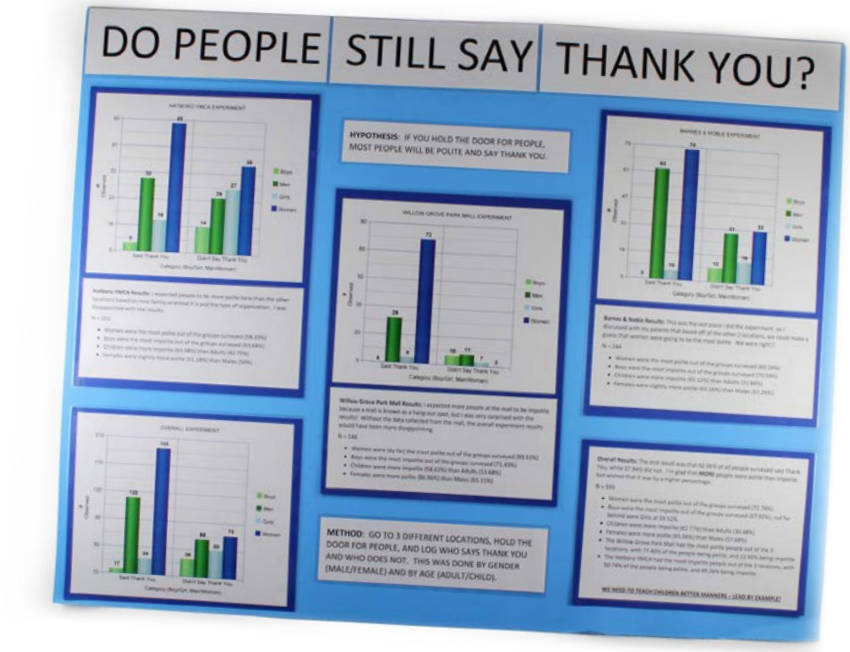
2016 Project Competition

Each year, the statistical project competition attracts a variety of submissions in which students from grades 7–12 conduct creative studies. The submission deadline for the project competition is June 1 to enable participation from high-school students who may have been preparing for the AP Statistics exam administered in mid-May. This deadline also makes it possible for teachers who might otherwise be busy at the AP Reading to assist with the competition judging. The project competition is especially useful for these students because it provides them with opportunities to apply all the statistical skills they have acquired throughout the school year to solve real-world problems of interest to them.

Results of the project competition and a list of the judges can be found in the online edition of *Amstat News* at *http://magazine.amstat.org*.





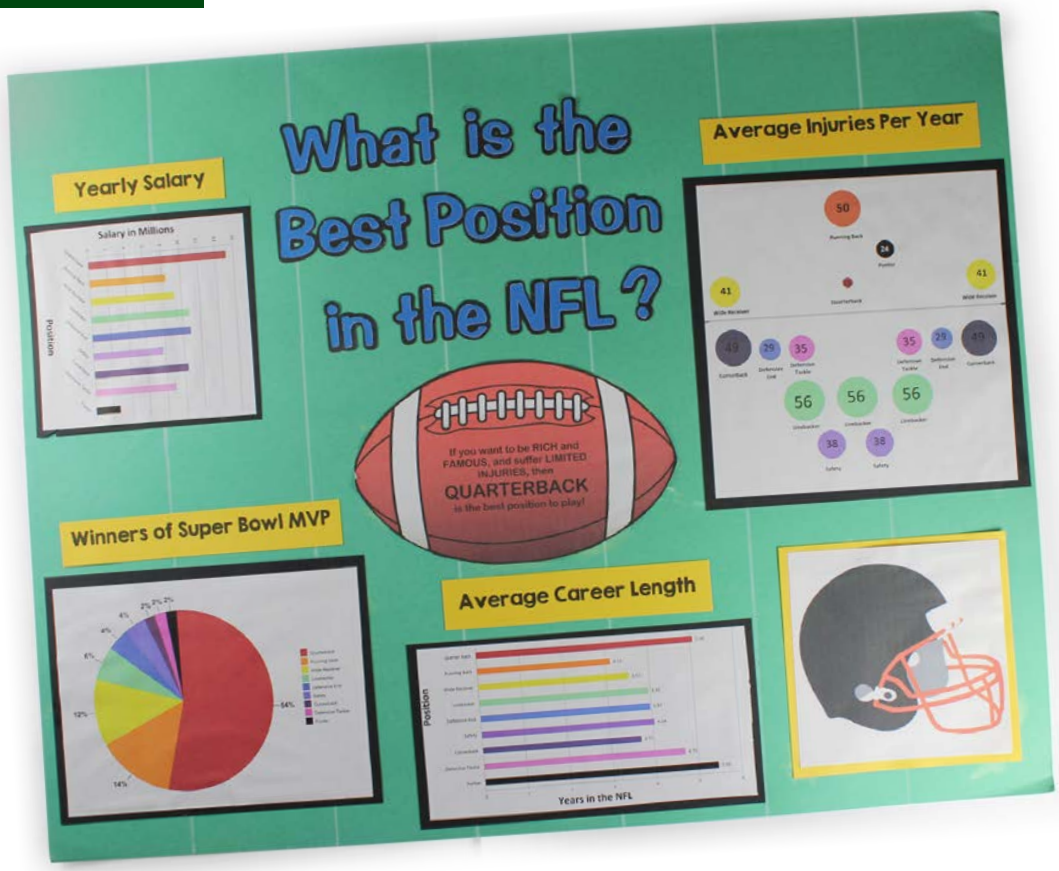


GRADES K–3

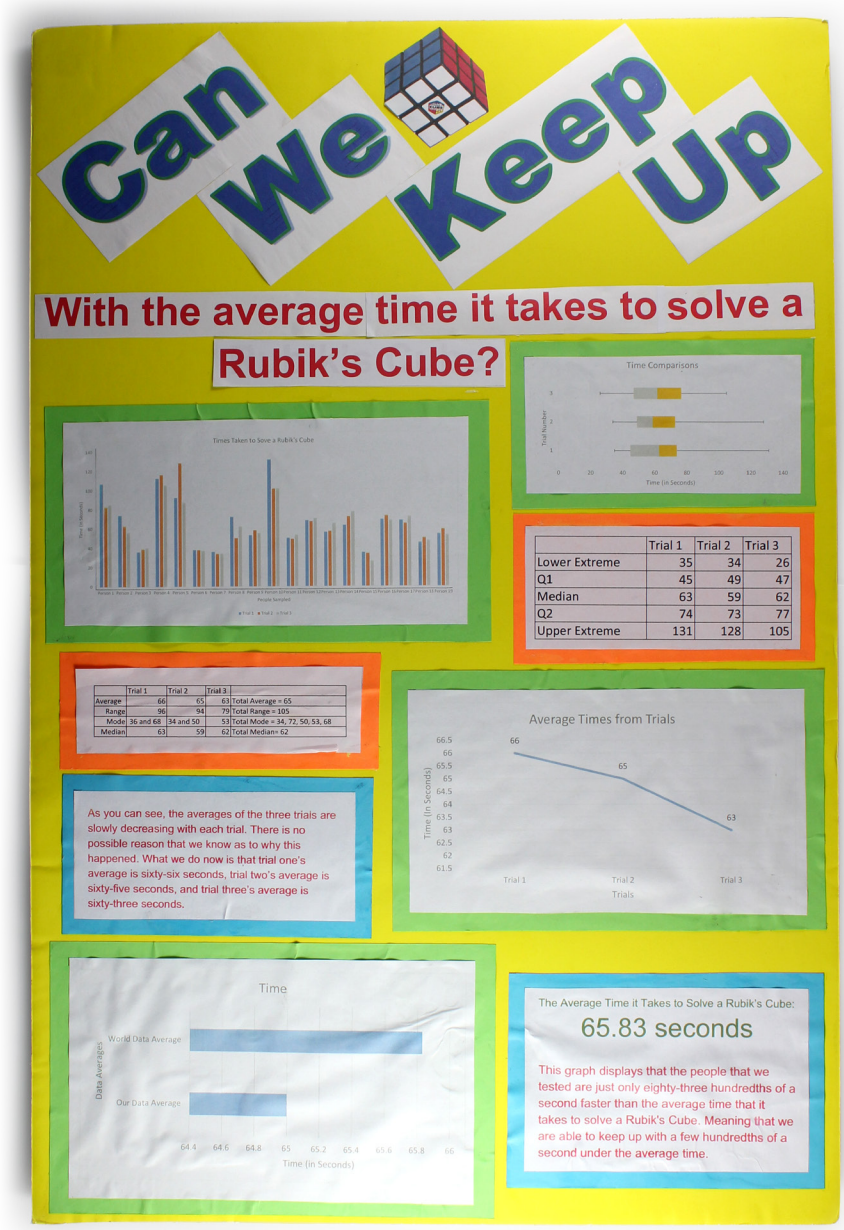
FIRST PLACE
Quinn Shields
2nd Grade Predictions
Carderock Springs Elementary
Bethesda, Maryland

SECOND PLACE
Madeline Krassner
Can You Really Improve Your Chances of Winning Rock-Paper-Scissors?
Overlook Elementary
Abington, Pennsylvania

THIRD PLACE
Aeryn Gibbons
Do People Still Say Thank You?
Roslyn Elementary
Roslyn, Pennsylvania


What is the Best Position in the NFL?
Yearly Salary
Salary in Millions
Position
Average Injuries Per Year
If you want to be RICH and FAMOUS, and suffer LIMITED INJURIES, then
QUARTERBACK
is the best position to play!
Winners of Super Bowl MVP
Average Career Length
Position
Years in the NFL


Are the Boston Celtics Hitting Homeruns?
Research Question
Morris Peterson (NBA player) said the 3-point shot is the homerun of basketball. I want to know if the Boston Celtics have improved their percentage of 3-point shots made compared to other field goals made since the 3-point shot was invented in 1979.
Data Resource
My resource for my data is databasketball.com. This website gives offensive and defensive statistics on the Western and Eastern Conferences. I looked at number of attempts and successes from 1980-2010 for the Boston Celtics for both the 3-point shots and all other field goal shots.
The success percentage was calculated as:
# of successes / # of attempts x 100%
The Boston Celtics attempts in field goals and three-point shots by year
The Boston Celtics field goal and three point success rates by year
NBA 3-point line
New NCAA
Old NCAA
23 ft. 9 in.
Baseline
Basket
Conclusions
The Boston Celtics took more 3-point shots and improved their 3-point success percentage over time. But the field goal success for other shots stayed the same even though the number of these attempts decreased.
I think the Celtics have been practicing their 3-point shots, and hitting homeruns! Danny Ainge and Larry Bird should be proud!

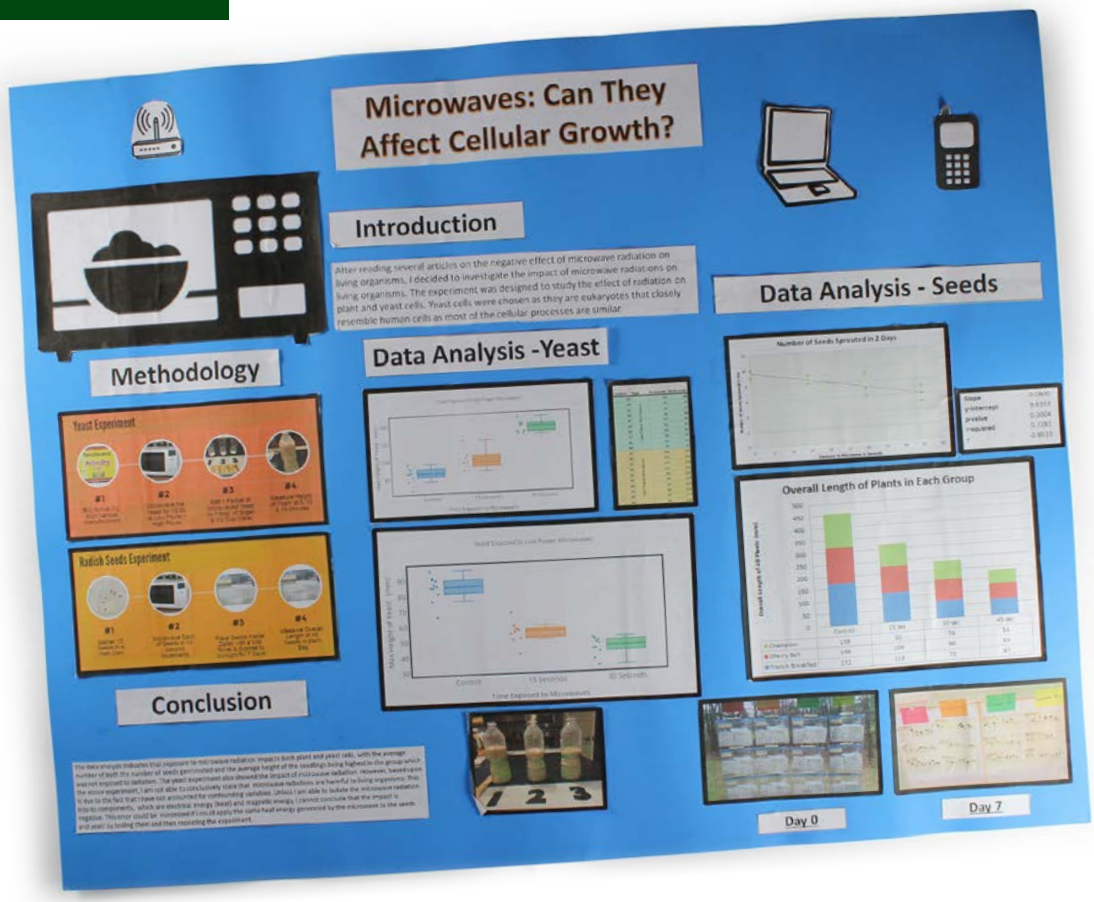


GRADES 4–6

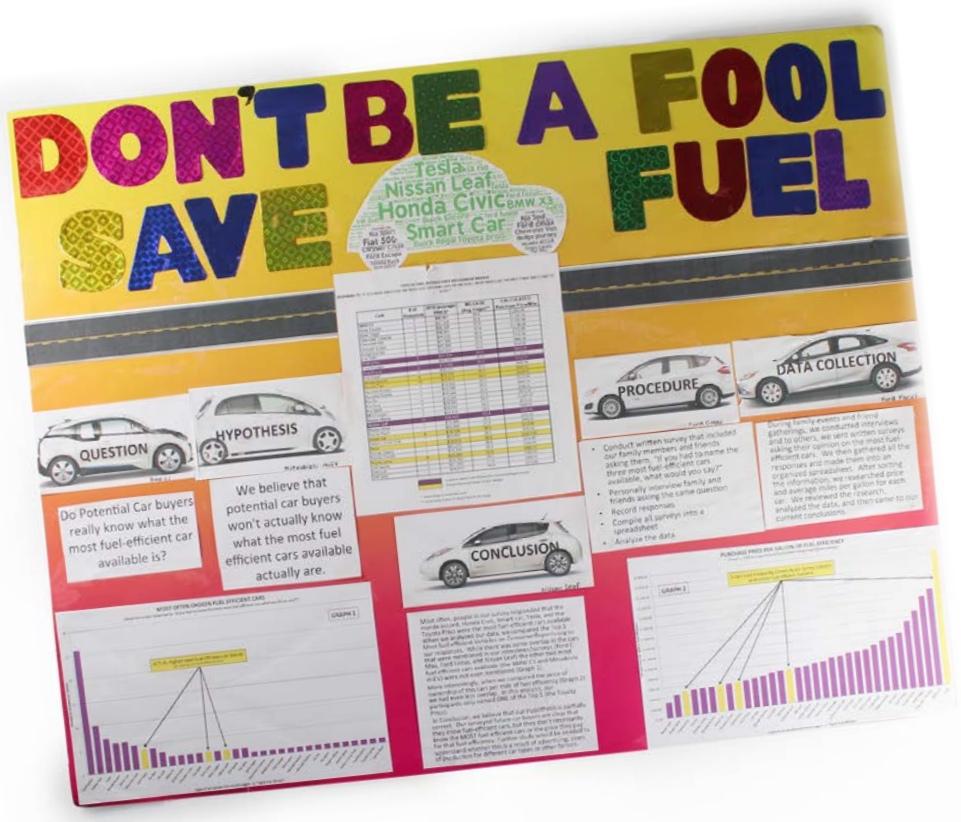
FIRST PLACE
Alexander Fiore
What Is the Best Position in the NFL?
Hyde Park Middle School
Las Vegas, Nevada

SECOND PLACE
Clyde Mauldin
Are the Boston Celtics Hitting Homeruns?
Mason Preparatory School
Charleston, South Carolina

THIRD PLACE
Aryaman Bisen, Fikreabe Getachew, Colin Jewell, and Aaron Luyen
Can We Keep Up with the Average Time It Takes to Solve a Rubik's Cube?
Hyde Park Middle School
Las Vegas, Nevada


Microwaves: Can They Affect Cellular Growth?
Introduction
Methodology
Data Analysis -Yeast
Data Analysis - Seeds
Conclusion
Yeast Experiment
Radish Seeds Experiment
Overall Length of Plants in Each Group
Day 0
Day 7


You're Welcome!
Chart 1
Percentage of Men and Women Who Said Thank You at McDonald's
Chart 2
Percentage of Men and Women Who Said Thank You at Cirella's
Chart 3
Total Percentage of People Who Said Thank You Individually or in a Group
Chart 4
Total Percentage of Children and Adults Who Said Thank You
Introduction
Conclusion
Cirella's
Italian Bistro & Sushi Bar
McDonald's

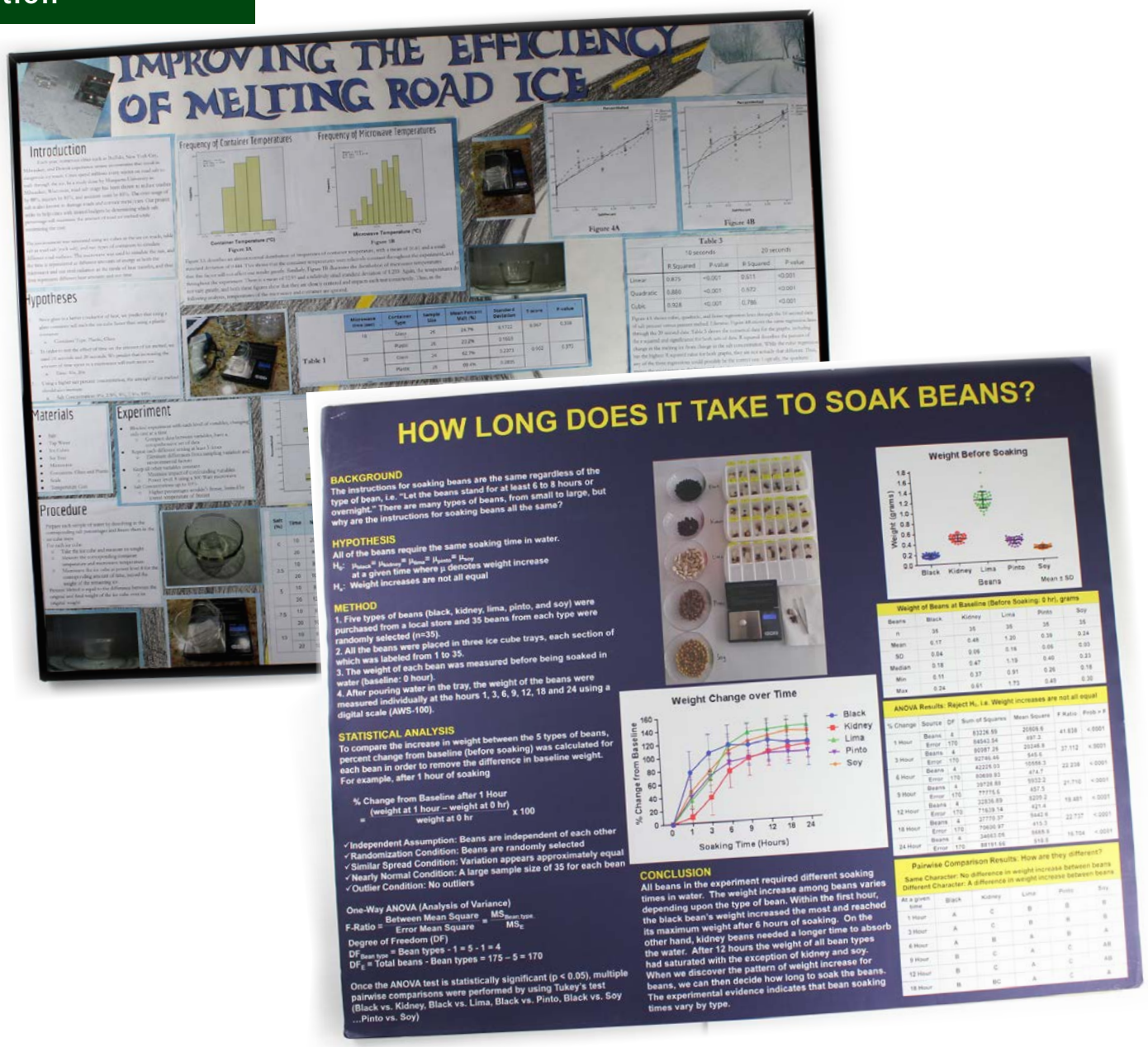


GRADES 7–9

FIRST PLACE
Dorris Dagama
Microwaves: Can They Affect Cellular Growth?
Forest Hills Central High School
Grand Rapids, Michigan

SECOND PLACE
Leo Wild
You're Welcome!
Half Hollow Hills High School East
Dix Hills, New York

THIRD PLACE
Timothy Buchanan and Garbriella Cilio
Don't Be a Fool, Save Fuel
Norwood Fontbonne Academy
Philadelpha, Pennsylvania



2016 Regional Poster Competition Leaders

Connecticut Chapter Statistical Poster Competition
Valerie Nazzaro, Wesleyan University
Jennifer McGinniss, Boehringer-Ingelheim
www.amstat.org/chapters/Connecticut/home/Poster/poster_index.htm

Kansas/Missouri Statistics Poster Contest
Ananda Jayawardhana, Pittsburg State University
www.pittstate.edu/department/math/stats_poster.dot

Michigan Statistics Poster Competition
Dan Frobish, Grand Valley State University
www.gvsu.edu/stat/mspc-homepage-22.htm

Nevada K–12 Statistics Poster Competition
David Thiel, Thrivent Financial
www.amstat.org/chapters/nevada

Ohio Statistics Poster Competition
Linda Quinn, Cleveland State University
Jerry Moreno, John Carroll University
www.bio.ri.ccf.org/ASA/poster.html

GRADES 10–12

FIRST PLACE
Amy Hung, Helen Li, Jessica Shu, and Selena Huang
Improving the Efficiency of Melting Road Ice
Lynbrook High School
San Jose, California

SECOND PLACE
Jin-Su Lee, Jin-Ju Lee, and Jin-Hee Lee
How Long Does It Take to Soak Beans?
Mission San Jose High School
Fremont, California

THIRD PLACE
Ethan Nguyen and Julia Nguyen
What Is College Worth?
Palo Verde High School
Las Vegas, Nevada

Pennsylvania Statistics Poster Competition
Pete Skoner, Saint Francis University Science Outreach Center
http://francis.edu/pa-statistics-poster-competition

Pullman, Washington Statistics Poster Competition
Dean Johnson, Washington State University
dean_johnson@wsu.edu

Washington Statistical Society Poster Competition (DC Metro Area)
Barnali Das, Lewin Group
www.amstat.org/education/posterprojects

ASA National Poster Competition
Leader: **Rodney Jee, Discover Financial Services**

Contact: **Rebecca Nichols,** ASA Director of Education,
rebecca@amstat.org
www.amstat.org/education/posterprojects

Students outside the regional competition areas submit their posters directly to the ASA office. The posters are then separately judged by the Washington Statistical Society as part of the Other Region. The best posters from each region are sent to the national judging. Information about regional poster competitions and winners is available on the individual regional poster competition websites. ■

CALL FOR NOMINATIONS

C. R. and Bhargavi Rao Prize

Penn State Department of Statistics awards the prize to a recognized leader in the field

C.R. Rao

Rao Prize Honorees

2003: Bradley Efron, Stanford University

2005: Jayaram Sethuraman, Florida State University (emeritus)

2007: Lawrence D. Brown, University of Pennsylvania

2009: Peter J. Bickel, University of California at Berkeley

2011: James O. Berger, Duke University

2013: Herman Chernoff, Harvard University (emeritus) and MIT

2015: Sir David Cox, University of Oxford

Members of the Rao Prize Committee are accepting nominations for the C. R. and Bhargavi Rao Prize for Outstanding Research in Statistics. The prize, awarded by the Penn State Department of Statistics, was established to recognize outstanding and influential innovations in the theory and practice of mathematical statistics, international leadership in directing statistical research, and pioneering contributions by a recognized leader in the field of statistics.

The Rao Prize is awarded in odd-numbered years to an individual working in the United States. The honoree receives a medal, cash prize, and invitation to visit Penn State to give a talk.

Nominations should include a letter describing the nominee's outstanding contributions to leadership and research in statistics, a current curriculum vita, and two supporting letters.

Submissions are due December 31 and should be emailed to *depthead@stat.psu.edu* or sent to Rao Prize Selection Committee Chair, 326 Thomas Building, Penn State University, University Park, PA 16802-2111.

C. R. Rao held the Eberly Chair in Statistics at Penn State from 1988–2001. He now serves as Holder Emeritus of the Eberly Chair in Statistics. He was the founding director of the Center for Multivariate Analysis. A President's National Medal of Science Laureate, Rao is recognized worldwide as a pioneer of modern statistical theory and one of the world's top five statisticians, with multifaceted distinctions as a mathematician, researcher, scientist, and teacher. His contributions to mathematics and statistical theory and applications have become part of undergraduate and graduate courses in statistics, econometrics, and electrical engineering at universities throughout the world.

For more information about the prize, visit *http://stat.psu.edu/rao-prize*. ■

Snee

Ronald Snee was honored with the American Society for Quality's (ASQ) Distinguished Service Medal, the highest distinction for contributions to the quality profession from the ASQ. He was recognized on May 15 in Milwaukee, Wisconsin, prior to ASQ's World Conference on Quality and Improvement.

The Distinguished Service Medal honors the lifetime contribution of any person who has been recognized as a long-term enabler, catalyst, or prime mover in the quality movement. Snee was awarded the medal for "creation of advanced experimental methodologies in statistical analysis and graphical representation of data, developing the concepts of statistical thinking and engineering for application as approaches to structured process development and analytical problem solving, and sustained contributions to the advancement of the body of knowledge for both business excellence and Lean Six Sigma initiatives."

Snee is founder and president of Snee Associates LLC, a firm dedicated to the successful implementation of process and organizational improvement initiatives. Prior to entering the consulting field, he worked at DuPont for 24 years in a variety of assignments including pharmaceuticals, statistical studies, manager of statistical software and engineering consultants, and process improvement. Snee also serves as adjunct professor in the pharmaceutical programs at Temple and Rutgers universities.

Snee earned his BA from Washington and Jefferson College and MS and PhD degrees from Rutgers University. He is a fellow of the ASQ and American Statistical Association and has been honored with the ASA's Deming Lecture and Dixon Statistical Consulting awards. Snee is an academician in the International Academy for Quality and has been awarded ASQ's Shewhart and Grant medals, as well as numerous other awards and honors. He is a frequent speaker and has published five books and more than 280 papers in the fields of statistics, quality, performance improvement, and management.

For more information about the ASQ's Distinguished Service Medal, visit *http://bit.ly/29Balf1*. ■

The 2016 recipients of the Roger Herriot Award for Innovation in Federal Statistics are **Thesia I. Garner** and **Kathleen S. Short.** They are being recognized for their important and extensive work in developing and refining alternative measures of poverty used to better understand the nature and scope of poverty in America.

Garner and Short have been calculating improved poverty measurements together for more than 20 years. Since 1995, Garner—working for the Bureau of Labor Statistics—and Short—working for the U.S. Census Bureau—have conducted joint research to develop new expenditure-based poverty calculations. To measure expenditures necessary to meet an economic threshold, Garner used BLS Consumer Expenditure Survey data. To measure the amount of income necessary to purchase those expenditures, Short used data from the Current Population Survey Annual Social and Economic Supplement. Their first set of thresholds and poverty statistics for 1990 through 1995 were published in the March 1998 *Monthly Labor Review*. Over the years, they have continued to research and implement additional components and improvements to their alternative methods and measures.

Their Supplemental Poverty Measure (SPM), though not intended to replace the official poverty measure, provides answers to essential policy questions about the impact government programs have on reducing poverty rates. The SPM is based on a more comprehensive set of expenses that show how the cost of purchasing basic needs contributes to poverty. Instead of being based on food costs in the 1950s as a budget share adjusted by the CPI inflation rate (as the official rate does), the SPM includes the costs of food, clothing, shelter, and utilities and adjusts for regional differences. Instead of using gross untaxed income, the measure adjusts for paid taxes, work and child care expenses, the Earned Income Tax Credit, cash-aid welfare, SNAP food assistance, and near-cash support.

Time series data from 2005 to 2010 measuring SPM were first published by Garner and Short in 2011, with annual results being published every year since. In 2015, the results were published at the same time as the official U.S. poverty measure. These most recent results of the SPM were published in September 2015 and can be found at *http://bit.ly/1OYjBsj*.

Roger Herriot was the associate commissioner of statistical standards and methodology at the National Center for Education Statistics (NCES) when he died in 1994. Soon after his death, the Social Statistics and Government Statistics sections of the American Statistical Association, along with the Washington Statistical Society, established the award, which is intended to recognize individuals or teams who, like Herriot, develop unique and innovative approaches to the solution of statistical problems in federal data-collection programs. ■

MORE ONLINE
View the names of previous winners and read more about Herriot's life at *http://bit.ly/29GEApl.*

Call for Authors
ASA-CRC Series on Statistical Reasoning In Science and Society

Co-published by the American Statistical Association and Chapman & Hall/CRC Press

There is a growing recognition of the importance of statistical and probabilistic reasoning across many different aspects of everyday life. This is the case now more than ever in this era of data science where the volume, availability, and types of data have increased significantly. It is crucially important that statistical thinking is introduced to students early in their education. Professionals in nearly every field encounter data throughout their working lives, and the ability to reason statistically allows them to make data-driven decisions. For members of the general public, learning how to reason statistically enables them to better understand risk, make decisions in the face of uncertainty, and become more informed citizens.

This exciting new series features:

- A forum for books highlighting the important role of statistics in many fields
- Concepts presented while assuming little or no background in statistics and including few or no equations
- Short and inexpensive books of 100–150 pages that can be written and read in areasonable amount of time
- A broad audience including professionals across many fields, the general public and introductory statistics courses in high schools and many college departments.
- Topics involve statistics in wide-ranging aspects of professional and everyday life, including the media, science, health, society, politics, law, education, sports, finance, climate, and security
- Combined marketing of ASA and CRC

Series Editors

Nicholas Fisher, University of Sydney

Nicholas Horton, Amherst College

Deborah Nolan, University of California, Berkeley

Regina Nuzzo, Gallaudet University

David Spiegelhalter, University of Cambridge

For more details visit: www.crcpress.com/go/asacrc

sectionnews

Physical and Engineering Sciences

Greg Piepel, Pacific Northwest National Laboratory

Funds are available through the SPES Marquardt Memorial Speakers Program for applied statisticians to visit colleges and universities to give a seminar and meet with students and professors. SPES reimburses the host institution for up to $1,000 (previously $500) to cover the expenses of the speaker's visit. The speaker provides information to students about what an applied statistician does; how applied statisticians solve problems in science, engineering, technology, and business; and the nontechnical skills required to be a successful applied statistician.

The Marquardt Industrial Speakers Program was established by SPES in the early 1990s to encourage careers in applied statistics. If you are part of an institution interested in having a speaker or a SPES member interested in being on the speakers list (or working directly with a local institution to set up a visit), contact Greg Piepel at *greg.piepel@pnnl.gov* or (509)-375-6911. ■

Q&P, SPES Invite All to 60th Annual Fall Technical Conference

Members of the Quality and Productivity Section (Q&P) and Section on Physical and Engineering Sciences (SPES) invite you to attend this year's Fall Technical Conference in Minneapolis, Minnesota, October 6–7. The conference is the premier forum in which to discuss topics at the interface of statistics and quality, leading to a more effective use of statistics to improve quality.

The theme for this year is "Statistics and Quality: Twin Pillars of Excellence." The program includes a range of talks about subjects such as experimental design, reliability analysis, Big Data, statistical process control, and industrial case studies. Highlights include a plenary address by Lynne Hare from Statistical Strategies LLC, the W. J. Youden Memorial Address by Joanne Wendleberger from Los Alamos National Laboratory, and a talk by ASA President Jessica Utts.

Short courses will take place October 5. Q&P is sponsoring a course titled "Design and Analysis of Experiments with R," which will be taught by John Lawson from Brigham Young University. SPES is sponsoring a course titled "Methods for Designing and Analyzing Mixture Experiments," which will be taught by Greg Piepel from MIXSOFT.

For more information and to view the complete program, visit *http://bit.ly/29VIoAV.*

Government Statistics

The Government Statistics Section and Washington Statistical Society are sponsoring the Seasonal Adjustment Practitioners Workshop (SAPW 2016) taking place November 4 at the Bureau of Labor Statistics in Washington, DC. There is no cost to attend, but attendees must register at Eventbrite (*www.eventbrite.com*) by searching for the name of the workshop.

This workshop will provide an opportunity for those who work with official statistics and seasonal adjustment to present interesting issues, problems, and applied research to a knowledgeable audience. There will be a plenary session in the morning with four speakers and a discussant, followed by contributed sessions in the afternoon.

The organizers are soliciting abstracts for presentations of interest to practitioners of seasonal adjustment. The abstracts should be fewer than 100 words in length for a presentation of about 15–20 minutes. Possible topics include the following:

- Seasonal adjustment methodology (model-based seasonal adjustment, high-frequency data, etc.)
- Issues in seasonal adjustment production (residual seasonality, revisions policy, etc.)
- Calendar effect estimation and adjustments (trading and working day, moving holidays)
- Seasonal adjustment software
- Seasonal adjustment diagnostics

Papers may cover any subject relating to seasonal adjustment methodology or the process of seasonally adjusting series at a statistical agency. Send your abstract of 100 or fewer words to *esmd.seasonal.workshop@census.gov* by August 15.

Even if you do not plan to present a paper, register for the conference and join the conversation. SAPW 2016 is organized by the U.S. Census Bureau and Bureau of Labor Statistics. ■

Professional Opportunity listings may not exceed 65 words, plus equal opportunity information. The deadline for their receipt is the 20th of the month two months prior to when the ad is to be published (e.g., May 20 for the July issue). Ads will be published in the next available issue following receipt.

Listings are shown alphabetically by state, followed by international listings. Vacancy listings may include the institutional name and address or be identified by number, as desired.

Professional Opportunities vacancies also will be published on the ASA's website (*www.amstat.org*). Vacancy listings will appear on the website for the entire calendar month. Ads may not be placed for publication in the magazine only; all ads will be published both electronically and in print.

Rates: $320 for nonprofit organizations (with proof of nonprofit status), $475 for all others. Member discounts are not given. For display and online advertising rates, go to *www.amstat.org/ads.*

Listings will be invoiced following publication. All payments should be made to the American Statistical Association. All material should be sent to *Amstat News*, 732 North Washington Street, Alexandria, VA 22314-1943; fax (703) 684-2036; email *advertise@amstat.org*.

Employers are expected to acknowledge all responses resulting from publication of their ads. Personnel advertising is accepted with the understanding that the advertiser does not discriminate among applicants on the basis of race, sex, religion, age, color, national origin, handicap, or sexual orientation.

Also, look for job ads on the ASA website at *www.amstat.org/jobweb.*

Indiana

■ Lilly is looking for an Industrial Statistician to join a top talent team! This is an exciting opportunity to collaborate with scientist and engineers in our large molecule development group and play a hands-on role in developing the medicines. Please view the job details at the link below and apply to begin a career at a company that changes lives! Job description/application: *http://bit.ly/27187BR* EOE.

Massachusetts

■ The MGH Institute of Health Professions, a graduate school in Boston affiliated with Partners HealthCare, invites applications for a full-time faculty position as assistant/associate professor of quantitative methods. The successful candidate will have expertise in methods for human subjects studies with competence in both biomedical and

Perelman School of Medicine
University of Pennsylvania

University of Pennsylvania Perelman School of Medicine
Biostatistics - Associate or Full Professor Tenure Track

The Department of Biostatistics and Epidemiology at the Perelman School of Medicine at the University of Pennsylvania seeks candidates for several Associate or Full Professor positions in the tenure track. Applicants must have a Ph.D. or equivalent degree.

A position requiring expertise in high-dimensional statistical methods for big data in health services research, such as genomics and/or electronic health records, is among the available positions.

Applicants will focus primarily on the development of innovative leading-edge statistical methodology with secondary emphasis on collaborative research projects within the Perelman School of Medicine. A demonstrated track record as the principal investigator of methodological research supported by extramural grant funding is required. There is a rich mix of ongoing biomedical research projects in the Perelman School of Medicine to provide motivation and opportunities for the development of novel statistical methods on wide ranging topics.

Candidates are expected to have a strong commitment to teaching and must demonstrate outstanding research productivity. Primary teaching responsibilities include participation in Penn's Center for Clinical Epidemiology and Biostatistics academic programs.

The Graduate Group in Epidemiology and Biostatistics, jointly with the Department of Statistics in the Wharton School, offers degree programs leading to both the Doctor of Philosophy (PhD) and Master of Science (MS) in Biostatistics.

Review of applications will begin on September 30, 2016 and will continue to be accepted after this date, until the positions are filled. The expected start date is July 2017 or later.

We seek candidates who embrace and reflect diversity in the broadest sense.

The University of Pennsylvania is an EOE. Minorities/Women/Individuals with disabilities/Protected Veterans are encouraged to apply.

Apply for this position online at: https://www.med.upenn.edu/apps/faculty_ad/index.php/g303/d4362

psychosocial sciences. She or he will consult with faculty members from different professions and teach statistics. Application Link: *http://bit.ly/29AVVMa.* The MGH Institute of Health Professions is an equal opportunity employer and is committed to enhancing the diversity of its faculty and staff. We welcome nominations and applications from individuals who would bring diversity of experience, thought and practice to the Institute's research, teaching and clinical missions. Applications from protected veterans and individuals with disabilities are strongly encouraged.

Texas

■ The University of Texas at El Paso seeks a statistical bioinformatician with expertise in big data analysis for a tenure-track Assistant professor position in the department of mathematical sciences. Successful candidates will develop productive research programs, mentor and teach at both the undergraduate and graduate levels, and collaborate with researchers in UTEP0x2019s NIH-funded Border Biomedical Research Center. Visit *http://www.utep.edu/employment* to view complete advertisement. EOE. ■

BIOSTATISTICS FACULTY
Memphis, TN

The Department of Biostatistics at St. Jude Children's Research Hospital (www.stjude.org/biostatistics) invites applications for three faculty positions at the Assistant or Associate Member (Professor) level depending upon qualifications. Candidates must have a PhD in Biostatistics or Statistics and a record of peer-reviewed publications showing evidence of (for Assistant Member, a potential of) productive methodological research.

For the 1st position, preference will be given to candidates with experience and statistical research interests in designing and conducting clinical trials and a commitment to collaborative research with clinical and laboratory investigators. Experience and/or interest in Systems Biology, Statistical Genomics or Bioinformatics are also highly desirable.

For the 2nd position, preference will be given to candidates with statistical research interests in Survival Analysis, Longitudinal Analysis or Multivariate Analysis and a commitment to collaborative research with clinical investigators. Experience in designing and conducting epidemiological studies is highly desirable.

For the 3rd position, preference will be given to candidates with statistical research interests in developing qualitative and quantitative methods in health outcomes research, which includes but not limited to, epidemiological studies, psychosocial studies and studies in health economics. Experience in designing clinical trials/therapeutic interventions and developing infrastructure for conducting collaborative international trials, particularly with the countries that have limited resources, is highly desirable. Continued independent statistical research motivated by biomedical collaborations is expected of the successful applicant.

The Department staff includes thirteen faculty positions, two post-doctoral fellows, twenty-two master's level biostatisticians, eight computer scientists and support staff. Applicants must demonstrate excellent oral and written communications skills and be proficient in computing. Compensation is very competitive and commensurate with experience. **Send letter of interest, CV, and have three reference letters sent to: kumar.srivastava@stjude.org or Dr. Deo Kumar Srivastava, Interim Chair, Dept. of Biostatistics, St. Jude Children's Research Hospital, 262 Danny Thomas Place, Memphis, TN 38105-3678.**

EE/AAE

Westat | An Employee-Owned Research Corporation®

EOE

Statistical Career Opportunities with Westat

Westat is an employee-owned corporation headquartered in Rockville, Maryland. We provide statistical consulting and survey research to the agencies of the U.S. Government and to a broad range of business and institutional clients. With a strong technical and managerial staff and a long record of quality research, Westat is a leader in the statistical services field.

We are currently recruiting for the following positions:

Senior Survey Sampling Statistician—International Surveys Westat is seeking a senior survey sampling statistician for work on international surveys in developing nations. This position requires a master's degree in statistics or survey research coupled with seven (7) or more years in sample survey design, or a PhD in statistics or survey research and five (5) or more years in sample survey design. Candidates would benefit from knowing SAS, R, and other statistical software packages although candidates are not required to do programming.

Senior Manager, Statistical Computing Unit This position requires candidates to have a strong statistical or other quantitative background and at minimum a master's degree in computer science, statistics, math, physics, or a related data science coupled with at least ten (10) years of experience in statistical or other data-intensive computing. Five (5) years of supervisory experience is also required.

Senior Survey Sampling Statistician This position requires a master's degree in survey sampling, statistics, survey research, or a related field with twelve (12) or more years in sample survey work or a PhD in survey sampling, statistics, survey research, or a related field and ten (10) or more years in sample survey work. Candidates would benefit from knowing SAS, R and other statistical software packages although candidates are not required to do programming.

Biostatistician Westat is seeking a biostatistician or statistician with experience analyzing health data. A master's degree in biostatistics or statistics and five (5) years of experience or a PhD in biostatistics or statistics is required.

Senior Biostatistician This position requires a master's degree in biostatistics or statistics and ten (10) years of experience, or a PhD in biostatistics or statistics and five (5) years of experience. Experience leading research teams, and knowledge of SAS or R is also required.

Westat is an Equal Opportunity Employer and does not discriminate on the basis of race, creed, color, religion, sex, age, national origin, veteran status, disability, marital status, sexual orientation, citizen status, genetic information, gender identity, or any other protected status under applicable law. To apply, go to **www.westat.com/careers**.

www.westat.com

Cleveland Clinic
Taussig Cancer Institute

Taussig Cancer Institute
Faculty Senior Biostatistician

The Taussig Cancer Institute at the Cleveland Clinic is seeking a Ph. D. faculty member to lead the cancer statistics program. Candidates should have a research track record and experience as a consulting or lead biostatistician in the analysis of clinical and translational cancer research. The Cancer Institute is a very active research organization with many investigator initiated projects and a robust presence at national meetings.

Additional requirements include experience in designing and directing research, supervising analyses of both large and small clinical and translational/genomic datasets, and expertise in general statistical methods. Expectations include publication of collaborative and primary research papers using clinically-derived electronic datasets, including the electronic medical record.

The level of appointment will be commensurate with experience and there are no requirements to obtain salary support. Master's level statistician personnel support will be provided.

The Cleveland Clinic is a not-for-profit, multi-specialty academic medical center that integrates clinical and hospital care with research and education. The Taussig Cancer Institute has been rated by U.S. News & World Report Magazine among the best cancer institutes in the country and is a part of the NCI designated Case Comprehensive Cancer Center.

Interested candidates should send their CV with a cover letter to Dr. Mikkael Sekeres with a copy to Victoria Mineff.

Mikkael Sekeres, MD
Vice Chair, Clinical Research
Director, Leukemia Program
Taussig Cancer Institute
Cleveland Clinic
sekerem@ccf.org

Victoria Mineff
Manager, Physician Relations
Taussig Cancer Institute
Cleveland Clinic
mineffv@ccf.org

The Cleveland Clinic Foundation is an Equal Opportunity Employer

AMSTATNEWS
ADVERTISING DIRECTORY

Listed below are our display advertisements only. If you are looking for job-placement ads, please see the professional opportunities section. For more job listings or more information about advertising, please visit *www.amstat.org*.

misc. products and services

CRC Press ..... p. 42
Oregon State University ..... p. 17
SAMSI ..... p. 25

professional opportunities

Cleveland Clinic ..... p. 46
St. Jude Children's Research Hospital ..... p. 46
University of Pennsylvania Perelman School of Medicine ..... p. 44
U.S. Census Bureau ..... p. 45
Westat ..... p. 46

software

JMP Software ..... cover 4
Minitab ..... centerfold
NCSS ..... p. 27
Salford Systems ..... centerfold
SAS Institute Inc. ..... cover 3
StataCorp LP ..... cover 2

SOCIAL CHATTER

What our followers are saying online

FOLLOW US

community.amstat.org

www.facebook.com/AmstatNews

@AmstatNews

www.instagram.com/AmstatNews

TELL US

This month, we're asking: Which historical figures inspire you? Whether statistician or no, what qualities did they possess that you try to bring to your own work? We hope you'll send us your answers! Be sure to tag @AmstatNews.

In July, we asked our followers what they would be if they weren't statisticians. Here is what a few of our members shared:

Photo courtesy of NASA

FitzPleasure
@Fitzpleasure26
An astronaut.

Steve Parisi
@stevenjparisi
a cherry picker of data

Ozena Wielgoszewska
@MeBozena
…a wannabe statistician :)

Martí Casals
@CasalsTMarti
#Basketball #Analyst

Ryan Rosario
Tough question! Electrical engineer? Linguist? Air traffic controller? Any one of those I guess.

Vaishali Macwan
Engineer :(But glad I am not :D

Laura Thompson
Computer scientist. Or, just build systems.

Malope 'Lopez' Mampholo Farmer… can still do it part-time though

Elpida Mavromataki
Dancer

David Little Bored.

Pran Tonjokin Data Wizard

Nick Hepler Forest Ranger

Kel Zou Rock star!

Mandisa Mhlanga Madanha Baker.

Statistics

The release of SAS/STAT® 14.1 brings you more statistical techniques for your data analysis.

SAS/STAT 14.1 Highlights

Generalized additive models by penalized likelihood estimation. Apply this technique, which provides automatic model selection by optimizing model fitting criteria, to your large data problems.

Imputation for survey data. Employ single and multiple hot-deck and fully efficient fractional imputation methods to handle nonresponse.

Additional model selection methods. Use the LASSO method for selecting generalized linear models or the group LASSO method for selecting general linear models.

Classification and regression trees. Use familiar modeling syntax to specify trees and display results with ROC plots as well as tree diagrams.

Recent SAS/STAT Highlights

Analysis for spatial point patterns. Understand locations of random events, such as crimes or lightning strikes, and how other spatial factors influence event intensity.

Weighted GEE methods. Deal with dropouts in longitudinal studies with a method that produces unbiased estimates under the missing-at-random (MAR) assumption.

Item response models. Calibrate test items and evaluate respondents' abilities with item response models.

Proportional hazards regression models for interval-censored data. Apply these popular regression models in survival analysis when the data are interval-censored.

Bayesian choice models. Use Bayesian discrete choice models to model consumer decisions in choosing products or selection from multiple alternatives.

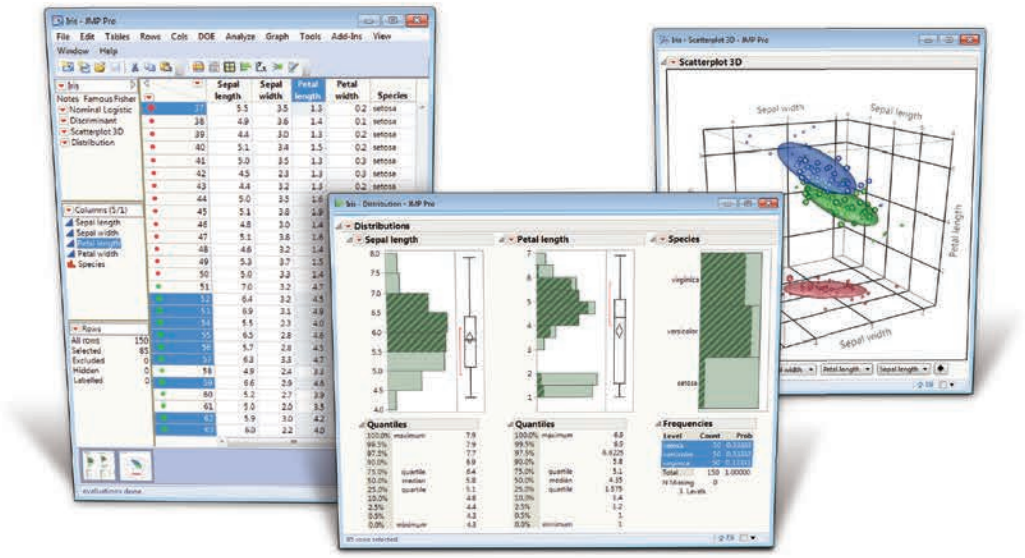
To learn more.

support.sas.com/statnewreleases

sas
THE POWER TO KNOW®

Join us at JSM for the
JMP Reception for Friends and Users
Sunday, July 31 4:30 - 6:30 p.m.

Please RSVP at: **jmp.com/jsm**



EXPLORE

Data analysis in flow

Introduced in 1989 with scientists and engineers in mind, JMP® software links powerful statistics to interactive graphics. It keeps data in flow, no matter whether it's small, tall or wide. A sampling of its capabilities:

- Regression, GLM and ANOVA
- Generalized Regression: Ridge, Lasso, Elastic Net*
- Mixed Models and Repeated Measures*
- Univariate and Bivariate Analysis
- Multivariate Analysis
- Data Mining Capabilities: Cross-Validation, Multi-Layer Neural Networks, Bootstrap Forests, Gradient-Boosted Decision Trees, Model Comparison*
- Cluster Analysis
- Nonlinear Modeling
- Graphical Application Development Environment
- Interactive Query Building for Connecting to Databases
- Automatic Recoding, Outlier Detection and Missing Value Imputation
- SAS,® R, MATLAB and Microsoft Excel Connections
- Time Series Analysis
- Design of Experiments
- Consumer and Market Research Methods
- Reliability and Survival Analysis
- Quality and Process Control
- Data Visualization, Mapping and Animated Graphs

*JMP Pro Only

Try JMP® 12 software! jmp.com/trial

jmp® Statistical Discovery.™ From SAS.

Available for Mac® and Windows

sas® THE POWER TO KNOW®